

# Metagenómica, Metatranscriptómica y Microbioma Humano

Bioinformática para biotecnología BIT120

20 abril 2017

Eduardo Castro-Nallar, PhD

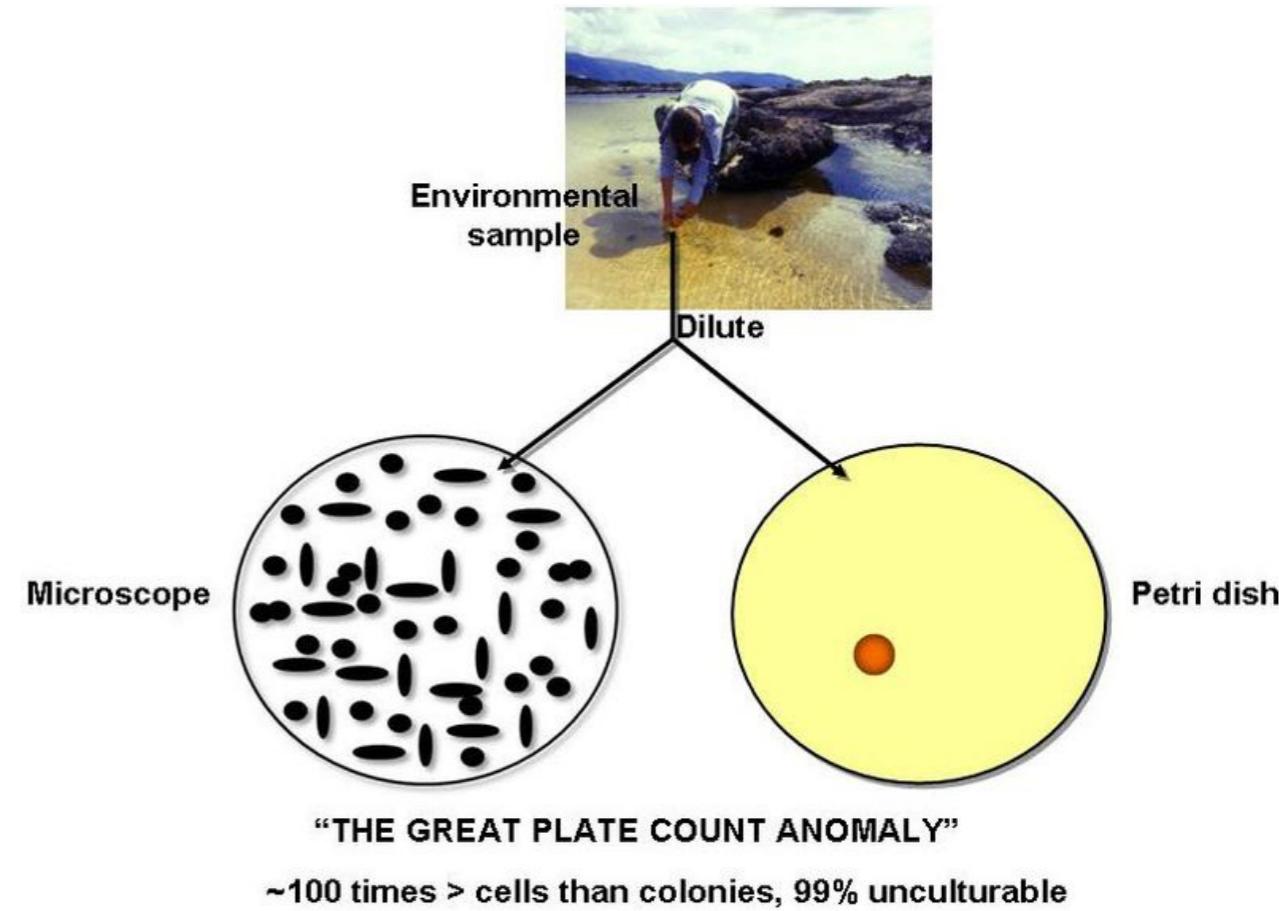
Center for Bioinformatics and Integrative Biology

[www.cbib.cl](http://www.cbib.cl)

[www.castrolab.org](http://www.castrolab.org)

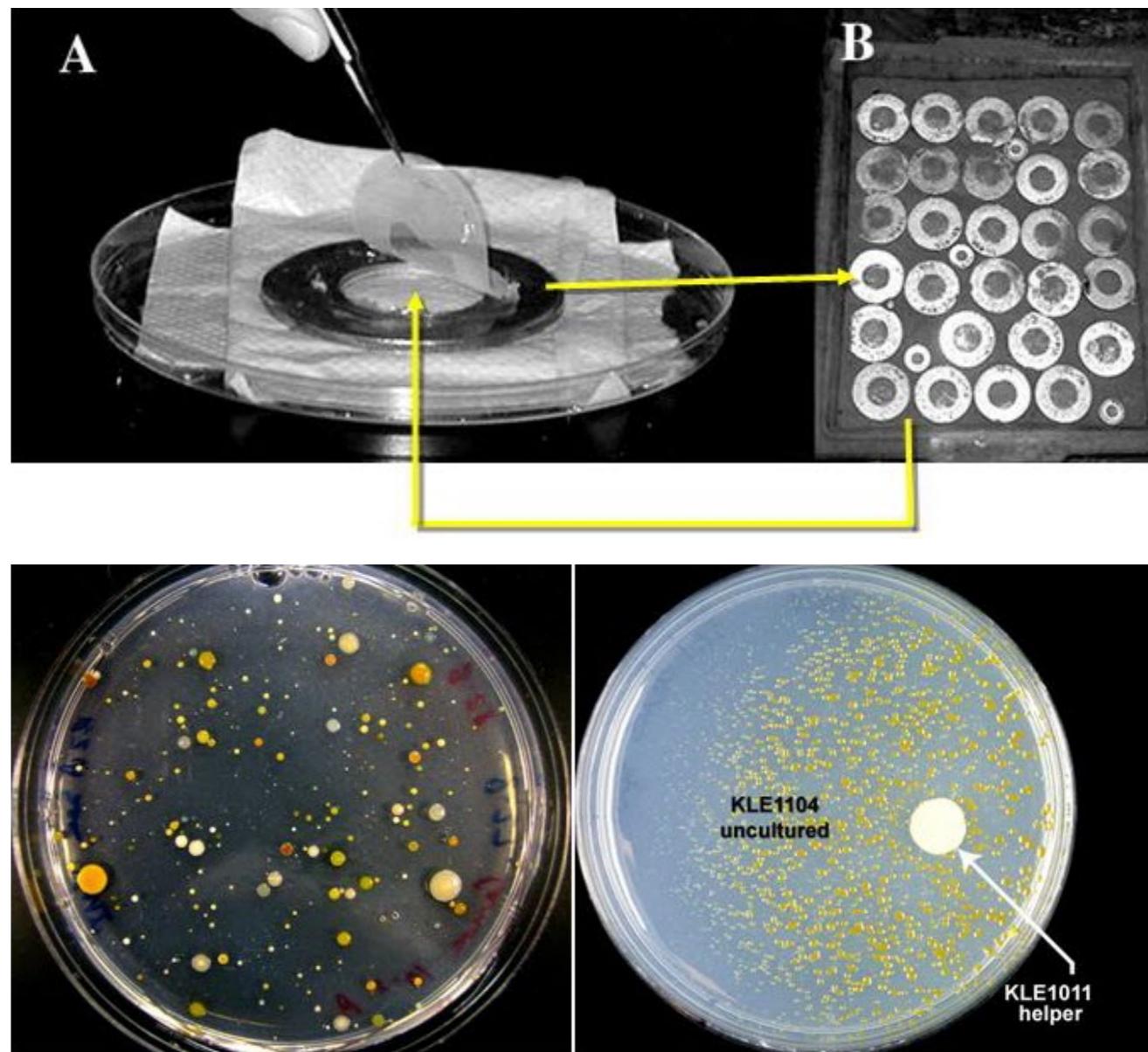
# Los microorganismos son incultivables (la mayoría)

- “The Great Plate Count Anomaly”
- Toma una muestra de suelo, mezcla con agua, vórtex, dejar decantar, diluye el sobrenadante y toma dos gotas
- Una gota a placa de Petri con medio LB, y la otra gota al microscopio óptico
- 100:1



# Los microorganismos son incultivables (la mayoría)

- Paradoja: genomas similares, uno crece y otro no. Deberían crecer en el mismo medio.  
*Bacillus*; *Pseudomonas*
- No solo no conocemos sus requerimientos nutricionales
- Crecer microorganismos en su ambiente natural
- “Factores de crecimiento”



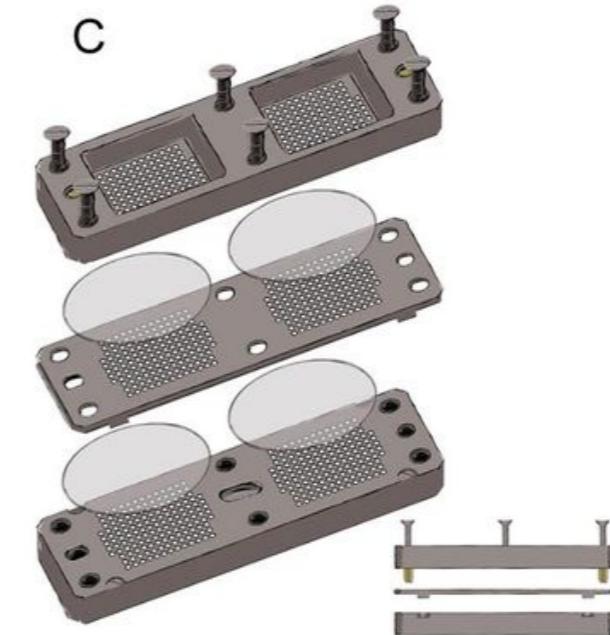
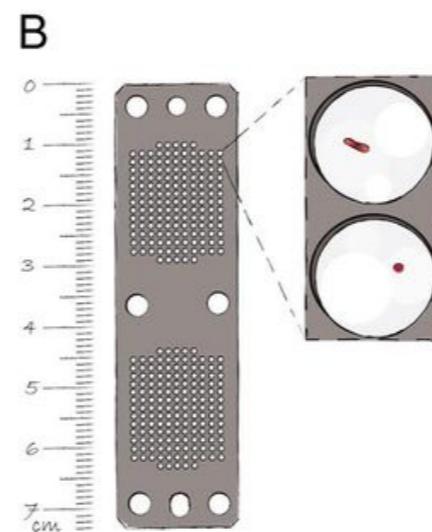
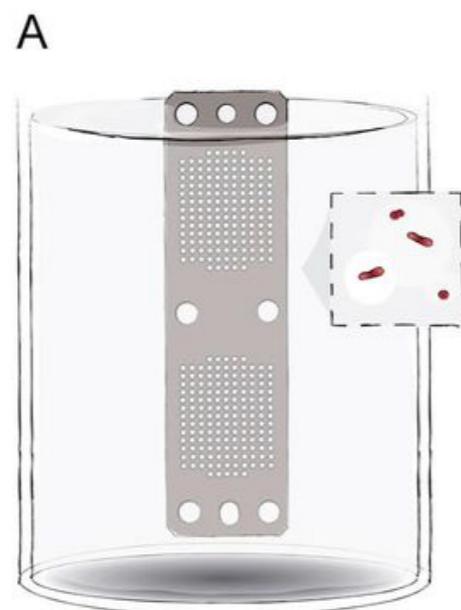
*Maribacter polysiphoniae* sp.  
99.8% identity by 16S rRNA gene

*M. luteus* sp.

