Hans-Peter Klenk

Culturomics
as tool in research and service in culture collections

Session 5: BRCs in the area of *omics*

ECCO XXXIV - *European Culture Collections as tools in research and biotechnology*

Institut Pasteur, Paris, 28 May 2015

BRCs in the era of *omics*

The English-language neologism **omics** refers to a field of study in biology ending in *-omics*, such as

- **Gen*omics***
- Prote*omics*
- Metabol*omics*
- Lipid*omics*
- Transcript*omics*
- Metage*nomics*

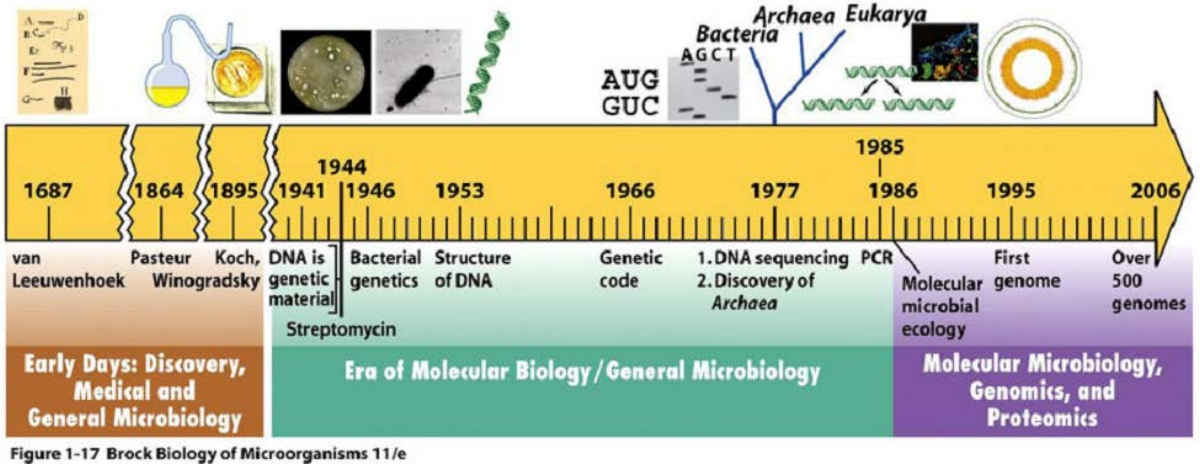


Figure 1-17 Brock Biology of Microorganisms 11/e

The related suffix *-ome* is used to address the objects of study of such fields, such as the genome, proteom, metabolom, respectively.

omics aims at the **collective high-throughput characterization and quantification of pools of biological molecules** that translate into the structure, function, and dynamics of an organism or organisms.

- What is '**Culturomics**'?
- and how does '*Culturomics*' fit to other *omics* fields of studies?

Microbial culturomics: paradigm shift in the human gut microbiome study

J.-C. Lagier[1,*], F. Armougom[1,*], M. Million[1], P. Hugon[1], I. Pagnier[1], C. Robert[1], F. Bittar[1], G. Fournous[1], G. Gimenez[1], M. Maraninchi[2], J.-F. Trape[3], E. V. Koonin[4], B. La Scola[1] and D. Raoult[1]

1) *Aix Marseille Université, URMITE, UM63, CNRS 7278, IRD 198, INSERM 1095,* 2) *Service de Nutrition, Maladies Métaboliques et Endocrinologie, UMR-INRA U1260, CHU de la Timone, Marseille, France,* 3) *IRD, UMR CNRS 7278-IRD 198, Route des Pères Maristes, Dakar, Sénégal and* 4) *National Centre for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA*

Seminal paper by Jean-Christoph Lagier and the team around Didier Raoult at Aix-Marseille Université, September 2012

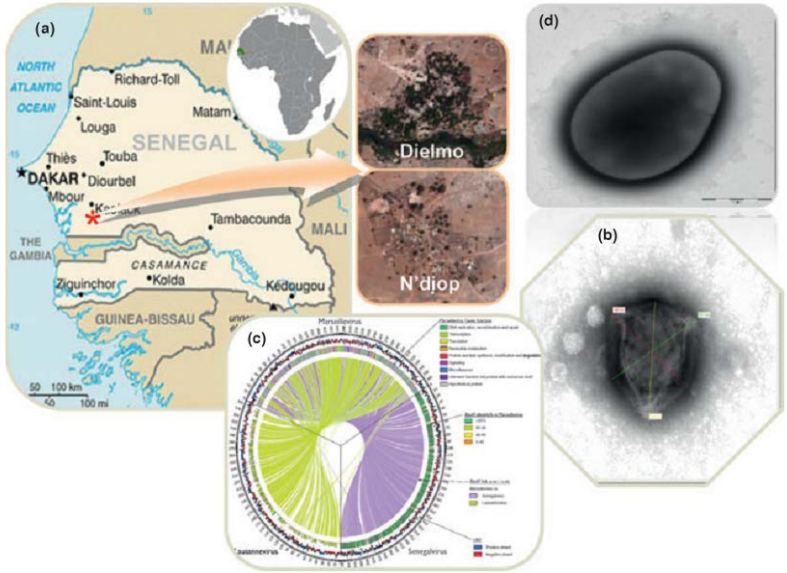


FIG. 1. The source of material for culturomics and the record-breaking virus and bacterium from the human gut. (a) The geographical locations of the Dielmo and N'diop villages (Sources: Wikitravel.org and Google Earth) from which the two African stool samples analysed in this work were obtained. (b) Electronmicrograph of the giant Senegalvirus, which was isolated from a stool sample of an individual from N'Diop. (c) Comparison of the Senegalvirus genome with the genomes of related giant viruses, Marseillevirus and Lausannevirus. (d) Electronmicrograph of *Microvirga massiliensis* (the bacterium with the largest genome ever isolated from humans), which was isolated from the Dielmo stool sample.

Lagier et al., Clin Microbiol Infect **18**:1185-94 (2012)

Culturomics background

Cultivation:

- Human microbiota first described *via* bacterial cultures *(considered outdated)*
- Metagenomics now more frequently used, but ignore minority bacterial populations *(dark matter)*
- Replicated natural environments to reduce the “great plate count anomality” *(difference between microscopic and culture counts)*

Analysis techniques:

- 16S rRNA sequences enabled accurate identification of novel species
- MALDI –TOF allows rapid high-throughput identification of rare and new species

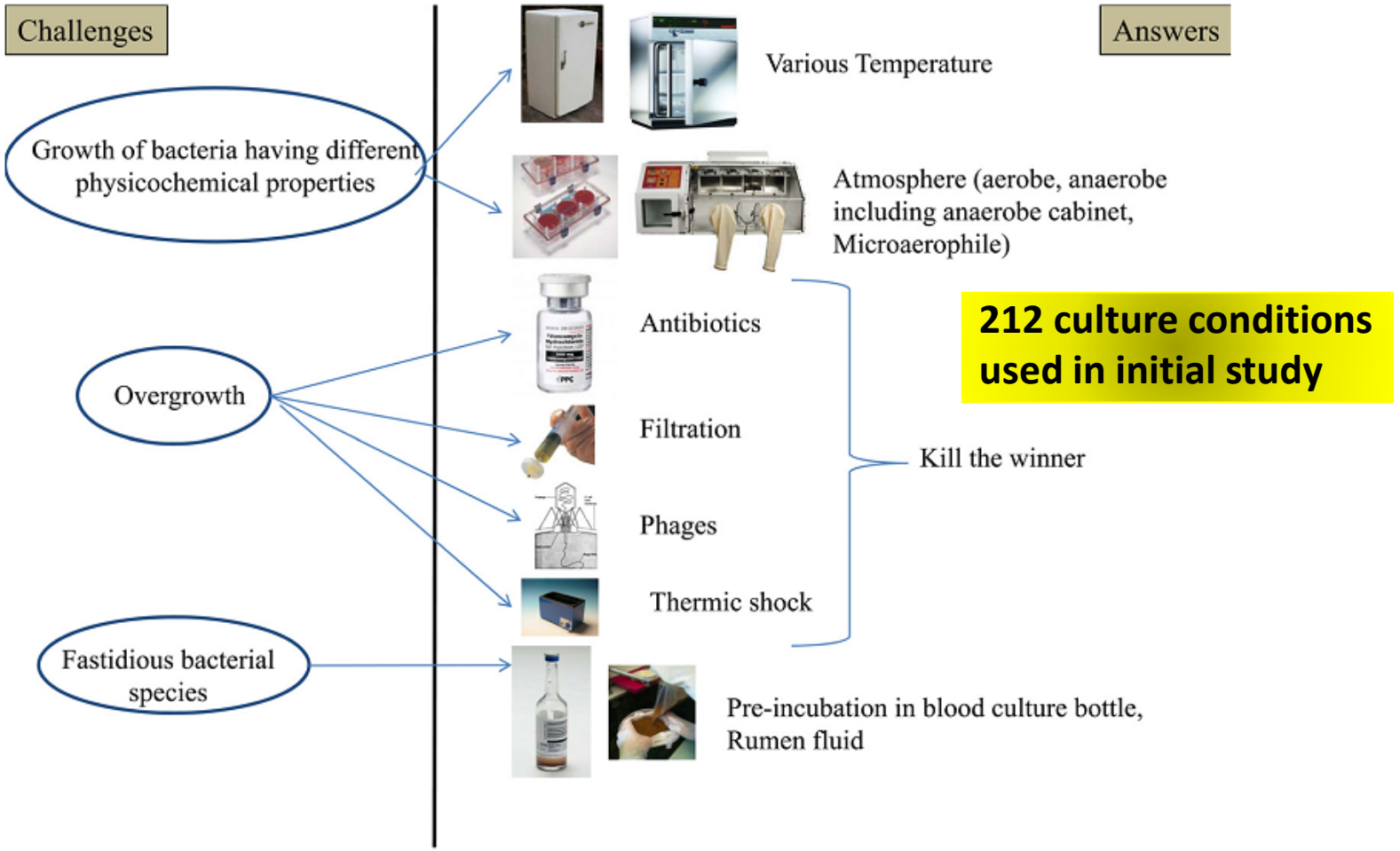
Aims:

- to generate levels of identification of cultures equivalent to those of pyrosequencing by combining novel culture conditions with rapid identifications *via* MALDI-TOF
- To create a major complement to metagenomics

The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota

Lagier et al., Clin Microbiol Rev **28**:237-64 (2015)

Challenges of culturomics and specific answers including techniques used to limit the overgrowth of common bacteria



The rebirth of culture in microbiology through the example of culturomics to study human gut microbiota

Lagier et al., Clin Microbiol Rev **28**:237-64 (2015)

