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

- **Genomics**
- **Proteomics**
- **Metabolomics**
- **Lipidomics**
- **Transcriptomics**
- **Metagenomics**

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. Lagier1,*, F. Armougom1,*, M. Million1, P. Hugon1, I. Pagnier1, C. Robert1, F. Bittar1, G. Fournous1, G. Gimenez1, M. Maraninchi2, J.-F. Trapé3, E. V. Koonin4, B. La Scola1 and D. Raoult1

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.
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 anomaly” (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
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

### Examples for culture conditions used for culturomics standardisation

<table>
<thead>
<tr>
<th>Culture conditions for culturomics standardization</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preincubation in aerobic blood culture bottle with rumen fluid and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in anaerobic blood culture bottle with rumen fluid and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in anaerobic blood culture bottle and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation under aerobic conditions in Trypticase soy broth and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation under anaerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under anaerobic conditions at 28°C</td>
</tr>
<tr>
<td>Preincubation under aerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under aerobic conditions at 28°C</td>
</tr>
<tr>
<td>Preincubation under anaerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation under aerobic conditions in 5% sheep blood broth and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in anaerobic blood culture bottle with stool filtered at 5 µm and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in aerobic blood culture bottle with stool filtered at 5 µm and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in aerobic blood culture bottle with 5 ml sheep blood and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in anaerobic blood culture bottle with 5 ml sheep blood and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>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</td>
</tr>
<tr>
<td>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</td>
</tr>
<tr>
<td>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</td>
</tr>
<tr>
<td>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</td>
</tr>
<tr>
<td>Preincubation under anaerobic conditions in marine broth and then 5% sheep blood agar under anaerobic conditions at 37°C</td>
</tr>
<tr>
<td>Preincubation in aerobic marine broth and then 5% sheep blood agar under aerobic conditions at 37°C</td>
</tr>
</tbody>
</table>

- **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

Improving bacterial culture due to diffusion chambers

300 times more colonies in diffusion chamber compared with petri-dish

Kaeberlein et al., Science 2002

Environmental cells mixed with agar

0.02 μm to 0.03 μm membrane

Air

Environmental substrate: soil, sediment...