# Los microorganismos son incultivables (la mayoría)

## Aternativas

- “Persuadirlos” para que crezcan en el lab
- Crecerlos *in situ*
- iChip



# Importancia

- Microorganismos manejan y sostienen toda la vida en el planeta
- Ciclo del nitrógeno, carbono, azufre
- Conversión de elementos y compuestos inorgánicos en compuestos orgánicos accesibles a otros organismos
- 99% no cultivable. Grupos taxonómicos completos no contienen un solo representante cultivable
- ¿Cuál es la significancia de la microbiología si estamos limitados a microorganismos cultivables?

# Importancia

- Fotosíntesis, fijación de nitrógeno, biodegradación y producción primaria



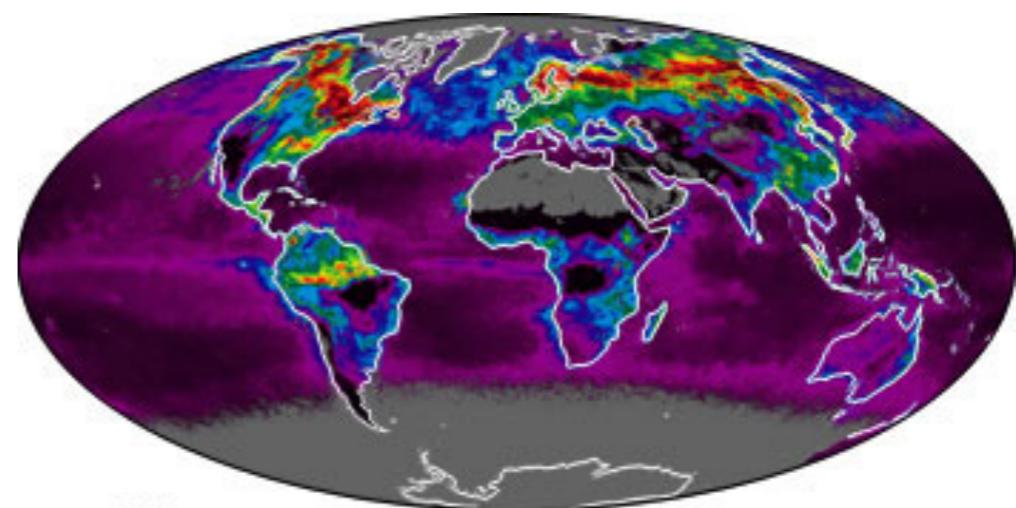
$N_2$  a compuestos orgánicos  
nitrogenados en suelo



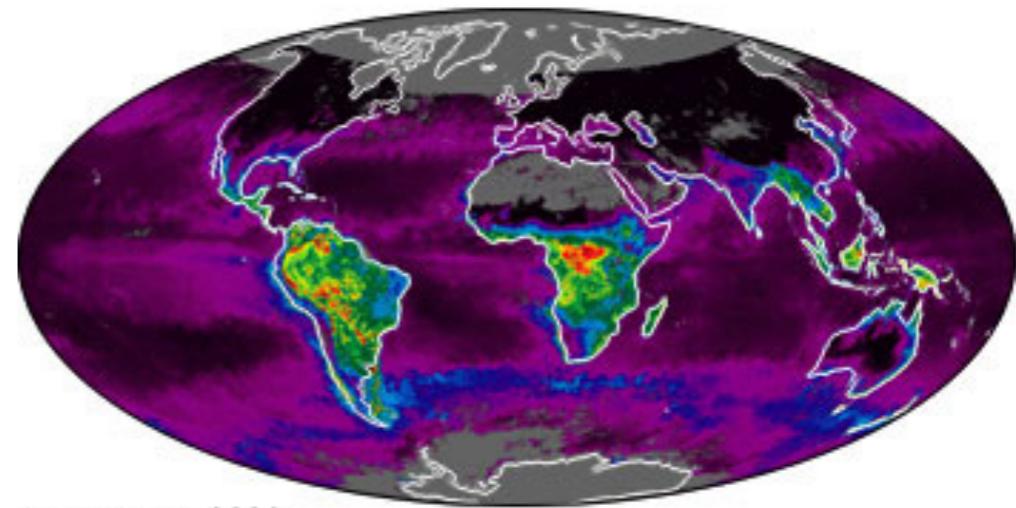
*Synechococcus*  
50% del  $O_2$   
mundial



Kg de C fijado



June 2002



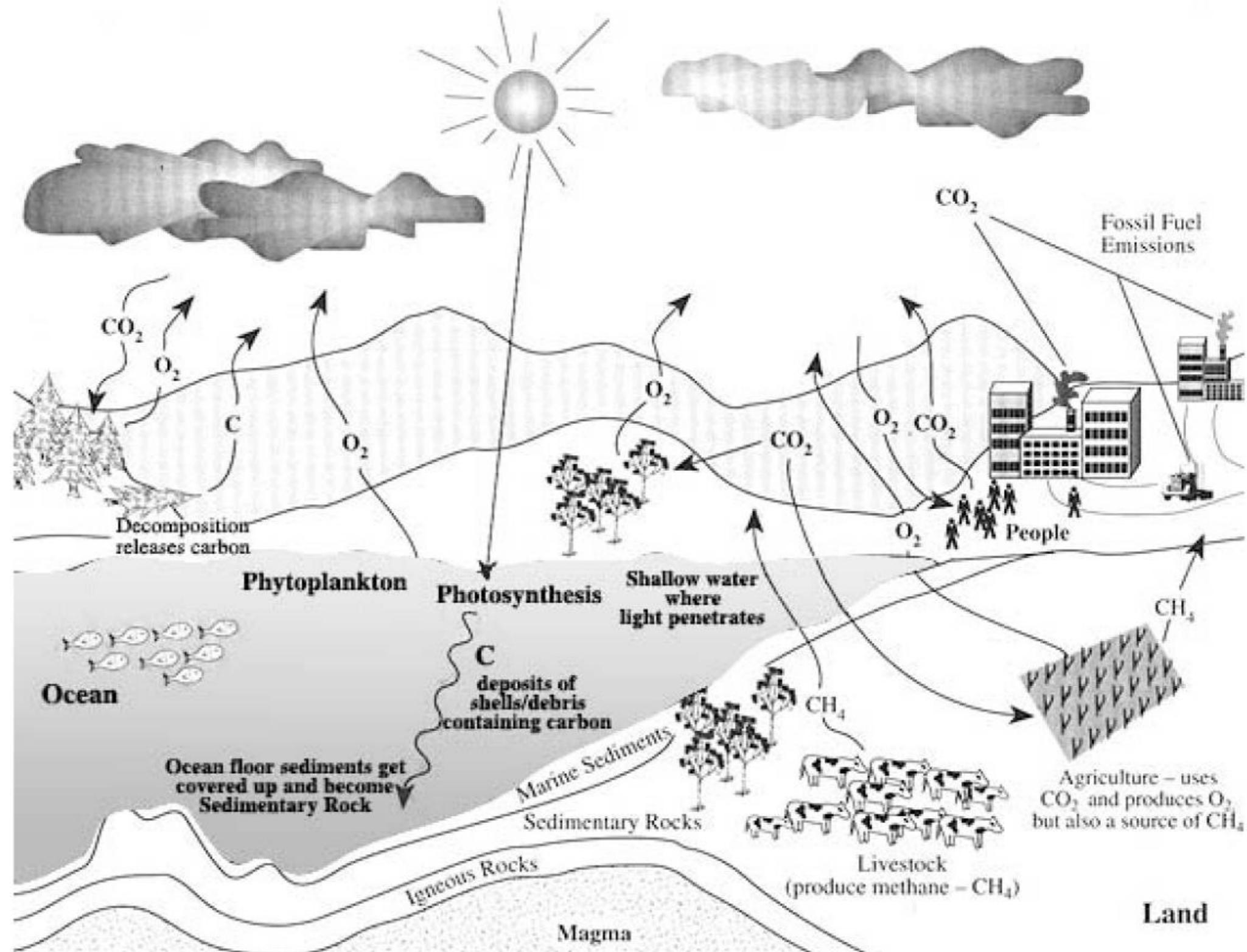
December 2002



Enzimas microbianas degradan  
compuestos orgánicos

# Importancia

Presentes en el  
ciclo del  
Carbono → en  
todas partes



# Solución

- Nudo = ¿Cómo acceder a la mayoría de la biodiversidad?
- Si no podemos desatar el nudo, simplemente lo cortamos = metagenómica
- Nos da acceso a la diversidad de especies y funciones metabólicas que existen



# Metagenomas: Genomas de múltiples orígenes

- *Meta* = más allá
- El estudio de los genomas en un medio ambiente más allá del organismo individual

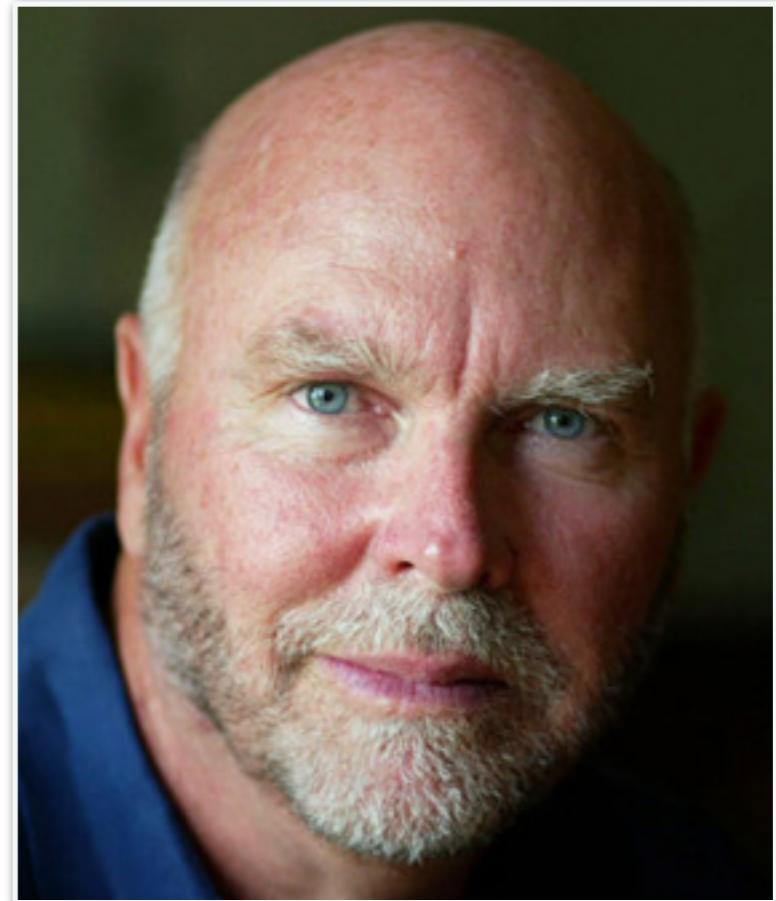
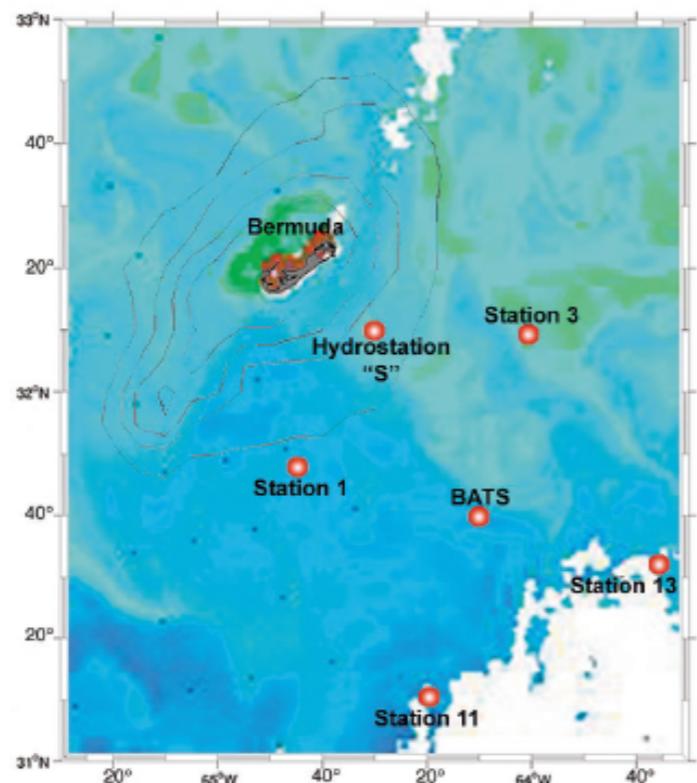
a prefix appearing in loanwords from Greek, with the meanings "after," "along with," "beyond," "among," "behind," and productive in English on the Greek model:  
*metacarpus; metagenesis.*

**Metagenomics** is the study of genetic material recovered directly from environmental samples. The broad field may also be referred to as environmental genomics, ecogenomics or community genomics.