Examples for culture conditions used for culturomics standardisation

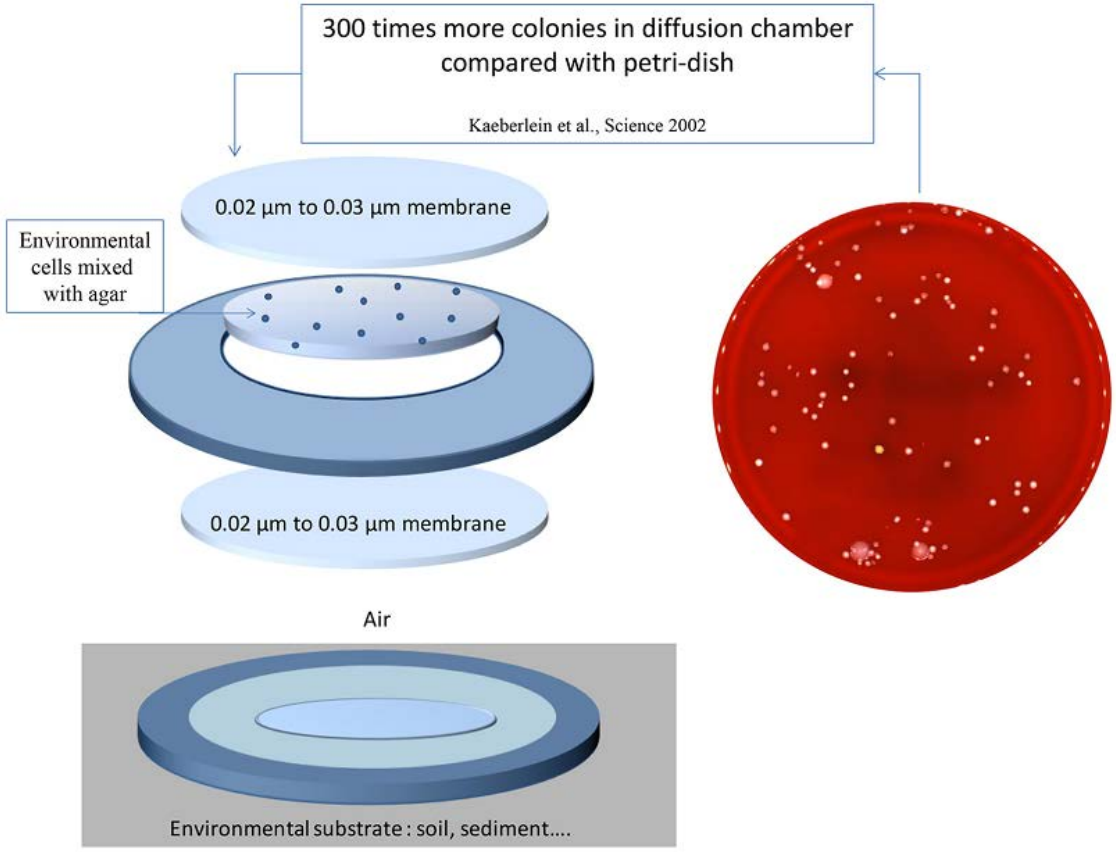
Culture conditions for culturomics standardization
Preincubation in aerobic blood culture bottle with rumen fluid and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle with rumen fluid and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation under aerobic conditions in Trypticase soy broth and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation under anaerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under anaerobic conditions at 28°C
Preincubation under aerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under aerobic conditions at 28°C
Preincubation under anaerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation under aerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle with stool filtered at 5 μm and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in aerobic blood culture bottle with stool filtered at 5 μm and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation in aerobic blood culture bottle with 5 ml sheep blood and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle with 5 ml sheep blood and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle after thermic shock at 80°C during 20 min and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in anaerobic blood culture bottle with 5 ml rumen fluid and sheep blood and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in aerobic blood culture bottle with 5 ml rumen fluid and sheep blood and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation under aerobic conditions in brain heart infusion broth with 5% sheep blood and then 5% sheep blood agar under aerobic conditions at 37°C
Preincubation under anaerobic conditions in marine broth and then 5% sheep blood agar under anaerobic conditions at 37°C
Preincubation in aerobic marine broth and then 5% sheep blood agar under aerobic conditions at 37°C

- **70 of 212 tested culture conditions were sufficient to grow 100% of all clones**
- **20 top culture conditions enabled growth of 73% of the clones**

The rebirth of culture in microbiology through the example of culturomics to study human gut Microbiota

Lagier et al., Clin Microbiol Rev **28**:237-64 (2015)

Improving bacterial culture due to diffusion chambers

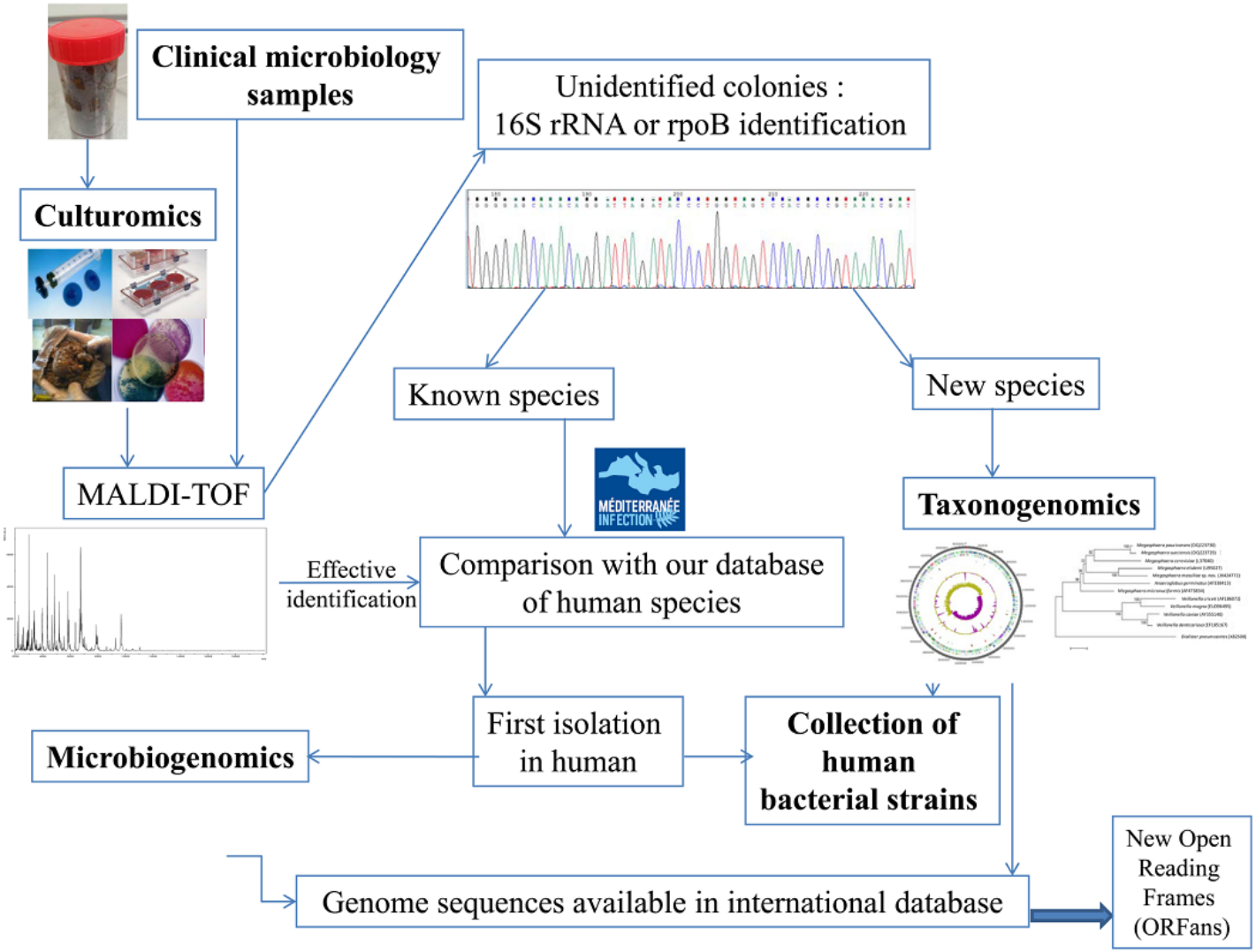


32,000 colonies obtained from **212 culture conditions** yielded **340 species** from **117 genera**, including **31 novel species** never seen before

The rebirth of culture in microbiology through the example of culturomics to study human gut Microbiota

Lagier et al., Clin Microbiol Rev **28**:237-64 (2015)

Overall process from clinical samples to genomic applications



The rebirth of culture in microbiology through the example of culuromics to study human gut Microbiota

Lagier et al., Clin Microbiol Rev **28**:237-64 (2015)

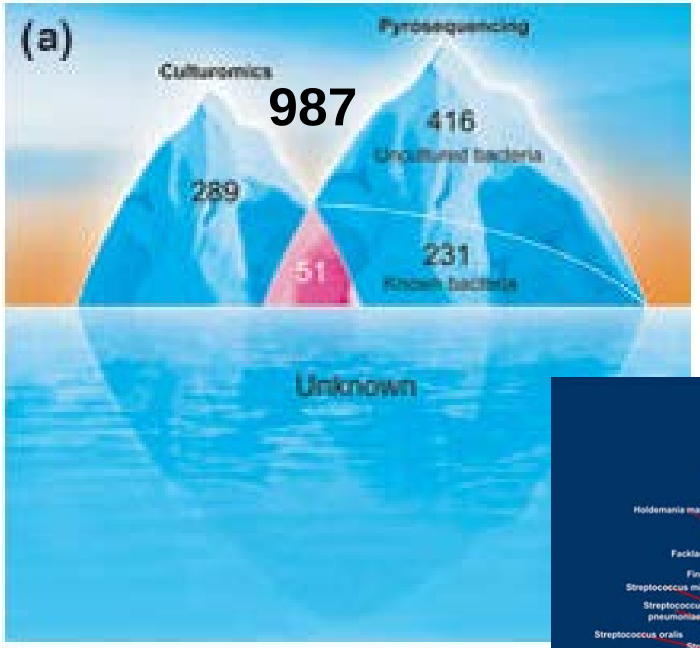
Characteristics of 23 n0vel bacterial species and genera cultures from two Senegalese stools

	Phylum	Initial culture conditions	Diameter (μm) (EM)	Genome size estimate (Mb)	ORFan (%)	Estimated GC content (%)	Genbank no.
N'Diop stool sample							
New species							
Oceanobacillus massiliensis	*Firmicutes*	Filtration brain–heart infusion 5% sheep blood 0.45-μm aerobe, 37°C	0.70	3.6	5.6	41	HQ586877
Bacillus timonensis	*Firmicutes*	Brain–heart infusion + sheep blood 5%, aerobe, 37°C	0.66	4.7	6.8	38.3	JF824810
Dielmo stool sample							
New species							
Kurthia massiliensis	*Firmicutes*	CNA aerobe 2.5% CO_2, 37°C	1.08	3.3	11.9	39.7	JF824795
Kurthia senegalensis	*Firmicutes*	Filtration 5% sheep blood agar 1.2-μm aerobe, 37°C	1.03	2.9	11.3	39.6	JF824796
Kurthia timonensis	*Firmicutes*	HTM, aerobe, 2.5% CO_2, 37°C	0.94	4.1	16.2	39	JF824797
Anaerococcus senegalensis	*Firmicutes*	*Brucella* anaerobe, 37°C	0.68	1.8	3	28.5	JF824805
Paenibacillus senegalensis	*Firmicutes*	Schaedler kanamycin vancomycin, aerobe, 37°C	0.66	5.7	10.7	48.3	JF824808
Bacillus massiliasenegalensis	*Firmicutes*	5% sheep blood agar, aerobe, 28°C	0.64	4.9	7.7	37.7	JF824800
Clostridium senegalense	*Firmicutes*	Inoculation in blood culture bottle for 5 days with 5 mL of sheep blood, 5% sheep blood agar, anaerobe, 37°C	1.05	3.9	11.5	29.3	JF824801
Peptoniphilus senegalensis	*Firmicutes*	Inoculation in blood culture bottle for 10 days with 5 mL of sheep blood, 5% sheep blood agar, anaerobe, 37°C	0.64	1.8	3.9	32.5	JF824803
Peptoniphilus timonensis	*Firmicutes*	Inoculation in blood culture bottle anaerobe for 14 days with 8 mL of rumen fluid, 5% sheep blood agar, anaerobe, 37°C	0.91	1.7	9.3	31	JN657222
Ruminococcus massiliensis[a]	*Firmicutes*	Inoculation in blood culture bottle anaerobe for 14 days with 8 mL of rumen fluid 5% sheep blood agar, anaerobe, 37°C	0.96	5.1	25	57	JN657221
Alistipes senegalensis	*Bacteroidetes*	Schaedler kanamycin vancomycin, anaerobe, 37°C	0.53	4	3.8	58.3	JF824804
Alistipes timonensis	*Bacteroidetes*	Inoculation in blood culture bottle anaerobe for 5 days, Schaedler kanamycin vancomycin, anaerobe 37°C	0.62	3.5	2.9	58.8	JF824799
Cellulomonas massiliensis	*Actinobacteria*	Passive filtration with Leptospira broth, 5% sheep blood agar, aerobic atmosphere, 37°C	0.48	3.4	7.9	73.9	JN657218
Aeromicrobium massiliense	*Actinobacteria*	5% sheep blood agar, aerobe, 37°C	1.04	3.3	10.5	72.6	JF824798
Brevibacterium senegalense	*Actinobacteria*	*Brucella*, aerobe, 37°C	0.68	3.4	9.6	69.9	JF824806
Enterobacter massiliensis	*Proteobacteria*	Phage T1 + T4, then 5% sheep blood agar, aerobe, 37°C	1.02	4.9	3	55.4	JN657217
Herbaspirillum massiliense	*Proteobacteria*	Passive filtration with Leptospira broth, 5% sheep blood agar, aerobic atmosphere, 37°C	0.44	4.2	8.1	59.7	JN657219
Microvirga massiliensis	*Proteobacteria*	MOD 2, aerobe, 37°C	2.28	9.35	24.1	59.2	JF824802
New genera							
Dielma fastidiosa	*Firmicutes*	Inoculation in blood culture bottle anaerobe for 10 days, brain–heart infusion, anaerobe, 37°C	0.59	3.6	10.5	40	JF824807
Senegalemassilia anaerobia	*Actinobacteria*	Inoculation in blood culture bottle anaerobe for 5 days. 5% sheep blood aga,r anaerobe, 37°C	0.70	2.3	6.3	61.8	JF824809
Timonella senegalensis	*Actinobacteria*	Inoculation in blood culture bottle anaerobe for 14 days with 8 mL of rumen fluid, 5% sheep blood agar, anaerobe, 37°C	0.59	3	11.9	61.3	JN657220