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

Overall process from clinical samples to genomic applications

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

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

<table>
<thead>
<tr>
<th>Phylum</th>
<th>Initial culture conditions</th>
<th>Diameter (µm) (EM)</th>
<th>Genome size estimate (Mb)</th>
<th>ORFan (%)</th>
<th>Estimated GC content (%)</th>
<th>Genbank no.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New species</strong></td>
<td><strong>Oceanobacillus messilensis</strong></td>
<td>Fimbriated</td>
<td>0.70</td>
<td>3.6</td>
<td>5.6</td>
<td>41</td>
</tr>
<tr>
<td></td>
<td><strong>Bacterium timorense</strong></td>
<td>Fimbriated</td>
<td>0.66</td>
<td>4.7</td>
<td>6.8</td>
<td>38.3</td>
</tr>
<tr>
<td><strong>Dieden stool sample</strong></td>
<td><strong>Kurthia massiliensis</strong></td>
<td>Fimbriated</td>
<td>1.08</td>
<td>3.3</td>
<td>11.9</td>
<td>39.7</td>
</tr>
<tr>
<td></td>
<td><strong>Kurthia senegalensis</strong></td>
<td>Fimbriated</td>
<td>1.03</td>
<td>2.9</td>
<td>11.3</td>
<td>39.6</td>
</tr>
<tr>
<td></td>
<td><strong>Kurthia timorense</strong></td>
<td>Fimbriated</td>
<td>0.94</td>
<td>4.1</td>
<td>16.2</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td><strong>Anarasea senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.68</td>
<td>1.8</td>
<td>3</td>
<td>28.5</td>
</tr>
<tr>
<td></td>
<td><strong>Poribacillus senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.66</td>
<td>5.7</td>
<td>16.7</td>
<td>40.3</td>
</tr>
<tr>
<td></td>
<td><strong>Bocillia massilisenealegisa</strong></td>
<td>Fimbriated</td>
<td>0.64</td>
<td>4.9</td>
<td>7.7</td>
<td>37.7</td>
</tr>
<tr>
<td></td>
<td><strong>Chloridium senealegisa</strong></td>
<td>Fimbriated</td>
<td>1.05</td>
<td>3.9</td>
<td>11.5</td>
<td>29.3</td>
</tr>
<tr>
<td></td>
<td><strong>Pectinilatus senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.64</td>
<td>1.8</td>
<td>3.9</td>
<td>32.5</td>
</tr>
<tr>
<td></td>
<td><strong>Pectinilatus timorense</strong></td>
<td>Fimbriated</td>
<td>0.91</td>
<td>1.7</td>
<td>9.3</td>
<td>31</td>
</tr>
<tr>
<td></td>
<td><strong>Ruminococcus massiliensis</strong></td>
<td>Fimbriated</td>
<td>0.96</td>
<td>5.1</td>
<td>25</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td><strong>Alseptes senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.53</td>
<td>4</td>
<td>3.8</td>
<td>58.3</td>
</tr>
<tr>
<td></td>
<td><strong>Alseptes timorense</strong></td>
<td>Fimbriated</td>
<td>0.62</td>
<td>3.5</td>
<td>2.9</td>
<td>58.8</td>
</tr>
<tr>
<td></td>
<td><strong>Cellulomonas massiliensis</strong></td>
<td>Fimbriated</td>
<td>0.48</td>
<td>3.4</td>
<td>7.9</td>
<td>73.9</td>
</tr>
<tr>
<td></td>
<td><strong>Aeromicrobium senealegisa</strong></td>
<td>Fimbriated</td>
<td>1.04</td>
<td>3.3</td>
<td>10.5</td>
<td>52.6</td>
</tr>
<tr>
<td></td>
<td><strong>Brevibacterium senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.68</td>
<td>3.4</td>
<td>9.5</td>
<td>69.9</td>
</tr>
<tr>
<td></td>
<td><strong>Enterobacter massiliensis</strong></td>
<td>Fimbriated</td>
<td>1.02</td>
<td>4.9</td>
<td>3</td>
<td>55.4</td>
</tr>
<tr>
<td></td>
<td><strong>Herbaspirillum senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.44</td>
<td>4.2</td>
<td>8.1</td>
<td>59.7</td>
</tr>
<tr>
<td></td>
<td><strong>Micococcus senealegisa</strong></td>
<td>Fimbriated</td>
<td>2.28</td>
<td>9.35</td>
<td>24.1</td>
<td>59.2</td>
</tr>
<tr>
<td><strong>New genera</strong></td>
<td><strong>Dielme fassioda</strong></td>
<td>Fimbriated</td>
<td>0.59</td>
<td>3.6</td>
<td>10.5</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td><strong>Senegalomassalea senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.70</td>
<td>2.3</td>
<td>6.3</td>
<td>61.8</td>
</tr>
<tr>
<td></td>
<td><strong>Timonellaceae senealegisa</strong></td>
<td>Fimbriated</td>
<td>0.59</td>
<td>3</td>
<td>11.9</td>
<td>61.3</td>
</tr>
</tbody>
</table>

**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**

**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.

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

Culturomics will benefit from 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

Legend:
- Intermediate result
- Method
- Final result

automated data handling pipeline
Scheuner and Goeker, unpublished
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*, GEBA

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
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.

Microbial encyclopaedia guided by evolution
Sequencing project reveals microbial cache of protein families.

Scientists Start a Genomic Catalog of Earth’s Abundant Microbes
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

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

1000 type strains from various culture collections and candidate type strains for sp. nov. from various researchers. Draft genomes. DNA production ongoing

1000 selected type strains; Draft and PacBio genomes. DNA production ongoing
Planctomycetales – two genera become five genera

Complete genome sequence of Planctomyces brasilienensis type strain (DSM 5305T), phylogenomic analysis and reclassification of Planctomyces 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 2 Scanning-electron micrograph of P. brasilienensis DSM 5305T highlighting stalks and crateriform structures on the cell surface.

Figure 5 Venn diagram depicting the intersections of sets of homologous proteins of P. maris, P. brasilienensis, 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].
*Escherichia coli* and subspecies

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

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 overlap 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 highly efficient and informative analyses of unexplored biodiversity.

- Culturomics might push the speed of accessing 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.