# Primer estudio metagenómico

- Acuñado en ~1998
- J Craig Venter en 2004
- Antes del secuenciamiento masivo
- Mar del Sargasso cerca de Bermuda



## RESEARCH ARTICLE

### Environmental Genome Shotgun Sequencing of the Sargasso Sea

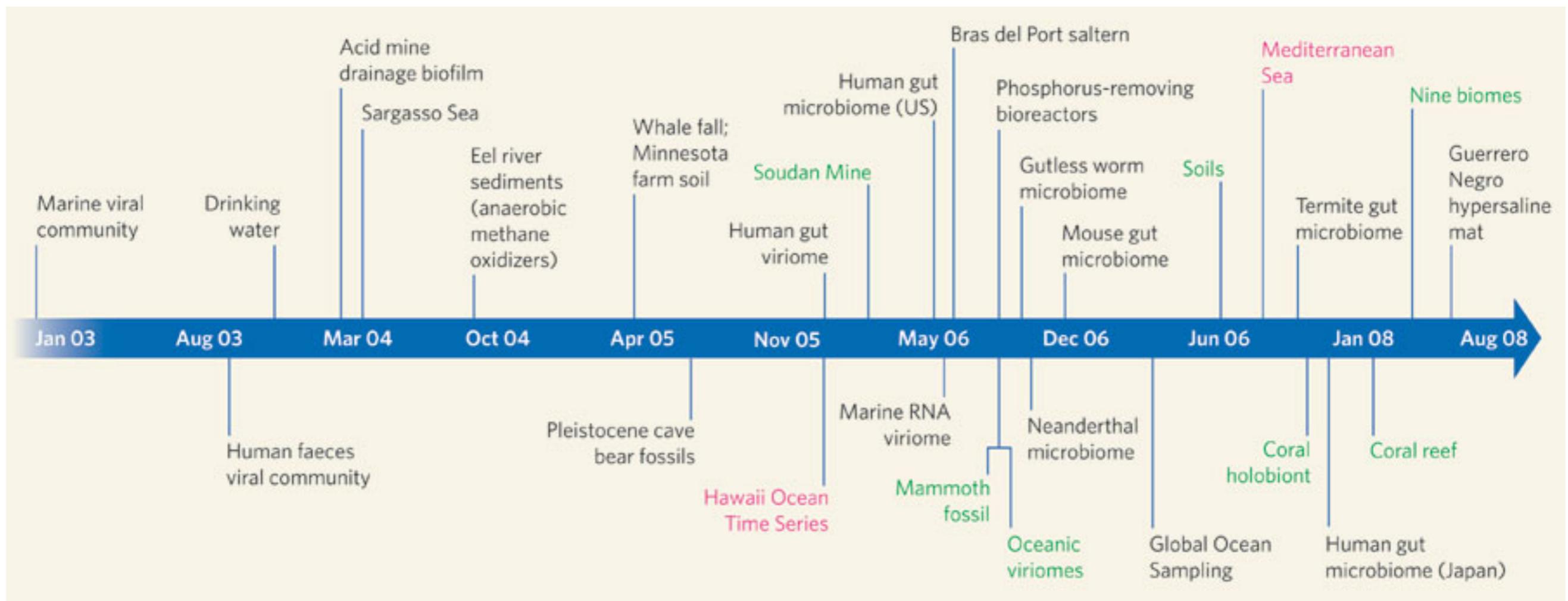
J. Craig Venter<sup>1,\*</sup>, Karin Remington<sup>1</sup>, John F. Heidelberg<sup>3</sup>, Aaron L. Halpern<sup>2</sup>, Doug Rusch<sup>2</sup>, Jonathan A. Eisen<sup>3</sup>, Dongying Wu<sup>3</sup>, Ian Paulsen<sup>3</sup>, Karen E. Nelson<sup>3</sup>, William Nelson<sup>3</sup>, Derrick E. Fouts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

+ Author Affiliations

✉ To whom correspondence should be addressed. E-mail: jcventer@tca.org

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

# Línea de tiempo metagenómica



- ★ The different technologies used are dye-terminator shotgun sequencing (black), fosmid library sequencing (pink) and pyrosequencing (green)

# Aplicaciones

# Búsqueda de enzimas industriales: bioprospección

- “Cosechar” la biodiversidad funcional de microorganismos
- “Screening” funcional o por identidad de secuencia

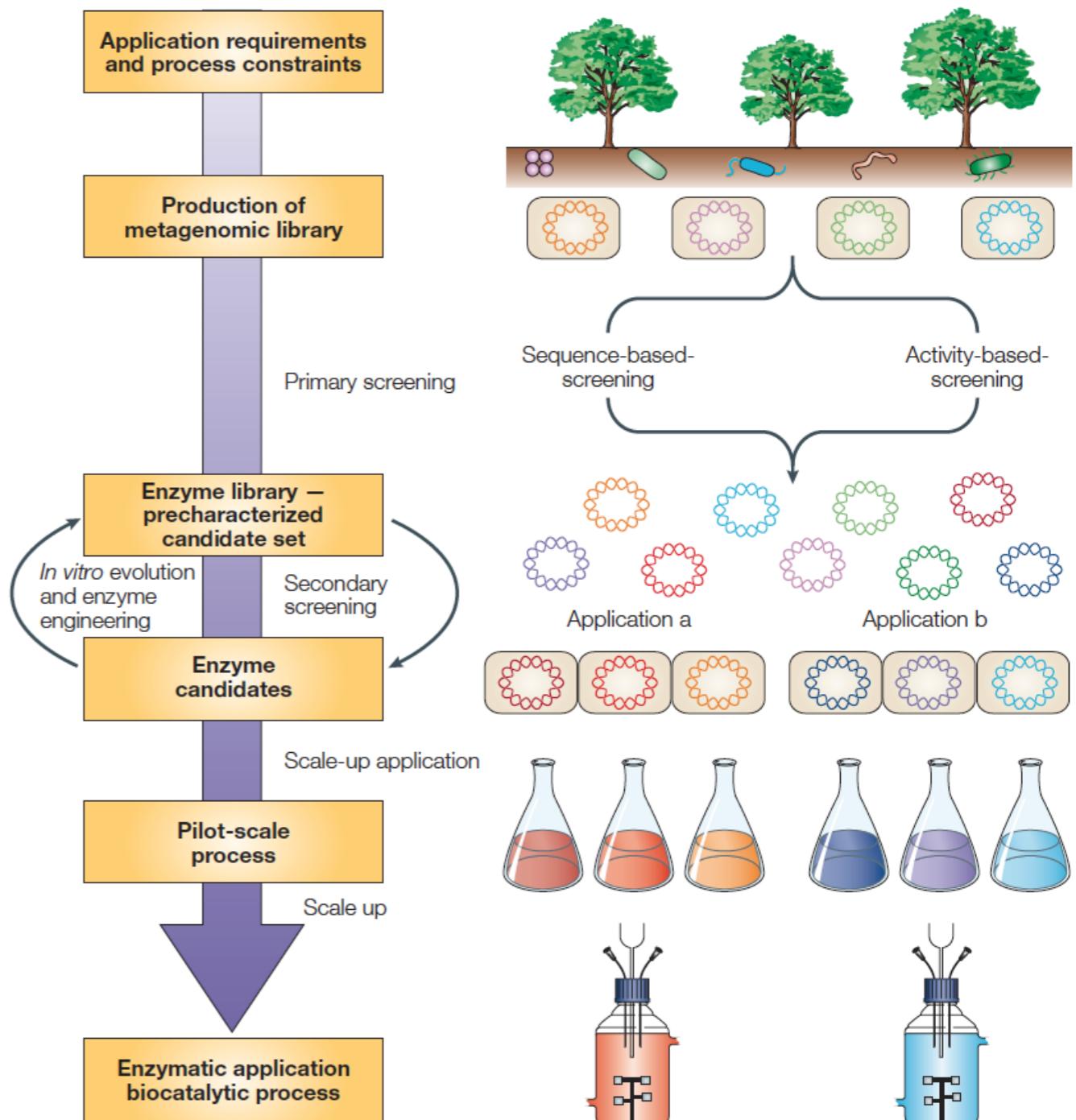


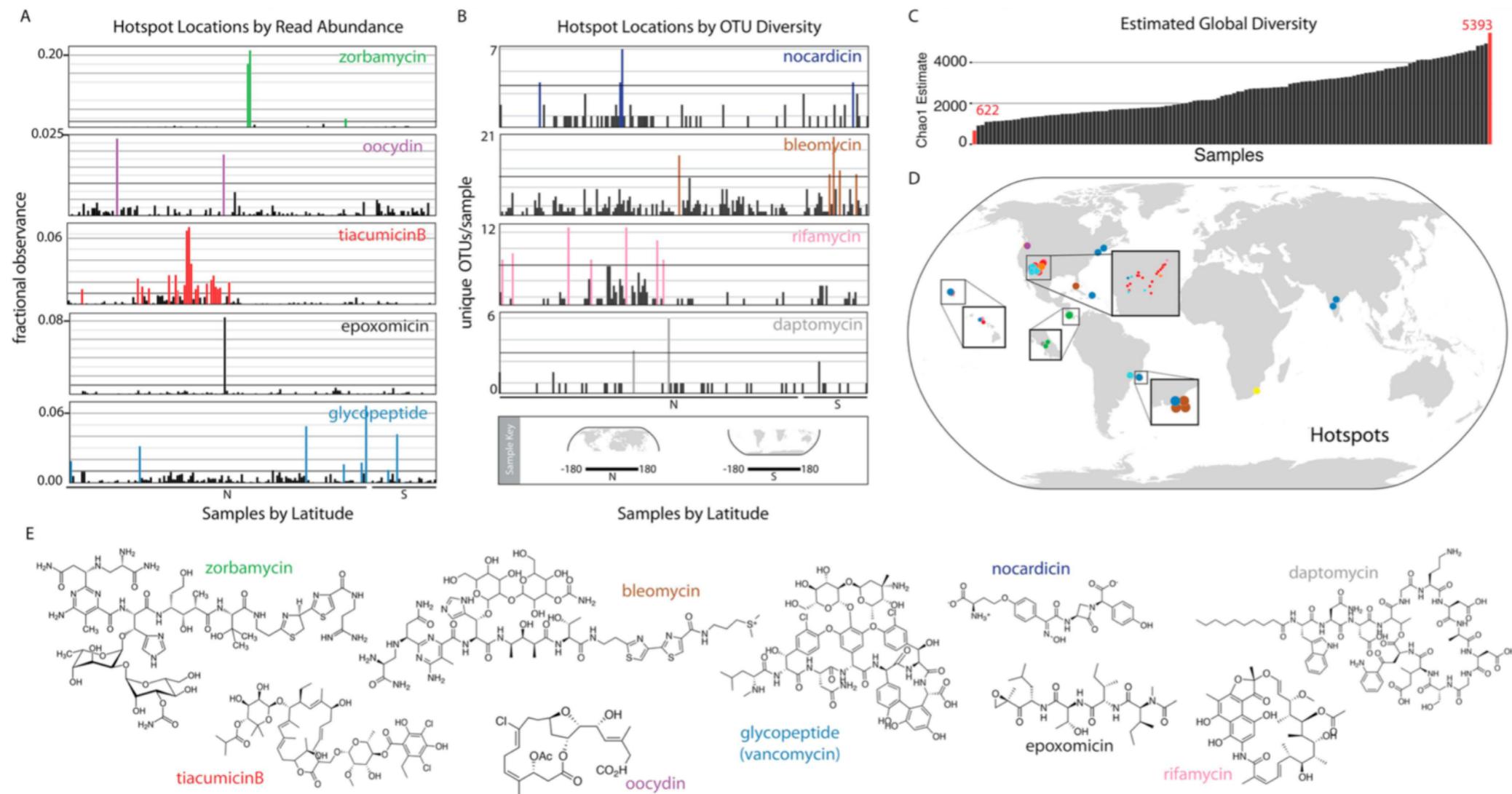
Table 1 | Activity-based screening for industrially relevant enzymes and biocatalysts from metagenomic libraries