Microbial culturomics: paradigm shift in the human gut microbiome study

Lagier et al., Clin Microbiol Infect **18**:1185-94 (2012)

Comparison of identification of bacteria in the human gut by culturomics and metagenomics



Microbial culturomics: paradigm shift in the human gut microbiome study

Lagier et al., Clin Microbiol Infect **18**:1185-94 (2012)

Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample

Pfleiderer et al., Eur J Clin Microbiol Infect Dis **32**:1471-81 (2013)

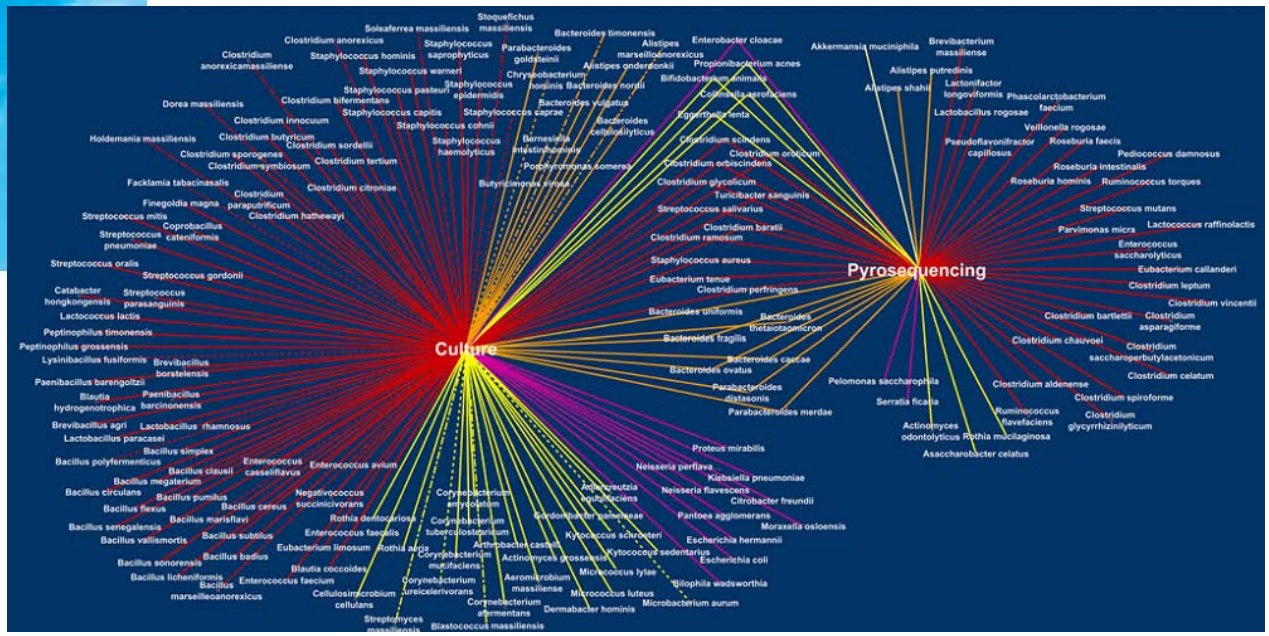
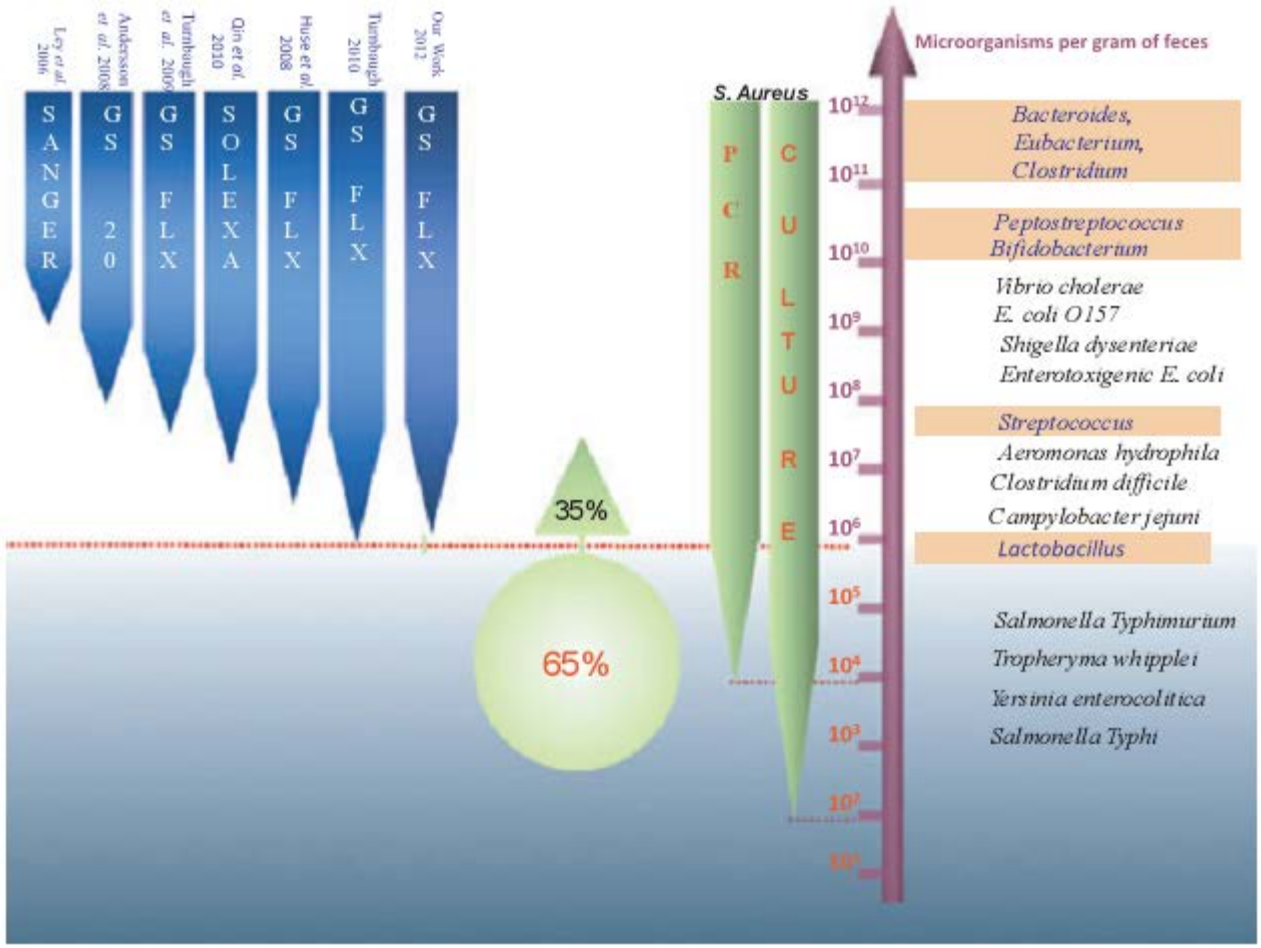


Fig. 2 Comparison of the pyrosequencing and culture results. The *broken lines containing dots and dashes* represent new bacterial species, while a simple *dotted line* represents a species isolated for the first time from the human gut. The different colors represent each phylum: *red*, *Firmicutes*; *orange*, *Bacteroidetes*; *yellow*, *Actinobacteria*; *pink*, *Proteobacteria*; *light yellow*, *Verrucomicrobia*

The detection thresholds of metagenomic and culturomic approaches



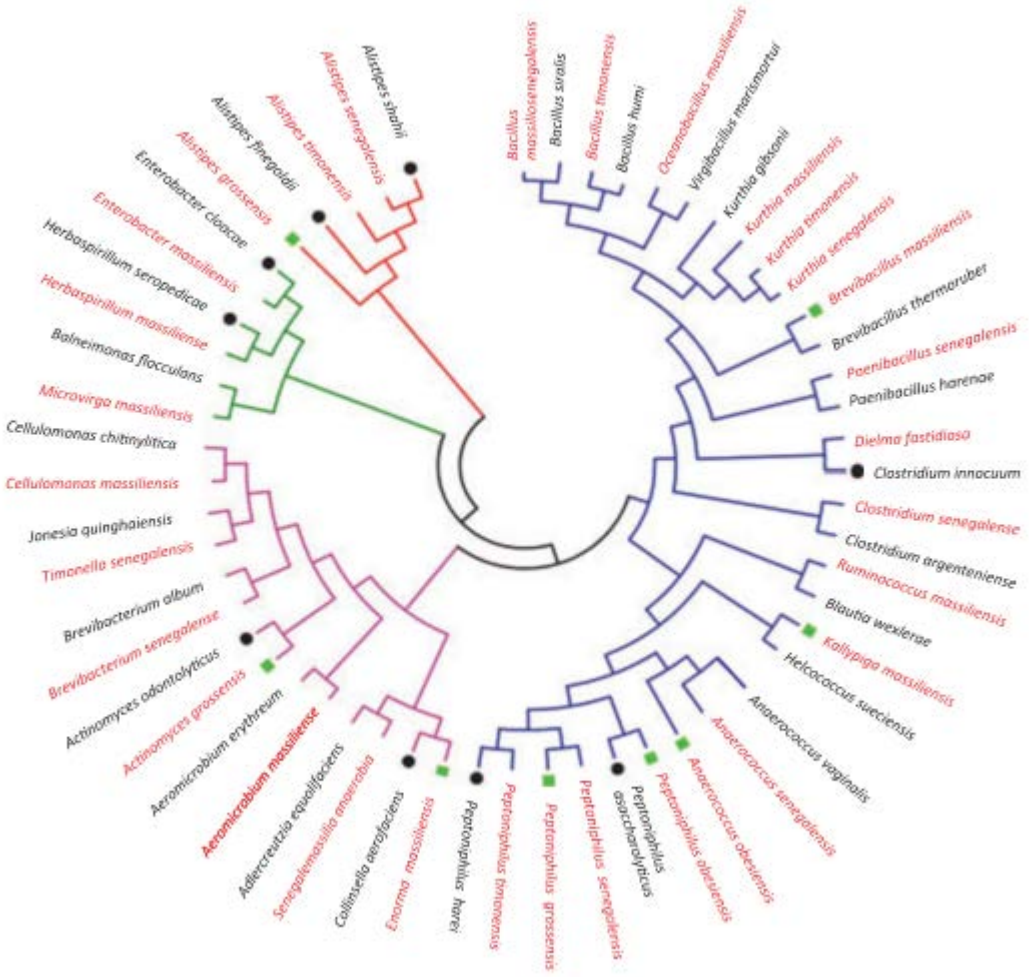
The detection threshold of metagenomic methods correlates with the concentration of bacteria in the investigated sample divided by the number of generated sequences.

The blue pointed shapes show the detection depth of different published metagenomic analyses of the human gut microbiome. The upper dotted red line shows the detection threshold of the most powerful available metagenomic methods, the middle line shows the detection threshold of PCR, and the lower line shows the detection threshold of culturomics. The latter two thresholds were determined by detection of *Staphylococcus aureus* that was added to the samples in varying concentrations (indicated by green pointed shapes). Among the 340 cultivated bacterial species, 29 were identified only after several days of incubation in an anaerobic blood culture bottle, so their concentrations in the original samples could not be estimated. Among the remaining 311 bacteria, 203 (65%) were found at concentrations of $<10^6$ CFU/g of stool, i.e. below the detection threshold of metagenomic methods

Microbial culturomics: paradigm shift in the human gut microbiome study

Lagier et al., Clin Microbiol Infect **18**:1185-94 (2012)

Phylogenetic tree representing the new bacterial species and genera obtained by culturomics



Phylogenetic tree representing the new bacterial species and genera obtained by culturomics.

Red labels indicate the new species found in the Senegalese patients and obese patient. Dark labels indicate the closest neighbour species defined as isolates and type in the RDP-II database.

Tree branches in red, dark green, purple and blue represent the phyla *Bacteroidetes*, *Proteobacteria*, *Actinobacteria,* and *Firmicutes*, respectively. Green squares denote new species found in the obese patient.