Function	Habitat	Library type	Average insert size (kb)	Number of clones screened	Library size (Mb)	Substrate	Number of hits	Hit rate (hit per Mb)	Ref.
Esterase/lipase	Forest soil	Plasmid	8	67,000	536	Tributyrin	98	1/5.5	*
Esterase/lipase	Forest soil	Fosmid	40	19,968	799	Tributyrin	47	1/17	*
Esterase/lipase	Sandy ecosystem	Fosmid	30	29,884	903	Tributyrin	49	1/18	*
Esterase/lipase	Sandy ecosystem	Fosmid	40	25,344	1,014	Tributyrin	29	1/35	*
Esterase/lipase	Soil	Plasmid	6	286,000	1,716	Tributyrin	3	1/572	63
Esterase/lipase	Soil	Plasmid	6	730,000	4,380	Triolein	1	1/4,380	63
Esterase/lipase	Soil	BAC	27	3,648	100	Bacto Lipid	2	1/50	64
Oxidation of polyols	Soil	Plasmid	3	900,000	2,700	1,2-ethanediol; 1,2-propanediol; 2,3-butanediol	15	1/180	65
Alcohol oxidoreductase	Soil/enrichment	Plasmid	4	400,000	1,600	Glycerol/1,2-propanediol	10	1/160	66
Amidase	Soil/enrichment	Plasmid	5	193,000	965	D-phenylglycine-L-leucine	7	1/138	67
Amylase	Soil	Plasmid	5	80,000	400	Starch	1	1/400	68
Amylase	Soil	BAC	27	3,648	100	Starch	8	1/12	64
Biotin production	Soil/excrement enrichment	Cosmid	35	50,000	1,750	Biotin-deficient medium	7	1/250	69
Protease	Soil	Plasmid	10	100,000	1,000	Skimmed milk	1	1/1,000	70
Cellulase	Sediment enrichment	λ phage	6	310,000	1,860	Carboxymethyl-cellulose	3	1/620	71
Chitinase	Seawater	λ phage	5	825,000	4,125	Methylumbelliferyl-diacylchitobioside	11	1/375	72
Dehydratase	Soil/sediment enrichment	Plasmid	4	560,000	2,240	Glycerol	2	1/1,120	38
4-Hydroxybutyrate conversion	Soil	Plasmid	6	930,000	5,580	4-Hydroxybutyrate	5	1/1,116	73
β-Lactamase	Soil	Plasmid	5	80,000	400	Ampicillin	4	1/100	68

The screening host for all studies was *Escherichia coli*. \*Unpublished results, K. Liebeton et al., BRAIN AG. BAC, bacterial artificial chromosome.

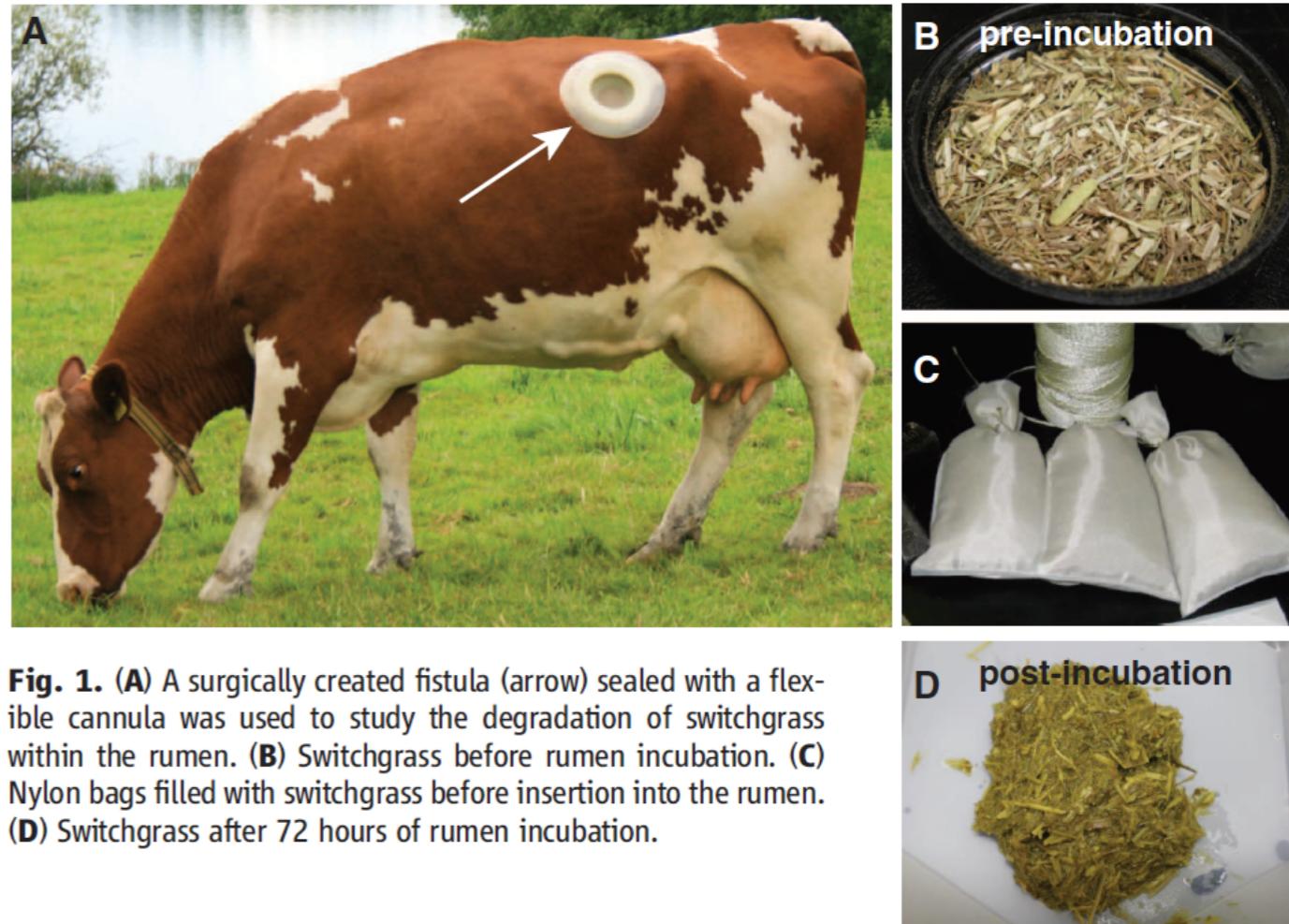
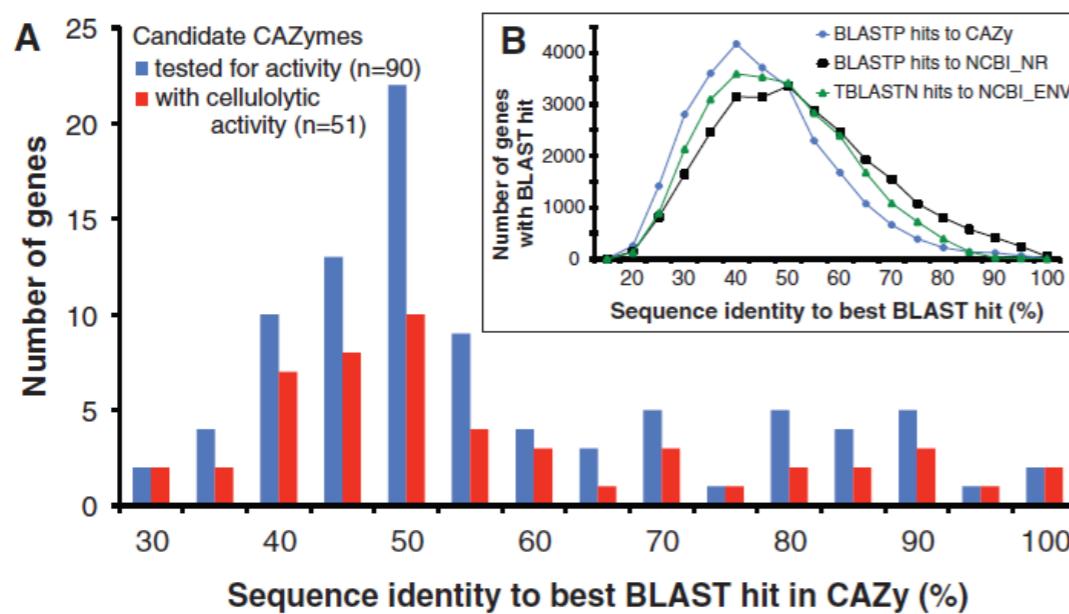
# “Screening” global

- Metabolismo secundario de bacterias
- ¿Dónde se encuentran “hotspots” de producción de moléculas de relevancia médica?

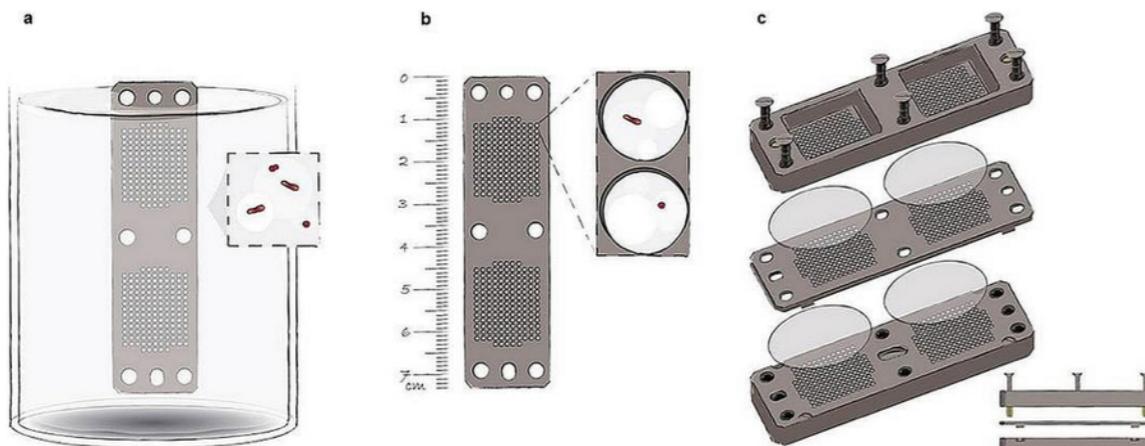


# Descubrir enzimas que degradan biomasa

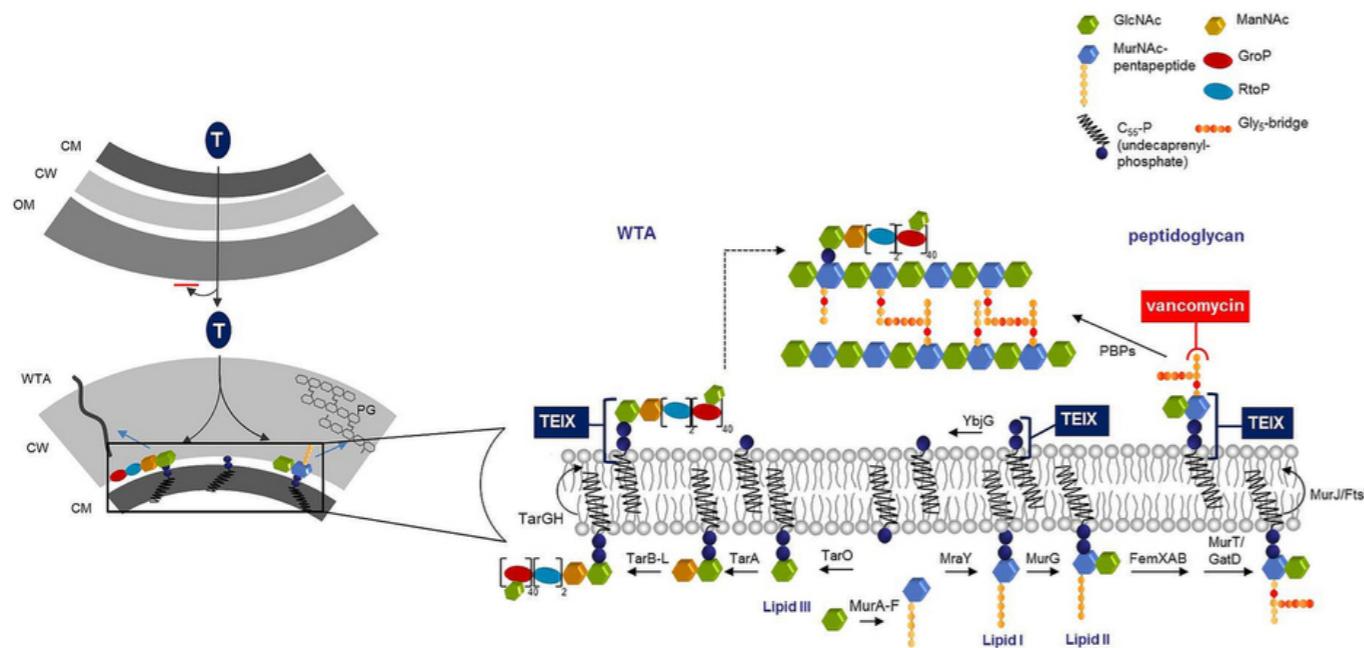
- Degradación de polisacáridos de plantas
- Producción de celulosa y combustible



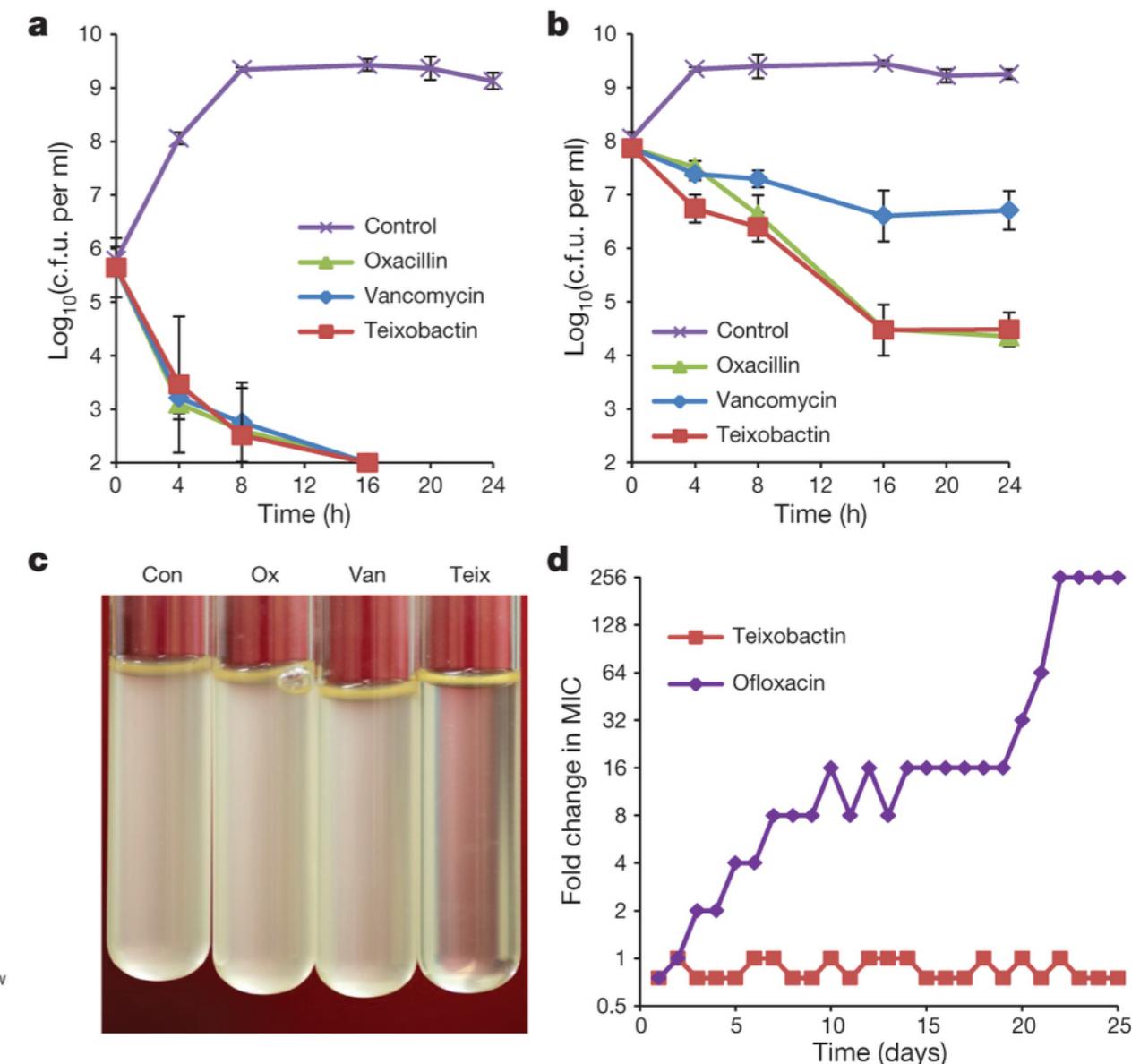
# Descubrir antibióticos



Dispositivo



Propuesta mecanismo



Pruebas funcionales

# Bio-minería

- BioSigma
- Identificación de genes y microorganismos
- Estrategia de *Systems Biology*

 Home About us Products & Services News Intellectual Property Work with us Suppliers Contact

[PRODUCTOS Y SERVICIOS >](#)

## System Biology and Microbiology



Bioleaching is a phenomenon of dissolution of minerals where the microorganisms responsible for the bioleaching activity are strongly affected by the process conditions. Comprehensive and integrated studies of each of their associated biological components provide information on how these organisms behave upon a variety of conditions and reveal clues of how to control the process to improve its performance and efficiency.

Precisely Systems Biology is the study of biological systems or components such as molecules, cells or whole organisms, internally or externally, interacting with each other and with their environment.

Due to the complex and dynamic behavior of these living systems, we use tools that integrate the information to predict the individual properties of the parts, understanding the functional features of these systems and their relationship with the environment. The analysis tools include the "omics" disciplines such as genomics, proteomics, metabolomics and bioinformatics.

BioSigma has a highly skilled team of experts on the use "omic" tools oriented to the bioleaching of low copper grade ores. We are equipped with a modern platform that is used on applied research projects and also on the continuous improvement of the monitoring and control of mining industrial processes. By January 2013, the company's achievements on these areas include the generation of intellectual property with 18 patent applications in Chile, also submitted to other 9 countries and with a total of 57 patents granted, which represent an important part of our technological innovations.

-  Bioleaching Technologies
-  Biomass Production and Inoculation
-  Biocharacterization Services
-  Laboratory of Biohydrometallurgy and Chemical Characterization
-  Project Development for Mining Resources Benefit
-  Bioremediation of Mining Products
-  Mineral Processing and Operation in Heaps and Dumps
-  System Biology and Microbiology

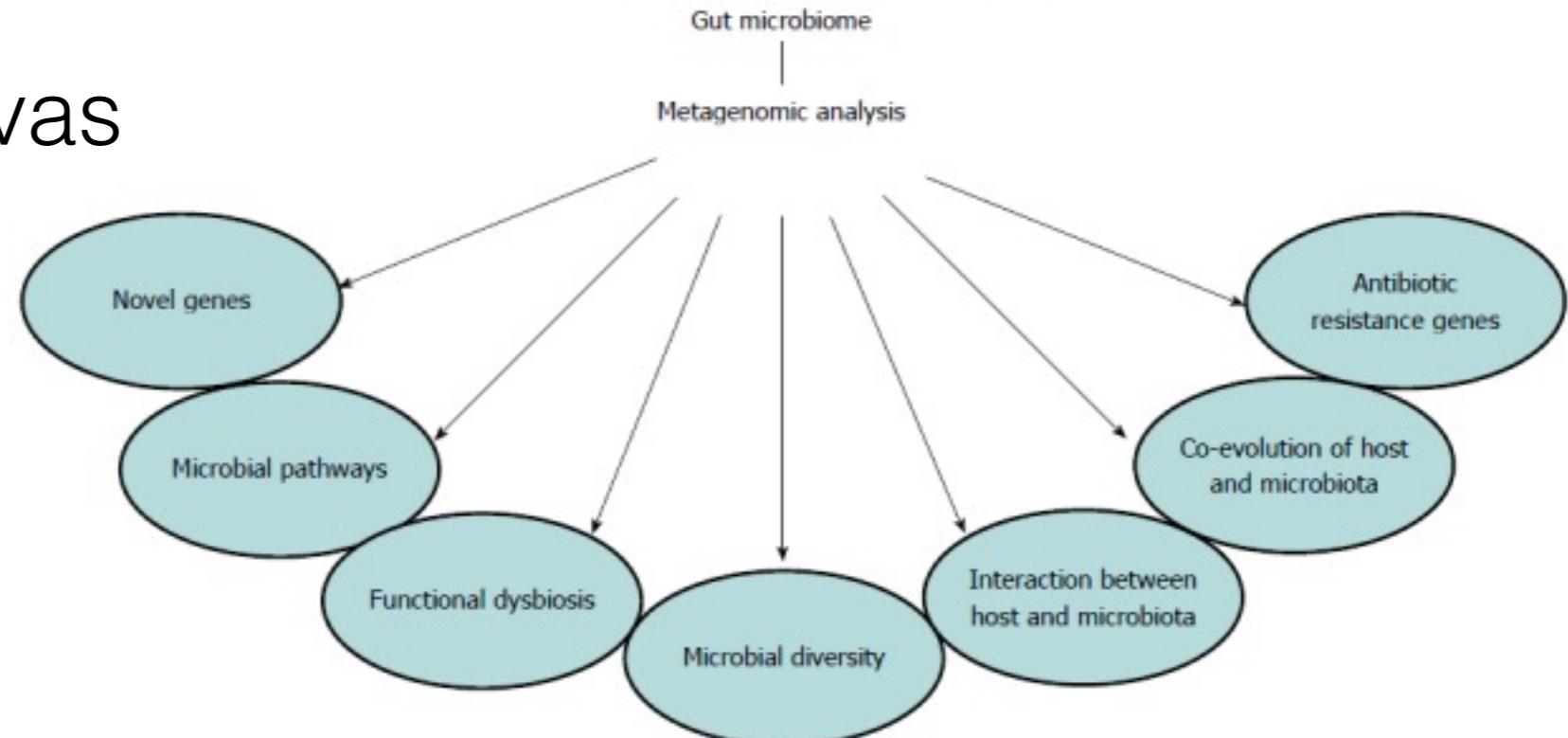
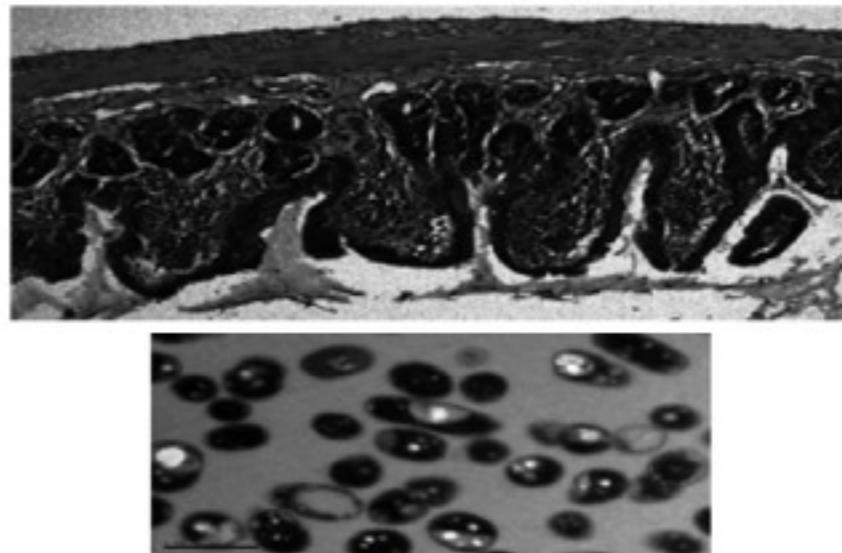
# Agricultura