Microbial culturomics: paradigm shift in the human gut microbiome study

Lagier et al., Clin Microbiol Infect **18**:1185-94 (2012)

Growing importance of new bacterial species isolated from human gut *via* culturomics

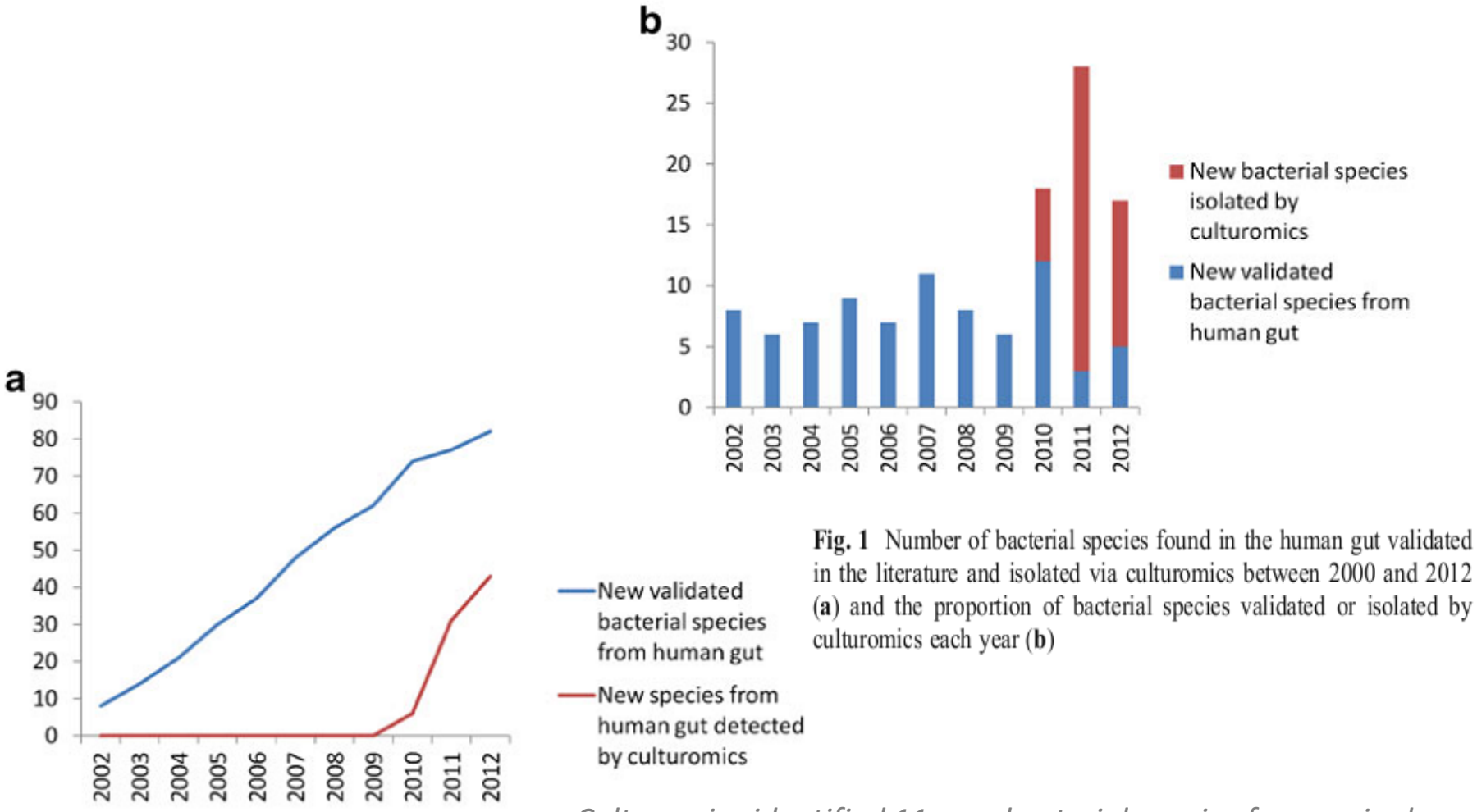


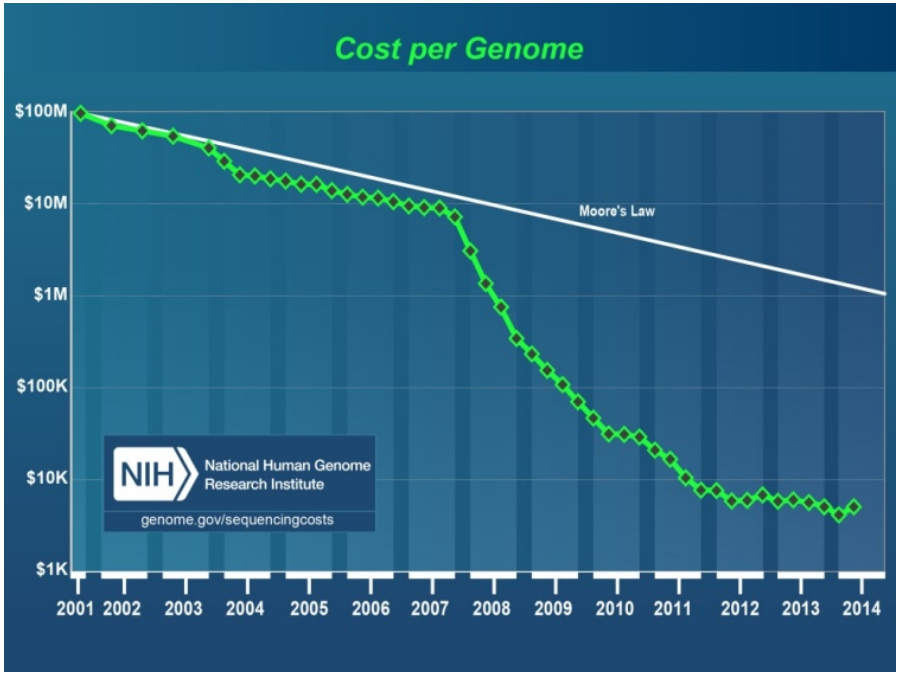
Fig. 1 Number of bacterial species found in the human gut validated in the literature and isolated via culturomics between 2000 and 2012 (**a**) and the proportion of bacterial species validated or isolated by culturomics each year (**b**)

Culturomics identified 11 new bacterial species from a single anorexia nervosa stool sample

Pfleiderer et al., Eur J Clin Microbiol Infect Dis **32**:1471-81 (2013)

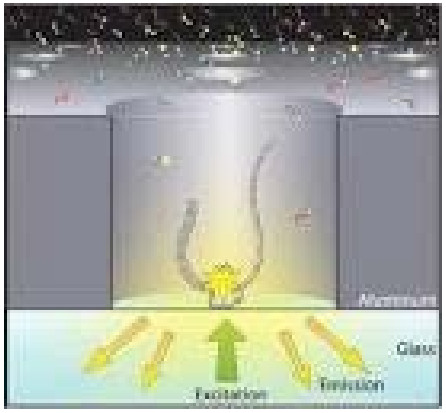
Culturomics will benefit form further progress in *omics* technologies

- Continued **decrease in costs for sequencing** as driver for further progress in *omics* fields
- The **availability of cultures** not only enables genome sequencing, comparative analysis, and functional genomics, but all downstream (biotechnological) applications