- UCD-Chile
- Caracterización de sistemas de rizósfera en viñas, e.g., *terroir*
- Compuestos bacterianos que promueven características deseables en el vino



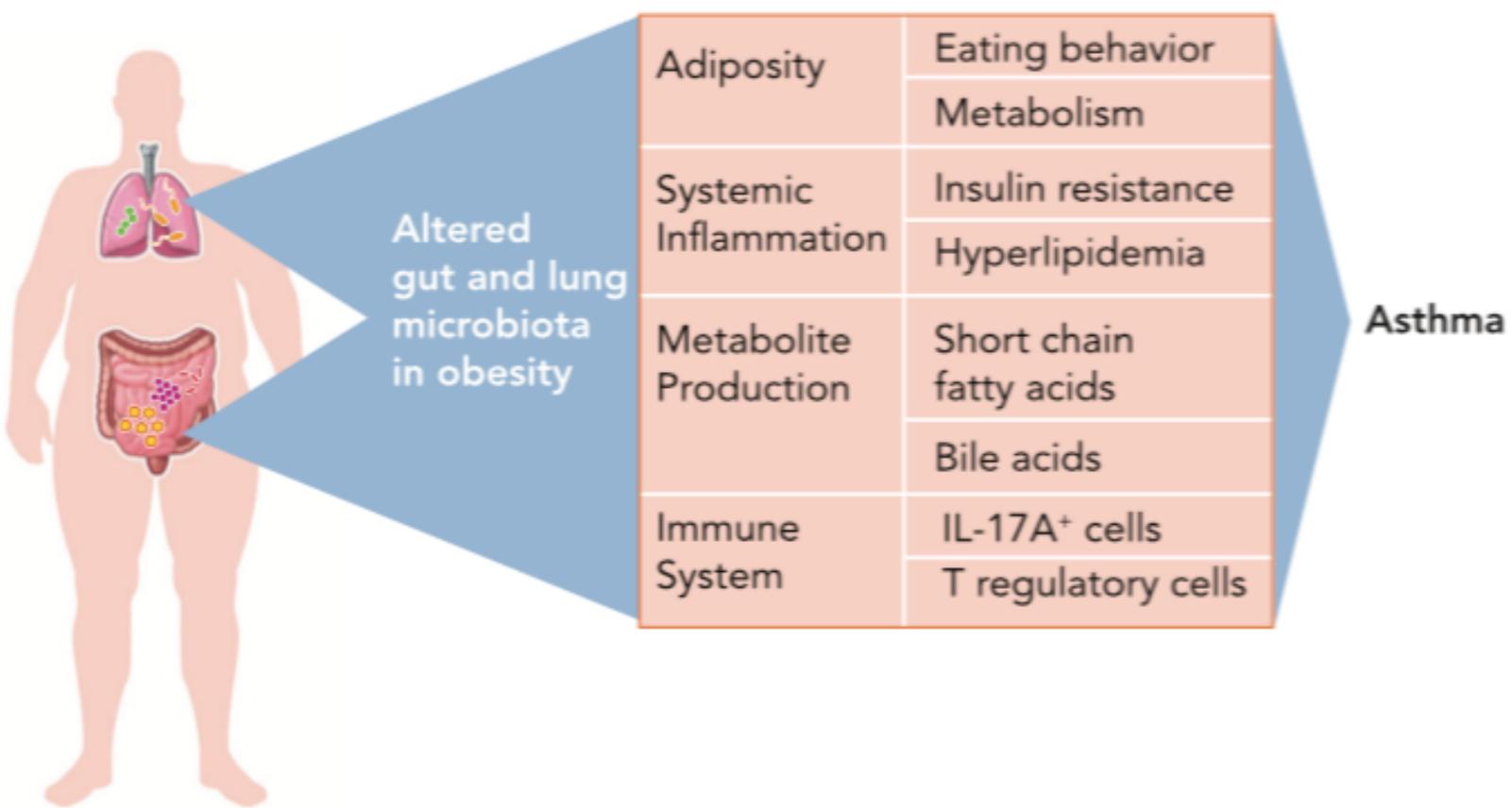
# Aplicación en microbiota intestinal

- El ecosistema es el individuo
- Asociaciones relativas a patologías y condiciones

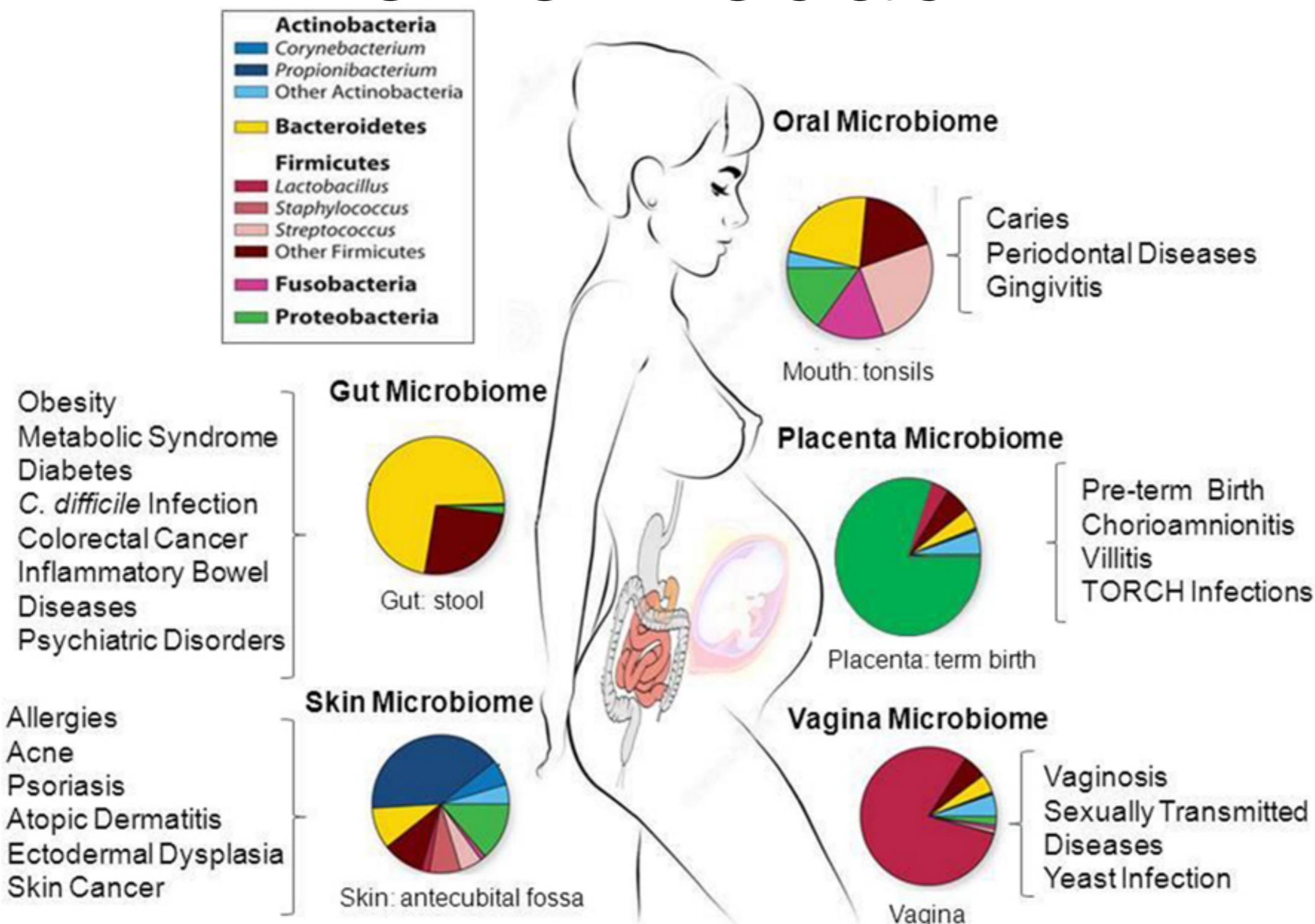


# Microbiota humana y enfermedad

- Asociaciones con patologías y condiciones
- Obesidad, alergias, asma, depresión
- Oportunidades terapéuticas

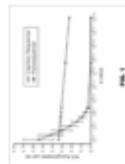


# Microbiota humana y enfermedad



# Importancia biotecnológica: Patentes

Gut microbiome as a biomarker and therapeutic target for ...



[www.google.com/patents/US20100172874](http://www.google.com/patents/US20100172874)

App. - Filed 10 Dec 2007 - Published 8 Jul 2010 - Peter J. Turnbaugh - The Washington University

The present invention relates to the gut microbiome as a biomarker and therapeutic target for energy harvesting, weight loss or gain, and/or ...

[Overview](#) - [Related](#) - [Discuss](#)

Compositions comprising a defined microbiome and methods ...



[www.google.com/patents/WO2015051323A1?cl=en](http://www.google.com/patents/WO2015051323A1?cl=en)

App. - Filed 3 Oct 2014 - Published 9 Apr 2015 - Frederic Bushman -

Trustees Of The University Of Pennsylvania

The invention relates to compositions comprising a defined microbiome and methods for using such compositions, including methods for replacing or supplementing a subject's gut microbiome, and methods for treating a disease, pathological condition, and/or iatrogenic condition of the gut.

**WO 2014152484 A1**

Microbiome



## ABSTRACT

The present invention provides freeze-dried compositions that include an extract of human feces and a cryoprotectant, and methods for making and using such compositions, including methods for replacing or supplementing or modifying a subject's colon microbiota, and methods for treating a disease, pathological condition, and/or iatrogenic condition of the colon.

Microbiome



[www.google.com/patents/WO2014152484A1?cl=en](http://www.google.com/patents/WO2014152484A1?cl=en)

App.

Second Genome, Inc.

In some embodiments, the quantitative measure is an index, such as, for example, a Microbiome Modulation Index (MMI). [0004] An aspect of ...

[Overview](#) - [Related](#) - [Discuss](#)

Methods of diagnosing and treating microbiome-associated ...



[www.google.com/patents/WO2011022660A1?cl=en](http://www.google.com/patents/WO2011022660A1?cl=en)

App. - Filed 20 Aug 2010 - Published 24 Feb 2011 - Bernat Olle - Puretech Ventures, Llc

Methods of diagnosing and treating microbiome-associated disease or improving health using interaction network parameters are provided.

[Overview](#) - [Related](#) - [Discuss](#)

Publication number	WO2014152484 A1
Publication type	Application
Application number	PCT/US2014/027391
Publication date	Sep 25, 2014
Filing date	Mar 14, 2014
Priority date	Mar 14, 2013
Inventors	Michael J. Sadowsky, 4 More »
Applicant	Regents Of The University Of Minnesota
Export Citation	BiBTeX, EndNote, RefMan
Patent Citations	(4), Classifications (10), Legal Events (3)

External Links: [Patentscope](#), [Espacenet](#)

# Importancia biotecnológica: Patentes

## Novel cellulase derived from metagenome, and preparation method therefor

WO 2014142529 A1

### ABSTRACT

The present invention relates to a novel cellulase derived from a metagenome, and a preparation method therefor, and more specifically, to: a cellulase (CelEx-CBR12) selected using a robot-based high-throughput screening system, from a metagenomic library extracted from suspended solids of bovine rumen; a gene encoding the same; and a method for preparing a recombinant cellulase. The cellulase CelEx-CBR12 selected in the present invention has exocellulase activity and endocellulase activity, and thus can be used for producing bioethanol using cellulosic biomass, and can be applied to various industries such as fibers, detergents, feed, food, pulp, paper production and the like.

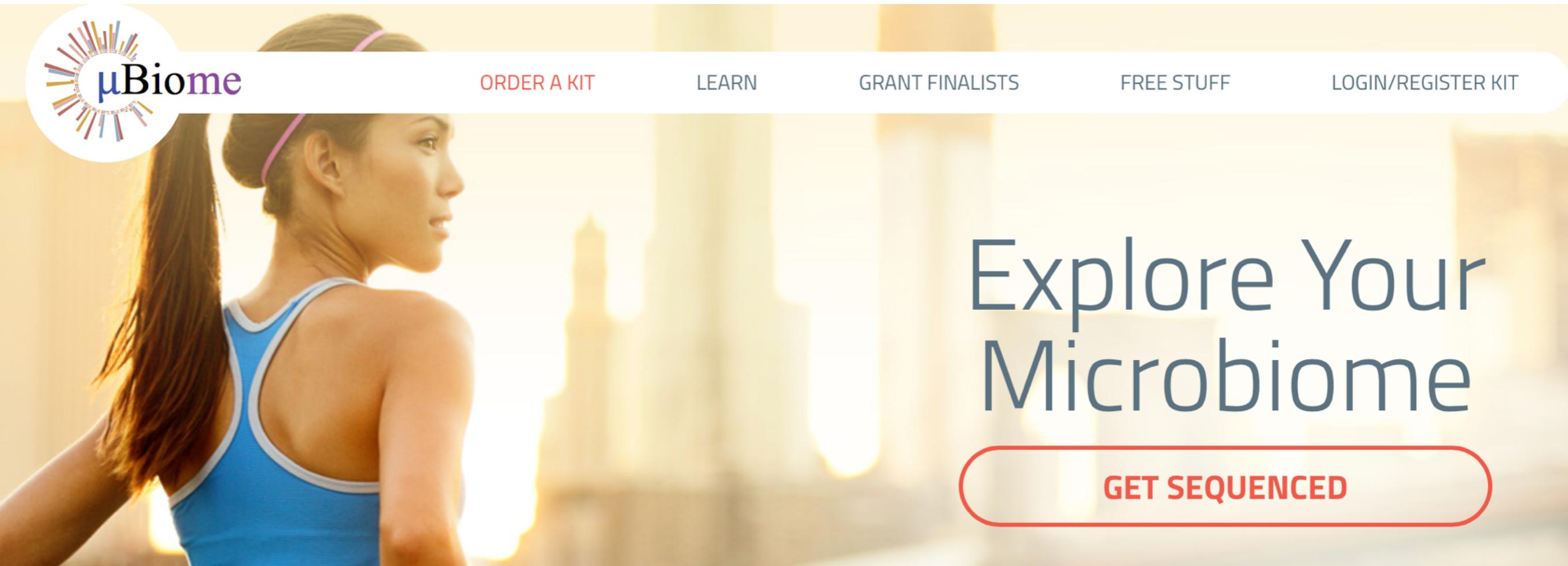
Publication number	WO2014142529 A1
Publication type	Application
Application number	PCT/KR2014/002024
Publication date	Sep 18, 2014
Filing date	Mar 11, 2014
Priority date	Mar 12, 2013
Inventors	Jae Jun Song, 송재준, Jong Hyun Choi, 최종현, Kyong-Cheol KO, 고경철, Yun Jon Han, 한윤전
Applicant	Korea Research Institute Of Bioscience, 한국생명공학연구원
Export Citation	<a href="#">BiBTeX</a> , <a href="#">EndNote</a> , <a href="#">RefMan</a>
Non-Patent Citations (7)	<a href="#">(7)</a>
Classifications (5)	<a href="#">(5)</a>
Legal Events (3)	<a href="#">(3)</a>

External Links: [Patentscope](#), [Espacenet](#)

### Recent patents related to microbial genomics and metagenomics

Patent number	Description	Assignee	Inventor	Date
US 9,192,361	Preparation method for a fecal microbiome transplant material and apparatus for the preparation of fecal material for a fecal transplant. A healthy patient's stool is collected and deposited in a first container, which has its open end attached to a second container to form a fecal processing device. Blender blades are positioned in the second container and a filter is mounted therein. Liquid is injected into the second container through a syringe port having a valve therein to block the escape of liquid from the connected containers. Intake and exhaust ports allow air to enter and escape from the connected containers while blocking the escape of liquid. The liquid and fecal matter in the connected containers is blended and the blended liquid is strained through the filter in the second container and removed through the syringe port for use in a fecal transplant.	Stevens CJ	Stevens CJ	11/24/2015
US 9,116,839	Methods and systems for identifying horizontally transferred genes, including defining a cuboid in a three-dimensional space, wherein the cuboid includes fragment points corresponding to the genomic fragments belonging to a plurality of sequenced microbial genomes, and dividing the cuboid into a plurality of grids. Then, selecting one or more grids corresponding to a selected genome and classifying each of the selected grids as one of majority, minority and mixed grids, based on number of fragment points corresponding to the selected genome in each of the selected grids. Further, at least one genomic fragment from the minority and the mixed grids is identified as the horizontally transferred gene based on a distance ratio assessment.	Tata Consultancy Services (Mumbai)	Mande SS, Mehra V, Ghosh TS	8/25/2015
US 9,080,211	Methods, compositions and kits for using a transposase and a transposon end for generating extensive fragmentation and 5'-tagging of double-stranded target DNA <i>in vitro</i> , then using a DNA polymerase for generating 5'- and 3'-tagged single-stranded DNA fragments without performing a PCR amplification reaction, wherein the first tag on the 5'-ends exhibits the sequence of the transferred transposon end and optionally, an additional arbitrary sequence, and the second tag on the 3'-ends exhibits a different sequence from the sequence exhibited by the first tag. The method is useful for generating 5'- and 3'-tagged DNA fragments for use in a variety of processes, including processes for metagenomic analysis of DNA in environmental samples, copy number variation analysis of DNA and comparative genomic sequencing, including massively parallel DNA sequencing.	Epicentre Technologies (Madison, WI, USA)	Grunenwald HL, Caruccio N, Jendrisak J, Dahl G	7/14/2015
US 9,040,101	Compositions and methods for treating diabetes, in particular, utilizing a gastrointestinal microbiome modulating composition.	MicroBiome Therapeutics (Broomfield, CO, USA)	Heiman ML, Stull DP, Peno JW	5/26/2015
US 8,980,553	Compositions and methods for enriching nontarget polynucleotides from a mixture of nontarget and target polynucleotides where differences between the target polynucleotides and the nontarget polynucleotides include the extent of modified bases that are present in a greater density in the target polynucleotides than in the nontarget polynucleotides. Useful in removing human genomic DNA from a mixture of DNAs obtained from human tissue samples to enrich for polynucleotides in a microbiome so as to characterize the microbiome by DNA sequencing.	New England Biolabs (Ipswich, MA, USA)	Feehery GR, Stewart F, McFarland J, Pradhan S	3/17/2015
US 8,865,410	Compositions and methods for preparing DNA sequencing libraries, in particular preparing DNA sequencing libraries from kilobase-scale nucleic acids. Also, methods for assembling short-read sequencing data into longer contiguous sequences; useful for various applications in genomics, including genome assembly, full-length cDNA sequencing, metagenomics and the analysis of repetitive sequences of assembled genomes.	University of Washington (Seattle)	Shendure J, Hiatt J, Patwardhan R, Turner E	10/21/2014
US 8,846,347	Methods for preparing DNA sequencing libraries by assembling short-read sequencing data into longer contiguous sequences for genome assembly, full-length cDNA sequencing, metagenomics and the analysis of repetitive sequences of assembled genomes.	University of Washington (Seattle)	Shendure J, Hiatt J, Patwardhan R, Turner E	9/30/2014
US 8,383,345	Methods for preparing DNA sequencing libraries for assembling short-read sequencing data into longer contiguous sequences and providing for genome assembly, full-length cDNA sequencing, metagenomics and the analysis of repetitive sequences of assembled genomes, generally comprising incorporating adaptor or vector sequences into at least one member of a target fragment library, amplifying the population of target fragment library members to produce a plurality of copies of amplified DNA molecules, fragmenting the plurality of amplified DNA molecules, incorporating additional adaptor sequences, and amplifying a region of at least one of the plurality of amplified DNA molecules.	University of Washington (Seattle)	Shendure J, Hiatt J, Patwardhan R, Turner E	2/26/2013

# Empresas - directo al consumidor



**μBiome**

ORDER A KIT    LEARN    GRANT FINALISTS    FREE STUFF    LOGIN/REGISTER KIT

Explore Your  
Microbiome

**GET SEQUENCED**

# Empresas - productos terapéuticos - diabetes



The image shows a screenshot of the MicroBiome therapeutics website. At the top left is the company logo, featuring a stylized globe icon followed by the text "MicroBiome therapeutics". A blue navigation bar below the logo contains links for "Home", "About Us »", "The Company »", "Technology »", "Business Development", "News »", "Resources »", and "Contact Us »". To the right of the navigation bar is a search bar with the placeholder text "search this site..." and a magnifying glass icon. The main content area features a large photograph of a gloved hand holding a petri dish with bacterial colonies against a background of DNA gel electrophoresis patterns. At the bottom of the page, a dark banner contains the text: "At MicroBiome Therapeutics, we are actively engaged in clinical studies of our lead product, NM504, in prediabetic and Type 2 diabetic patients. In Type 2 diabetes we are developing an adjunctive therapy to metformin, the most-widely prescribed first-line therapy for newly diagnosed patients."

# Empresas

**MIT  
Technology  
Review**

[Log in / Register](#) [Search !\[\]\(9ec46ccf39110b98e9de4be0362c59b6\_img.jpg\)](#)

[Topics+](#) [The Daily](#) [Magazine](#) [Business Reports](#) [More+](#)

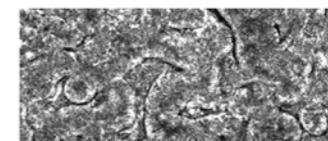
[Subscribe !\[\]\(8fef531e8b487e8fdd2f7578a0eb4330\_img.jpg\)](#)

**Biomedicine**

## Companies Aim to Make Drugs from Bacteria That Live in the Gut

**Scientific discoveries in recent years suggest that some serious conditions could be cured by adding “good” bacteria to your digestive tract.** Now several companies are racing to develop drugs that do so.