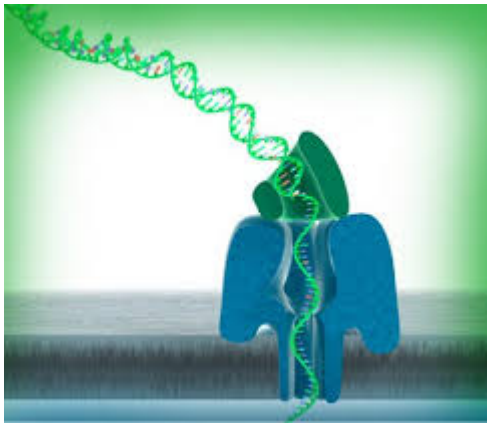


http:genome.gov/sequencingcosts

State of the art third generation single molecule sequencing technologies

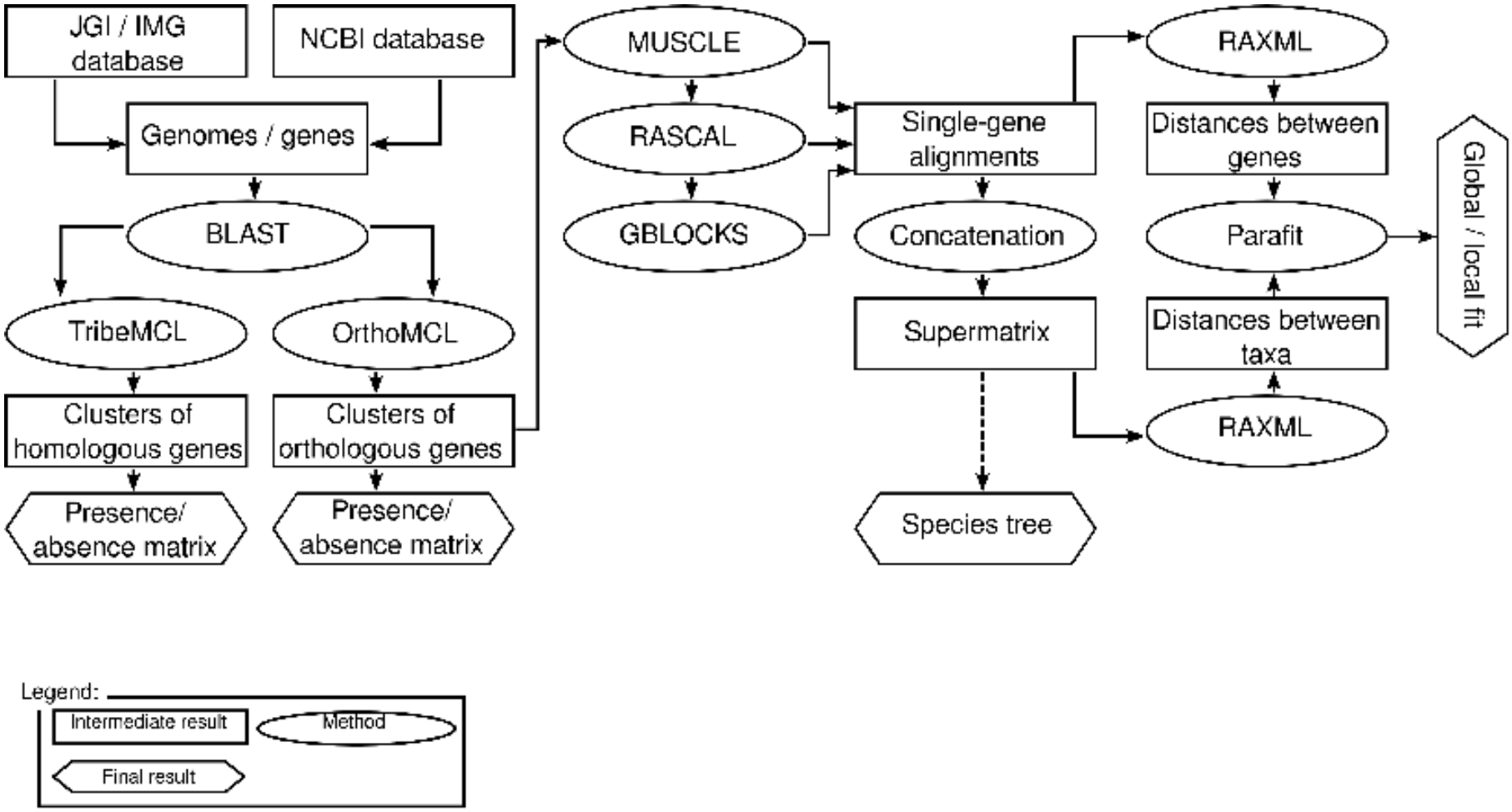


PacBio: sequencing by direct observation of synthesis



Oxford Nanopore: sequencing without synthesis of DNA

Phylogenomics analysis pipeline

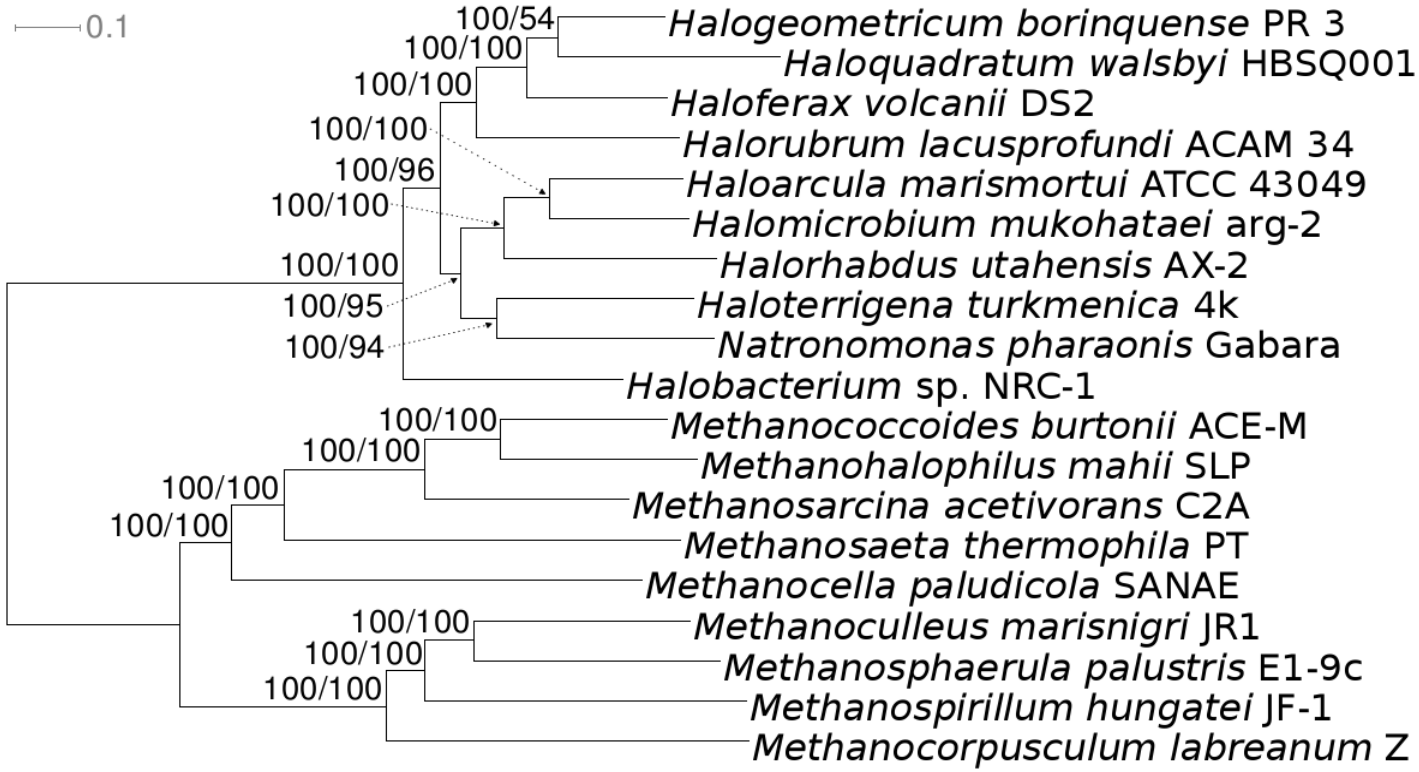


automated data handling pipeline

Scheuner and Goeker, unpublished

Phylogenomic trees with extremely high bootstrapping support

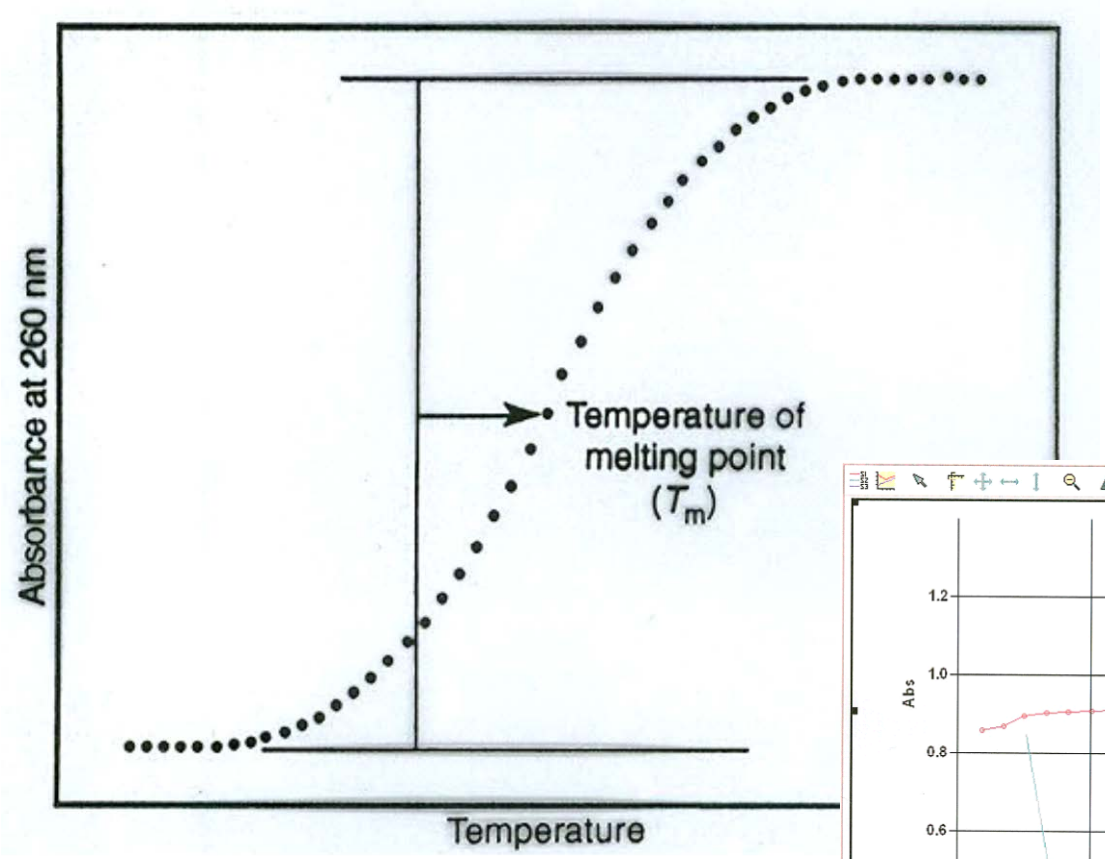
phylogenies from genome sequences and proteomes
- *generated by automated data handling pipelines*

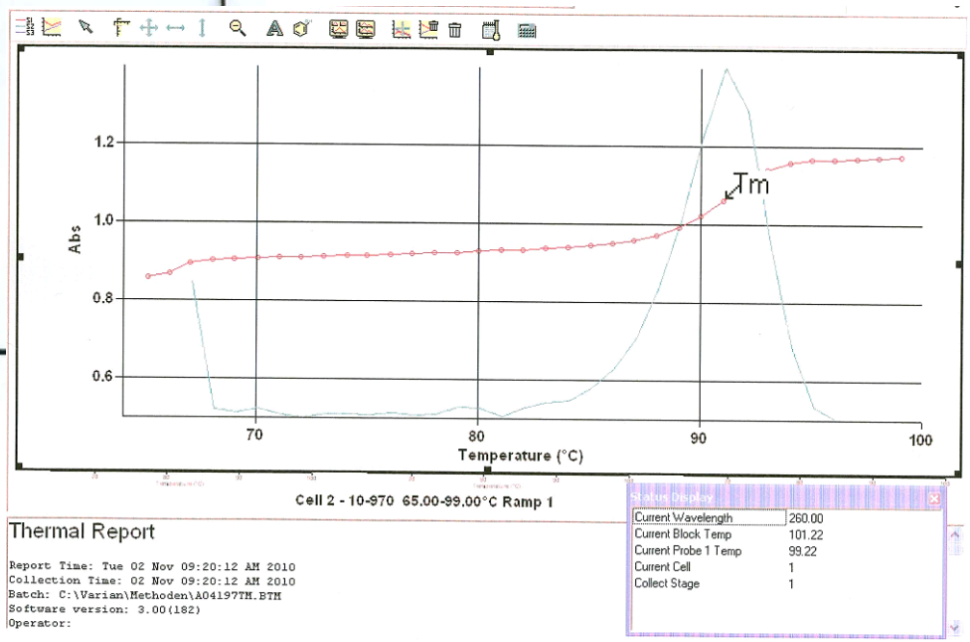


Insights into the diversity of catabolic metabolism from 10 haloarchaeal genomes

Anderson *et al*. (**2011**) *PLoS ONE* **6**:e120237

DNA-DNA Hybridization: the classical golden standard in species description

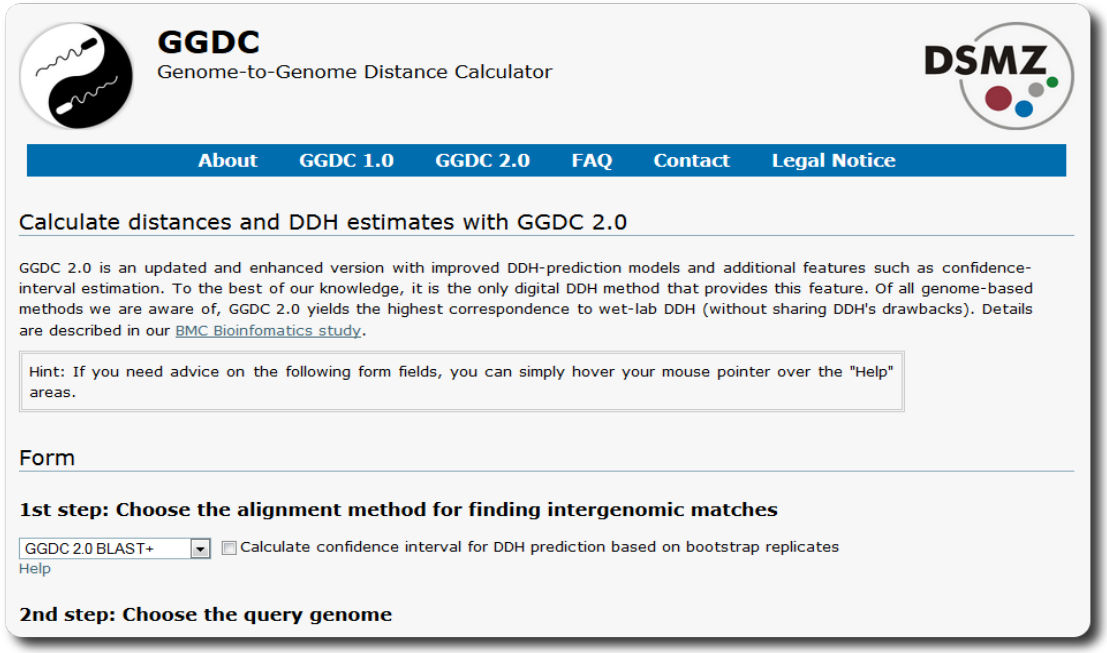




Digital alternatives to wet lab DNA-DNA hybridization: ANI and GGDC digital DDHs

> *Investigators are encouraged to propose new [...] genomic methods [...] provided [...] a sufficient degree of congruence between the technique used and DNA:DNA reassociation.*
>
> — *Stackebrandt et al. 2002*

Lack of congruency
⟹ inconsistencies in microbial taxonomy



GGDC is based on **GBDP**

- established ten years ago (Henz *et al.*, 2004)
- devised for assessing genome-based phylogenies
- most accurate known whole-genome phylogeny method (Patil and McHardy, 2013)

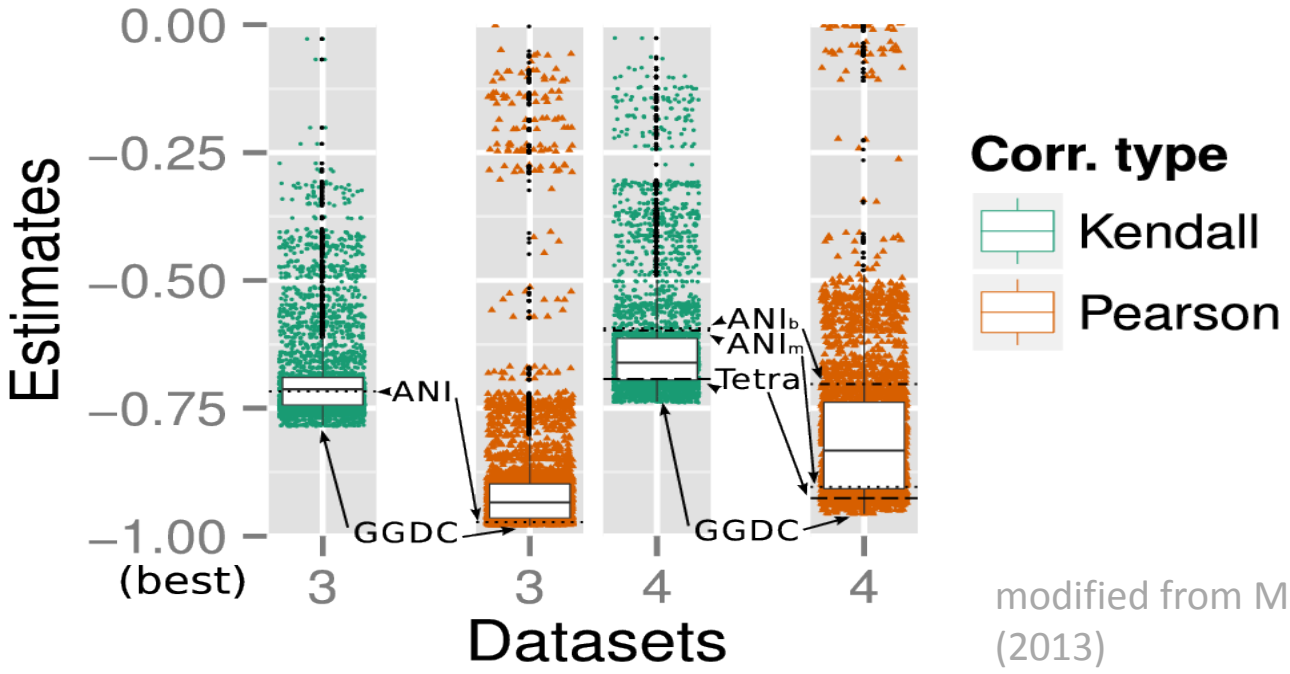
http://ggdc.dsmz.de

Comparability with wet lab DDHs
Correspondence to wet lab DDHs

Direct comparability with wet lab DDHs seems important

- microbiologists are used to the established DDH scale and thresholds
- digital DDH alternatives operate on a scale of their own (drawback)
- GGDC predicts digital DDH on the well-known DDH scale
- mimicking DDH on average as good as possible without mimicking its error rate is the aim

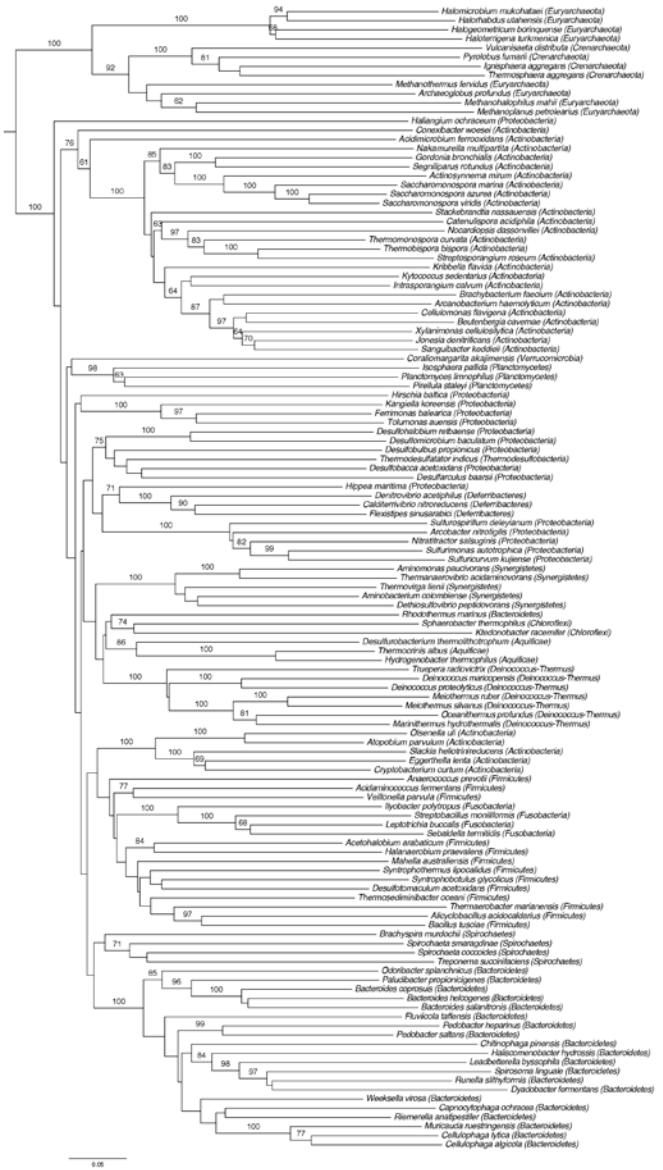
GGDC yields very high correspondence to wet lab DDHs



modified from Meier-Kolthoff *et al.* (2013)

Improved Phylogenetic Reliability

- GBDP primarily from phylogenomics
- branch support with pseudo-bootstrapping
- essential for phylogenetic analysis

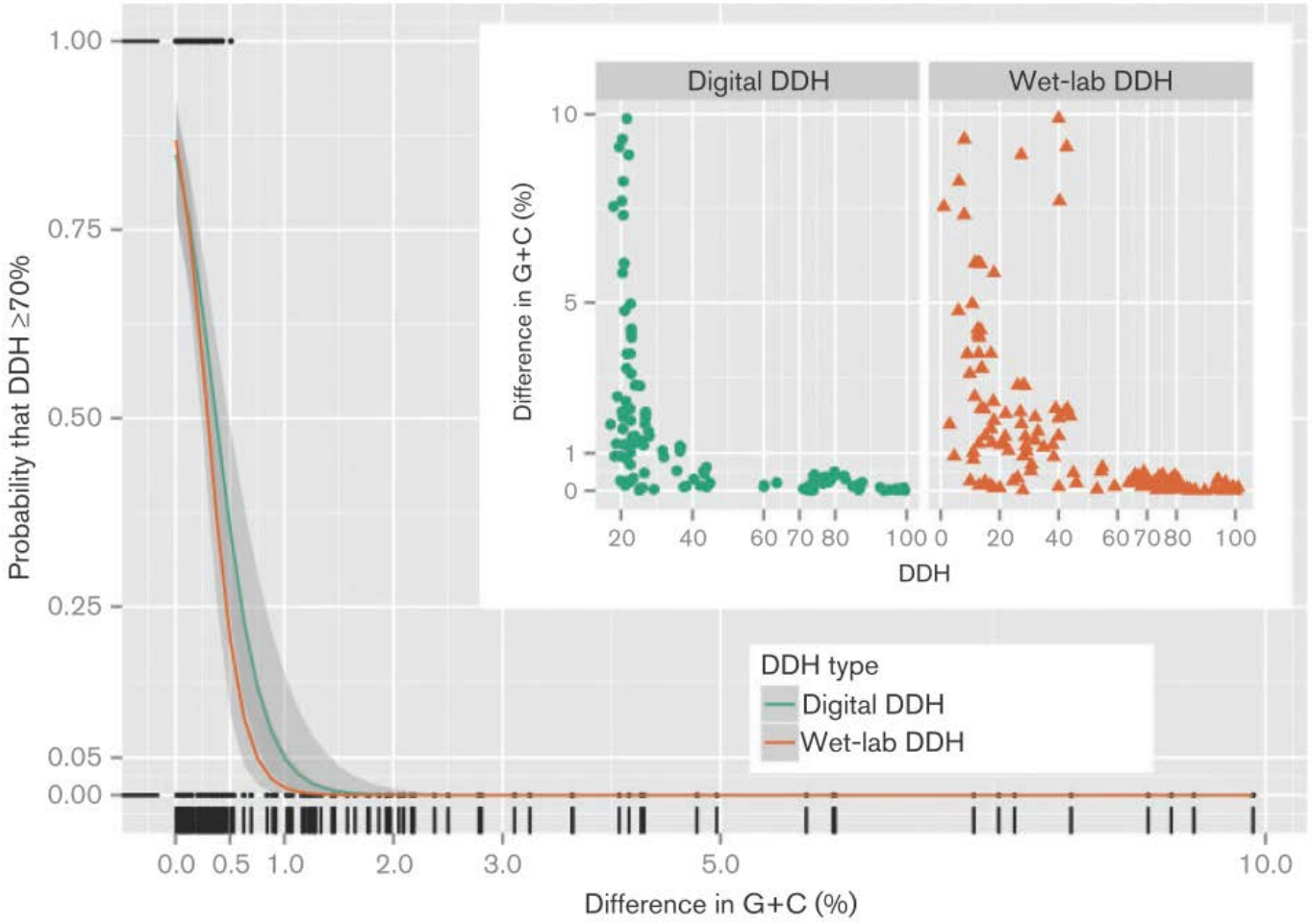


Genome-based GEBA tree

Meier-Kolthoff *et al*. (2013)

Taxonomic use of G+C Content in the Genomic Age

Within-species difference in G+C content <1% *if calculated from genome sequence*



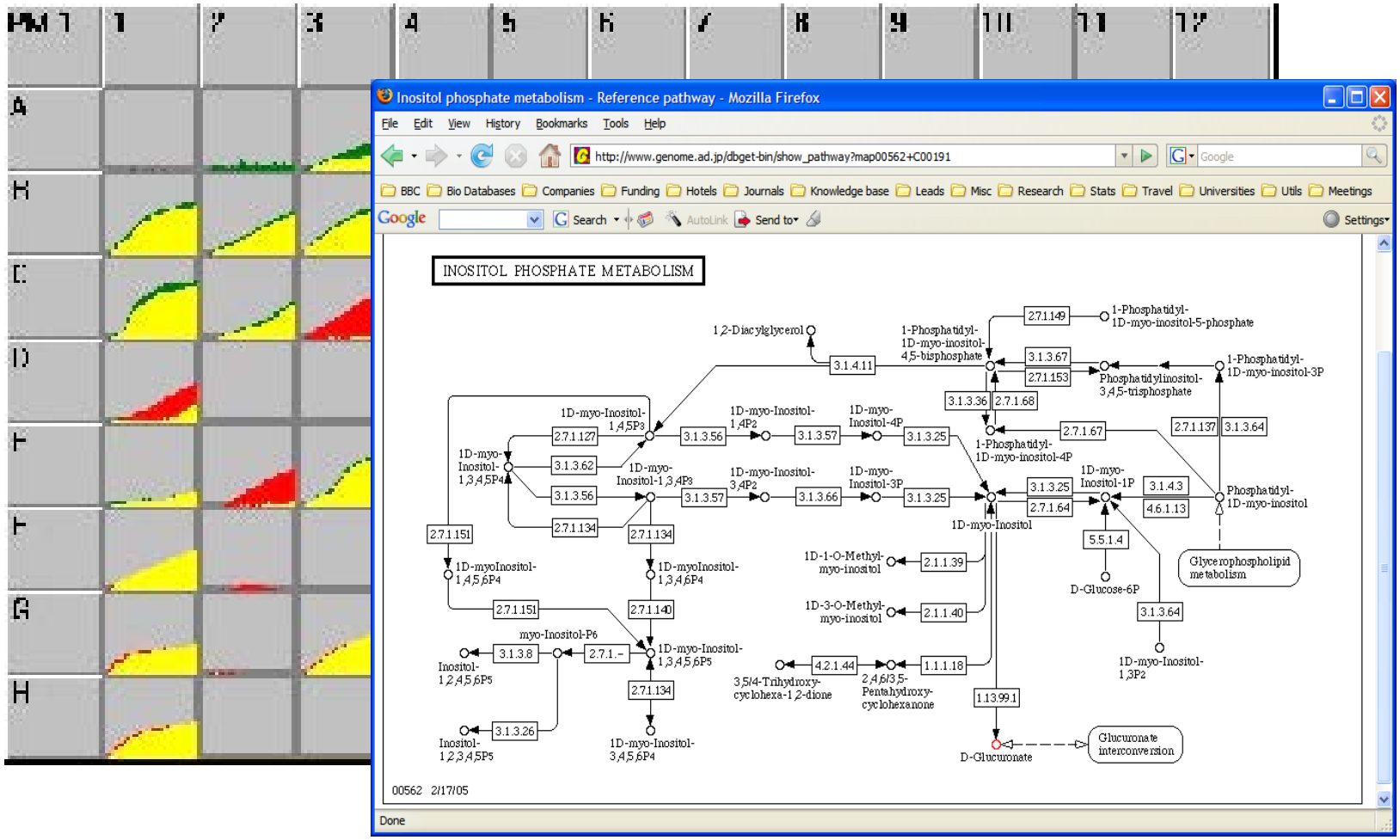
Meier-Kolthoff *et al.* (2014)

Confirmed in 9279 genome-sequence pairs

Availability of cultures also allows to analyse phenotypes

- via high throughput phenotyping
- based on pathways reconstructed from the genome sequence

Phenotype MicroArray-Technology for GEBA strains of model organisms



The Genomic Encyclopaedia of *Bacteria* and *Archaea,* GE*BA*

A systematic, genomic exploration
of all species of bacteria and archaea
with validly published names

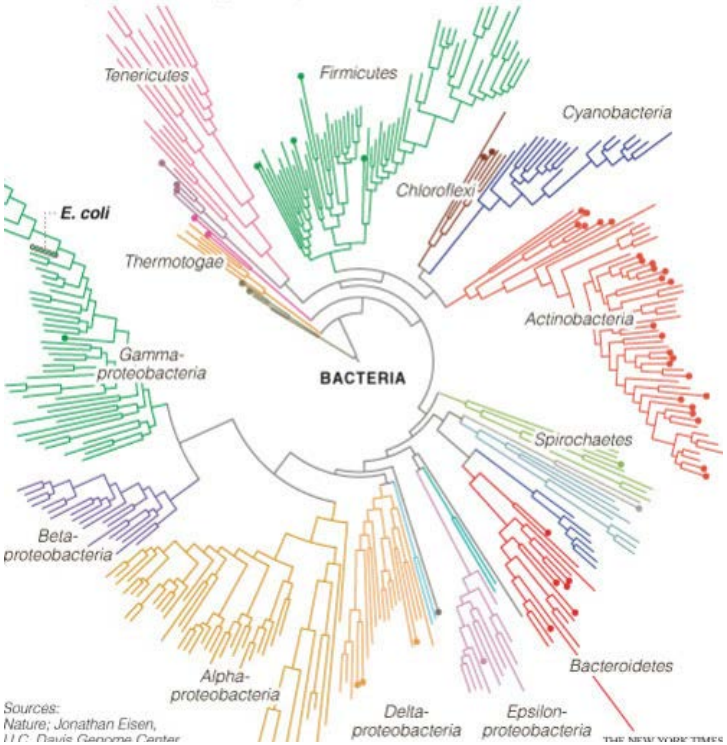
The ambitious but assuredly tractable goal of the project is to sequence the genome of at least one representative (type strain) of every bacterial and archaeal species that has a validly published name in conformance with the Bacteriological Code.

- standardized methods for DNA extraction, sequencing & annotation
- standardized dissemination of sequence data & metadata
- development of tools for species discrimination (dDDH) and genome-scale phylogenies
- taxonomic emendation of species & genus descriptions

First chapters - published 5 years ago

Filling Out the Branches

This "genome tree" shows relationships among the different species of bacteria that have had their genomes sequenced to date, with major phyla shown in different colors. A new project intended to expand the range and variety of sequenced microbes has completed its first 56 species, including the 53 species of bacteria marked below with dots.



December 29, 2009

Scientists Start a Genomic Catalog of Earth's Abundant Microbes

nature news

Published online 23 December 2009 | Nature | doi:10.1038/news.2009.1161

News

Microbial encyclopaedia guided by evolution

Sequencing project reveals microbial cache of protein families.

SCIENCE WATCH

THOMSON REUTERS

HOME | ABOUT THOMSON REUTERS | PRESS ROOM | CONTACT US | Site Search | SEARCH

What's Hot In... › What's Hot in Biology - Menu › Branching Out with Phylogenetically Driven Genome Sequencing

FOLLOW US | SHARE

BRANCHING OUT WITH PHYLOGENETICALLY DRIVEN GENOME SEQUENCING

WHAT'S HOT IN BIOLOGY, NOVEMBER/DECEMBER 2011

At #1 is a paper from Jonathan Eisen at the University of California, Davis, and a large group of colleagues, who set out to see if they could target microbes for sequencing specifically in order to understand better their evolution, phylogenetic history, and functioning.

The team's quest started from the observation that almost everything we knew about bacterial evolution and family trees was derived from just three of the 40 or so phyla of bacteria. The dominant phylogenetic tree was derived from the sequence of one small piece of RNA, and there were bits of the tree that made no sense. So Eisen and his colleagues set out to draw up a "Genomic Encyclopedia of Bacteria and Archaea" (GEBA).

They first made a list of branches of the tree that had little or no sequence data available and sent it to Hans-Peter Klenk at the DSMZ (the German Collection of Microorganisms and Cell Cultures), who identified about 200 microbes from those branches in the collection.

A woman works with human genetic material at a laboratory in Munich, May 2011. REUTERS/Michael Dald.

WHAT'S HOT IN...

- Biology
- Chemistry
- Medicine
- Physics
- Home

MORE ANALYSES

- Featured Analyses
- Special Topics

Second chapter - published last year

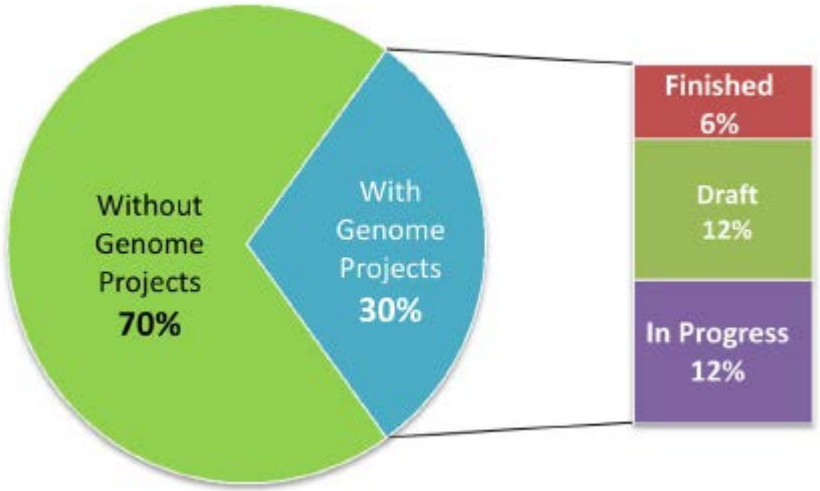
OPEN ACCESS Freely available online

PLOS | BIOLOGY

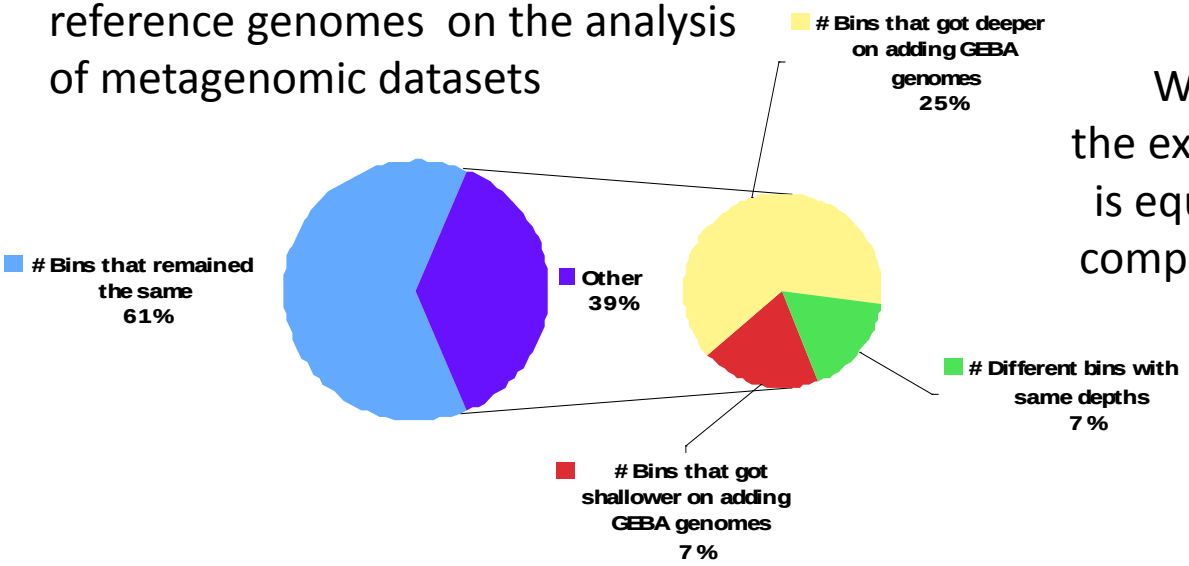
Community Page

Genomic Encyclopedia of Bacteria and Archaea: Sequencing a Myriad of Type Strains

Nikos C. Kyrpides[1,2*], Philip Hugenholtz[3], Jonathan A. Eisen[4], Tanja Woyke[1], Markus Göker[5], Charles T. Parker[6], Rudolf Amann[7], Brian J. Beck[8], Patrick S. G. Chain[9], Jongsik Chun[10], Rita R. Colwell[11,12], Antoine Danchin[13], Peter Dawyndt[14], Tom Dedeurwaerdere[15], Edward F. DeLong[16], John C. Detter[9], Paul De Vos[14,17], Timothy J. Donohue[18], Xiu-Zhu Dong[19], Dusko S. Ehrlich[20], Claire Fraser[21], Richard Gibbs[22], Jack Gilbert[23], Paul Gilna[24], Frank Oliver Glöckner[7,25], Janet K. Jansson[26], Jay D. Keasling[26,27], Rob Knight[28], David Labeda[29], Alla Lapidus[30,31], Jung-Sook Lee[32], Wen-Jun Li[33], Juncai MA[34], Victor Markowitz[1,26], Edward R. B. Moore[35], Mark Morrison[36], Folker Meyer[37], Karen E. Nelson[38], Moriya Ohkuma[39], Christos A. Ouzounis[40,41], Norman Pace[42], Julian Parkhill[43], Nan Qin[44], Ramon Rossello-Mora[45], Johannes Sikorski[5], David Smith[46], Mitch Sogin[47], Rick Stevens[37], Uli Stingl[48], Ken-ichiro Suzuki[49], Dorothea Taylor[6], Jim M. Tiedje[50], Brian Tindall[5], Michael Wagner[51], George Weinstock[52], Jean Weissenbach[53], Owen White[21], Jun Wang[44,54], Lixin Zhang[19,55], Yu-Guang Zhou[34], Dawn Field[56], William B. Whitman[57], George M. Garrity[6,50], Hans-Peter Klenk[5*]



Effect of the availability of reference genomes on the analysis of metagenomic datasets



Without the GEBA framework, the exploration of our microbial planet is equivalent to navigation without a compass, a map or stars by which one can fix their position.

Third chapter – ready for submission

GEBA pilot project w. extension (2007-2012) Jonathan Eisen
250 phylogenetically selected type strains. Complete genomes.
all sequences finished; about 175 published

GEBA phase I [KMG-1] **(2011-2014)** Nikos C. Kyrpides
1000 phylogenetically selected type strains. Draft genomes.
all sequences finished; about 20 published

GEBA phase II [KMG-II] **(2012-2015)** Hans-Peter Klenk
100 phylogenetically selected type strains; 800 type strains to complete genera and families; 100 candidate type strains for sp. nov.
Draft genomes. DNA production ongoing

GEBA phase III [KMG-III] **(2013-2016)** William Whitman
1000 type strains from various culture collections and candidate type strains for sp. nov. from various researchers. Draft genomes. DNA production ongoing

GEBA 'phase IV' [Actino-1000] **(2014-2017)** Hans-Peter Klenk
100o selected type strains; Draft and PacBio genomes. DNA production ongoing

Planctomycetales – two genera become five genera

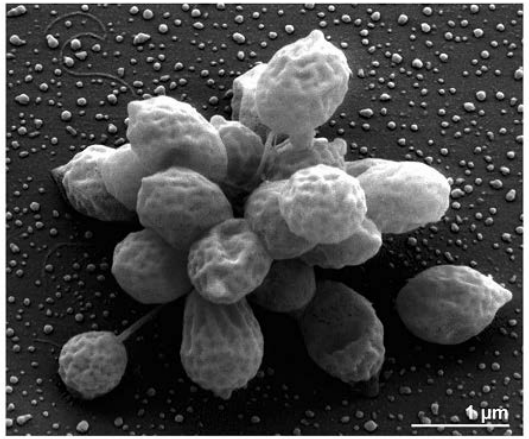


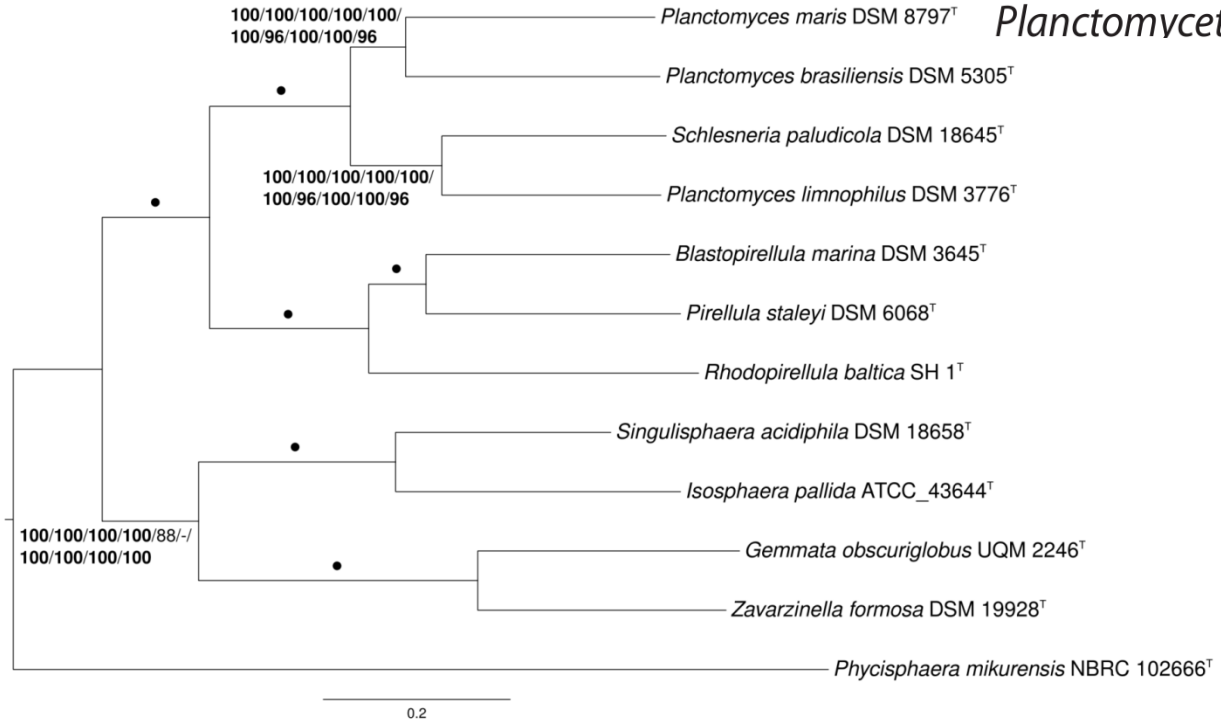
Figure 2 Scanning-electron micrograph of *P. brasiliensis* DSM 5305^T highlighting stalks and crateriform structures on the cell surface.

Scheuner *et al. Standards in Genomic Sciences* 2014, **9**:10
http://www.standardsingenomics.com/content/9/1/10

Standards in Genomic Sciences

EXTENDED GENOME REPORT Open Access

Complete genome sequence of *Planctomyces brasiliensis* type strain (DSM 5305^T), phylogenomic analysis and reclassification of *Planctomycetes* including the descriptions of *Gimesia* gen. nov., *Planctopirus* gen. nov. and *Rubinisphaera* gen. nov. and emended descriptions of the order *Planctomycetales* and the family *Planctomycetaceae*



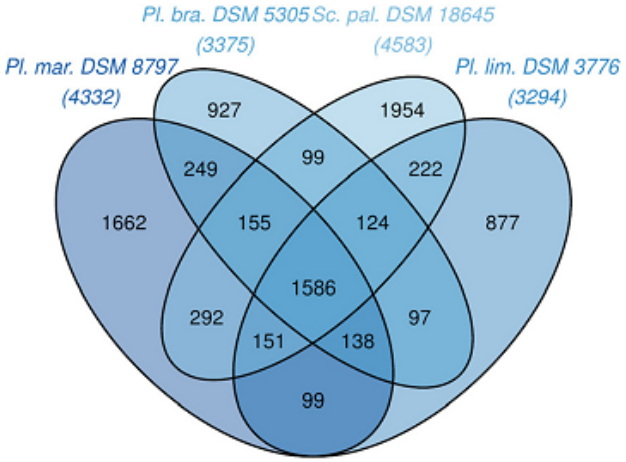


Figure 5 Venn diagram depicting the intersections of sets of homologous proteins of *P. maris*, *P. brasiliensis*, *P. limnophilus* and *S. paludicola*. Their cardinalities are given in parentheses; for the total number of proteins see Table 3 and the resources listed in Table 5. The Venn diagram was calculated with the corresponding R package [92].

Escherischia coli and subspecies

Meier-Kolthoff *et al. Standards in Genomic Sciences* #CITATION
#ARTICLE_URL_DISPLAY_TEXT_FOR_STAMPED_PDF

Standards in Genomic Sciences

EXTENDED GENOME REPORT Open Access

Complete genome sequence of DSM 30083^T, the type strain (U5/41^T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy

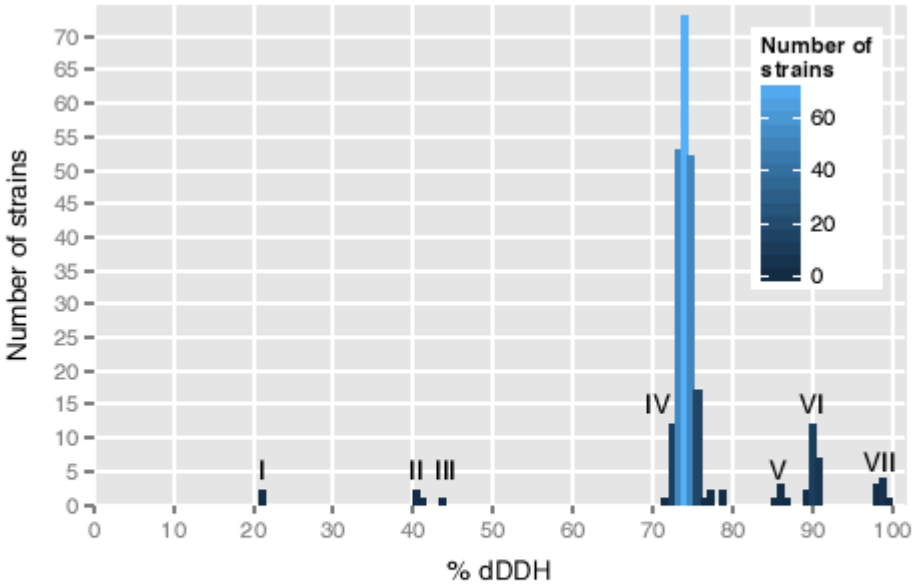


Figure 5 Histogram of the digital DDH similarities between the type strain, DSM 30083^T, and other genome-sequenced *E. coli* strains as well as outgroups. Interesting groups are marked by Roman numerals I-VII: *Escherichia hermannii* and *Shimwellia blattae* (I), *E. fergusonii and E. albertii* (II), *E.* sp. TW09308 (III), *E. coli* (IV-VII). Regarding the revised phylotypes from [66] (compare Figure 6), phylotype B2 is covered by dDDH groups V, VI, and VII with VII being the group containing (among other strains) type strain DSM 30083^T itself and its closest relative *E. coli* S88. IV marks the biggest group which includes phylotypes A, B1, D1, D2, D3, E, F1, F2 and *Shigella* I, IIa, IIb and III. The full list of dDDH values and affiliation to phylotypes is contained in Additional file 2.

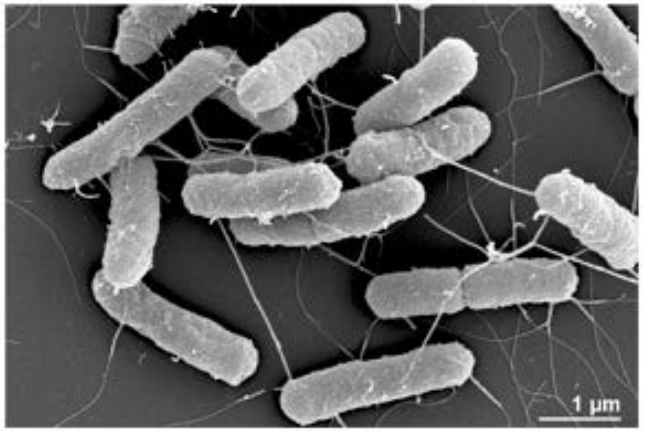
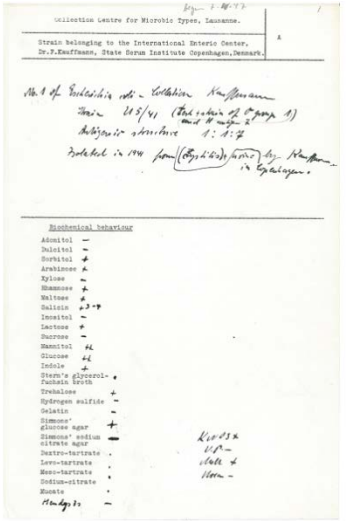
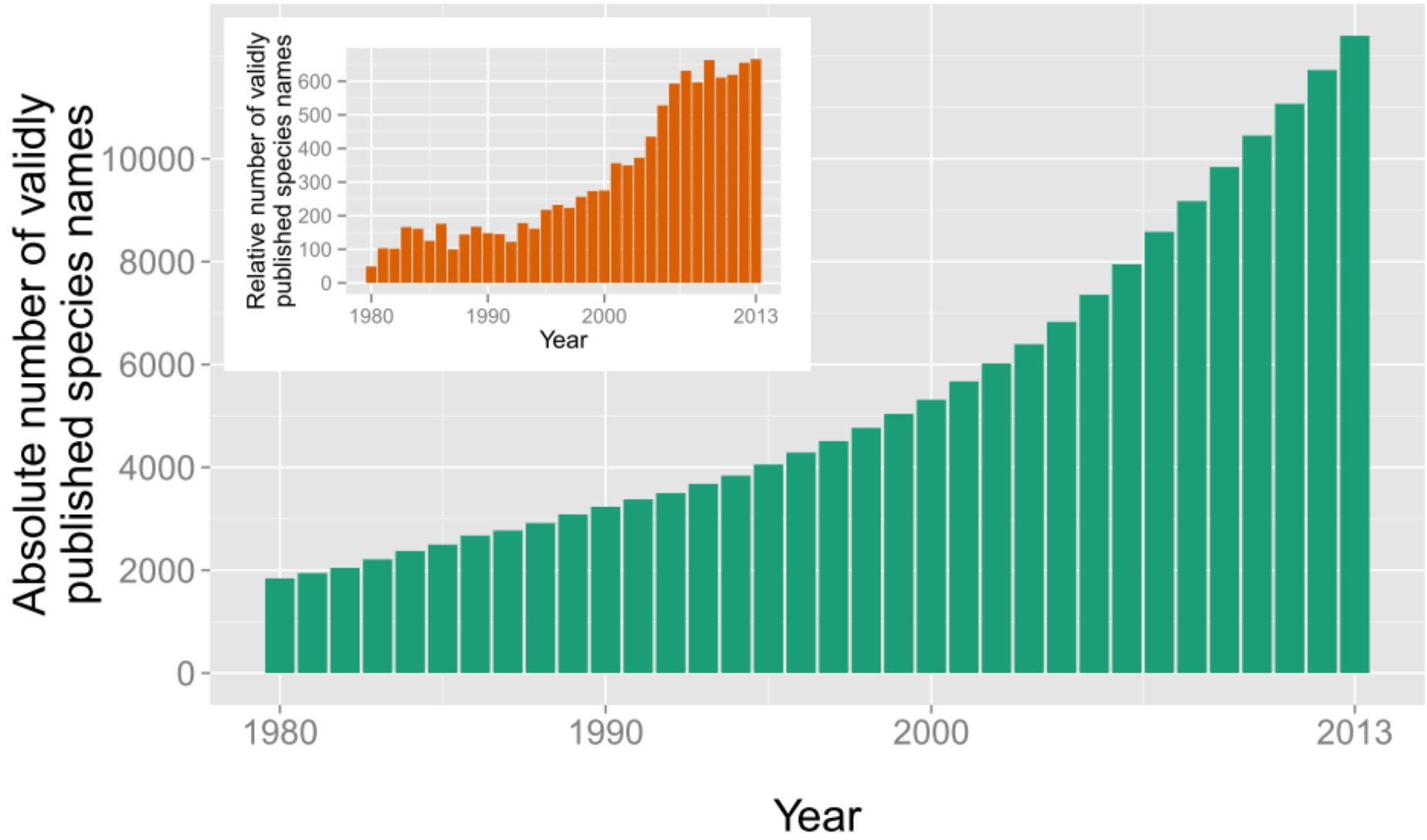


Figure 2 Scanning-electron micrograph of strain *E. coli* DSM 30083^T

Suggested dDDH threshold for differentiation of subspecies: 79-80%



Growth of number of validly published species names since 1980 – might dramatically increase through culturomics



Summary

- The collective high-throughput characterization and quantification of pools of cultures makes **'Culturomics'** a true *-omics* technique
- The poor overlab of strains identified *via* Culturomics and metagenomics makes culturomics a valuable complement to current analysis methods
- Only cultures allow efficient downstream functional analyses and biotechnology
- When combined with state of the art sequencing technologies culturomics allows higly efficient and informative analyses of unexplored biodiversity
- Culturomics might push the speed of accessig microbial biodiversity
- Within the next few years wet lab DNA-DNA-hybridizations (DDHs) will be replaced by digital DDHs (depending on the progress in GEBA)
- Soon we will see more and more whole genome sequence-based phylogenies and less 16S rRNA trees