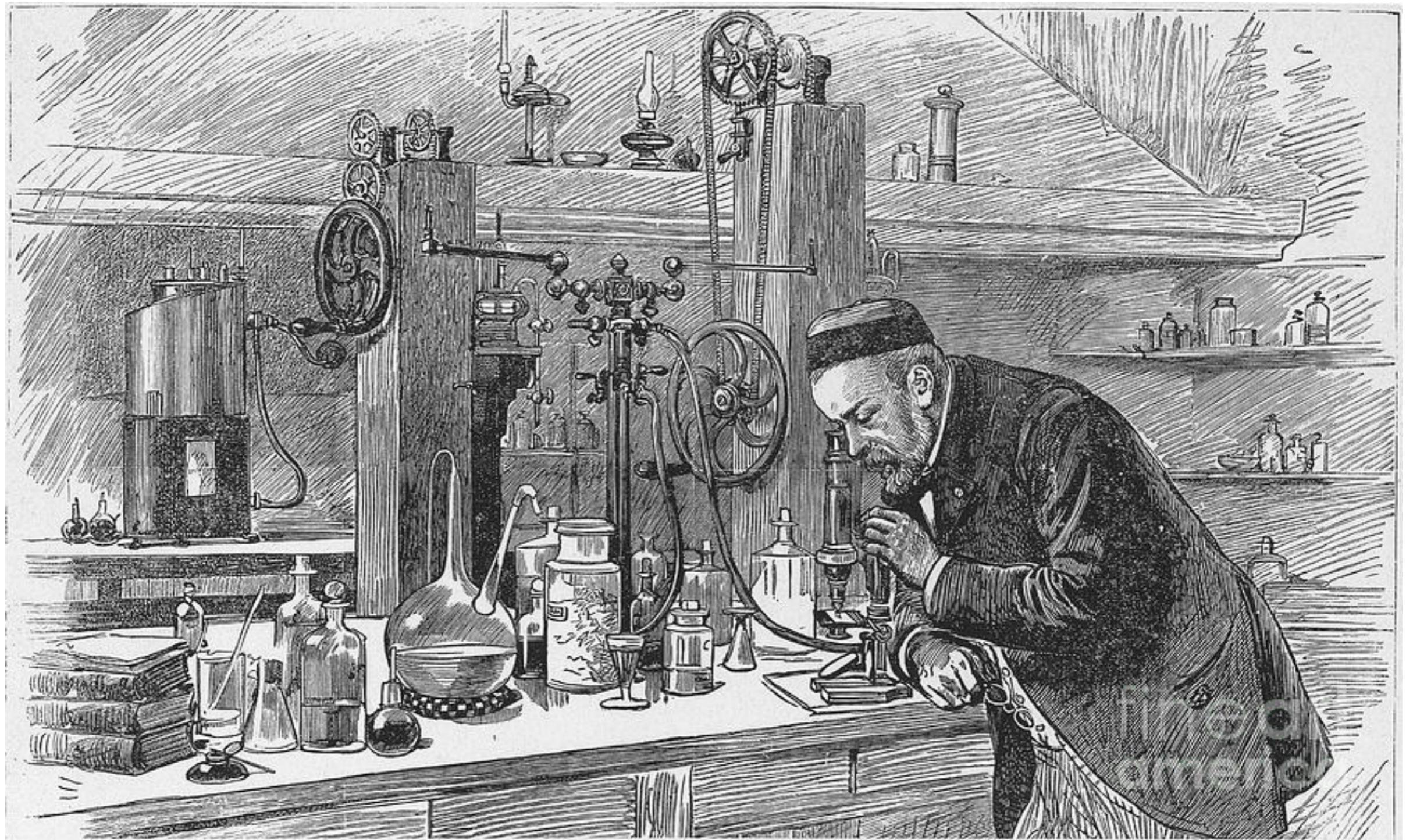
It’s a jungle in there: massive populations of microbes, immune cells, and cells of the gut tissue are interacting and exchanging countless chemical and physical signals. Disruptions to this complex ecosystem, often called the microbiome, have been linked not only to gastrointestinal problems but also to metabolic, immunological, and even neurological disorders.



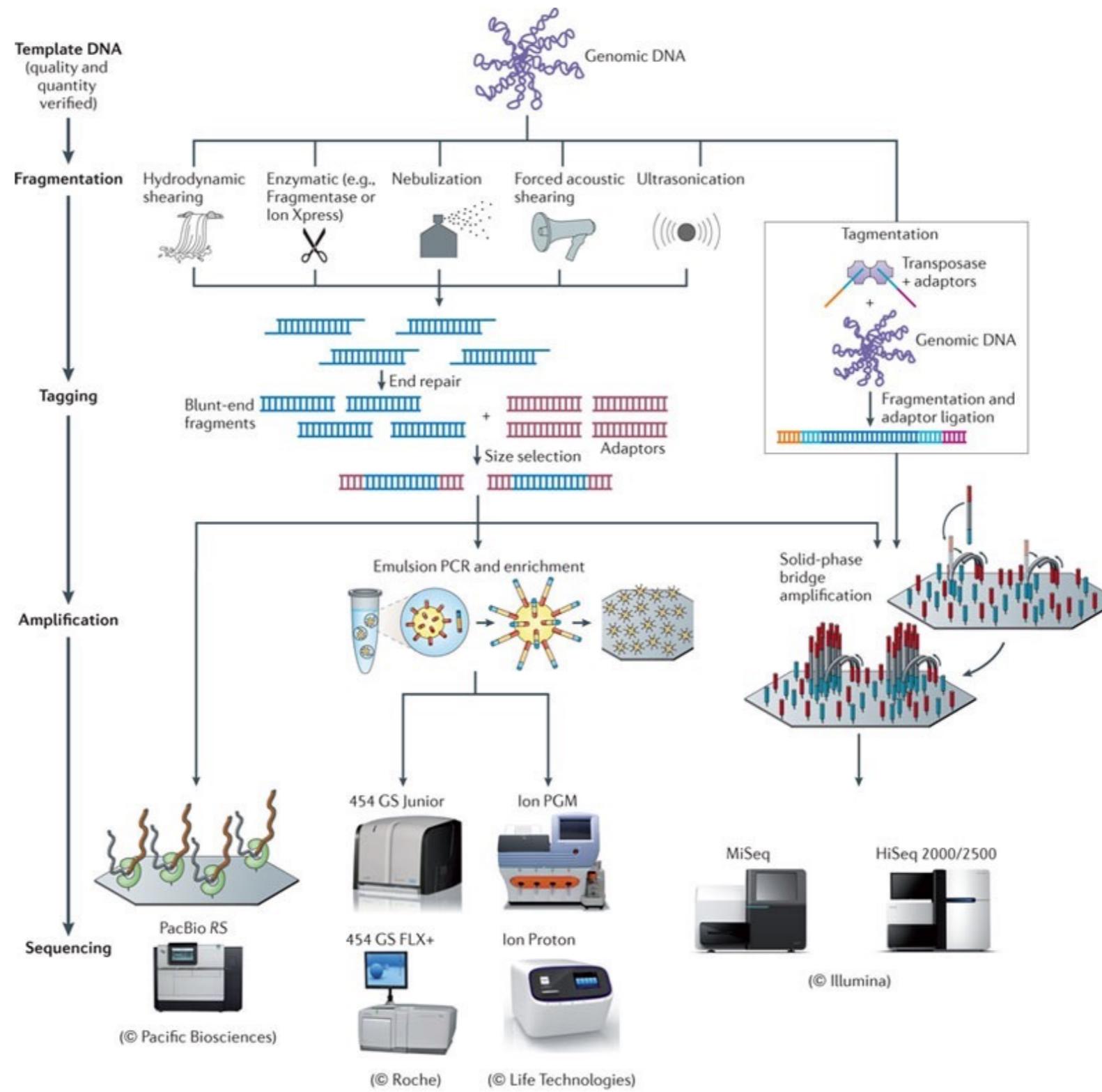
<https://www.technologyreview.com/s/545446/companies-aim-to-make-drugs-from-bacteria-that-live-in-the-gut/>

# Generación de metagenomas

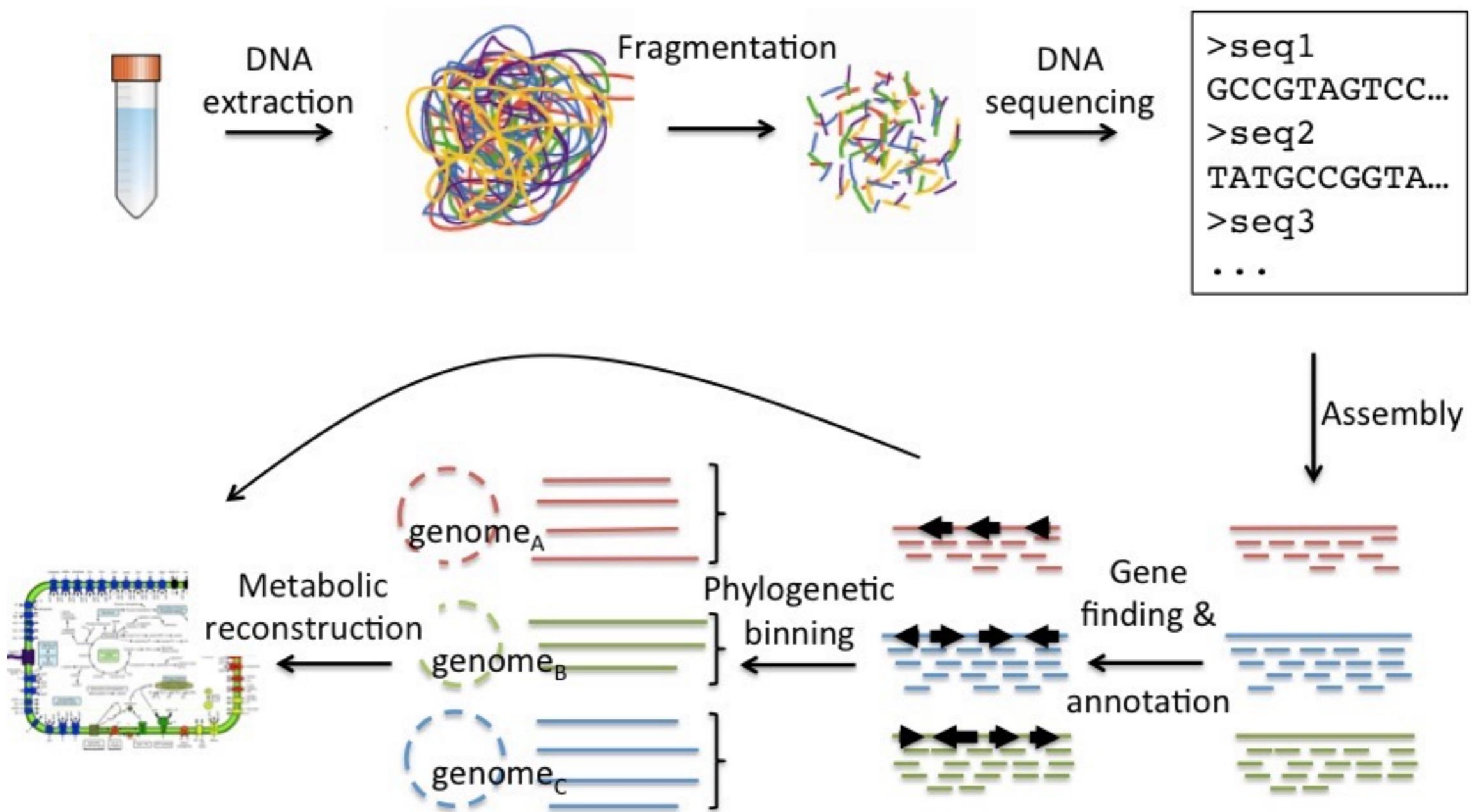
# Shotgun Sequencing: microscopio moderno



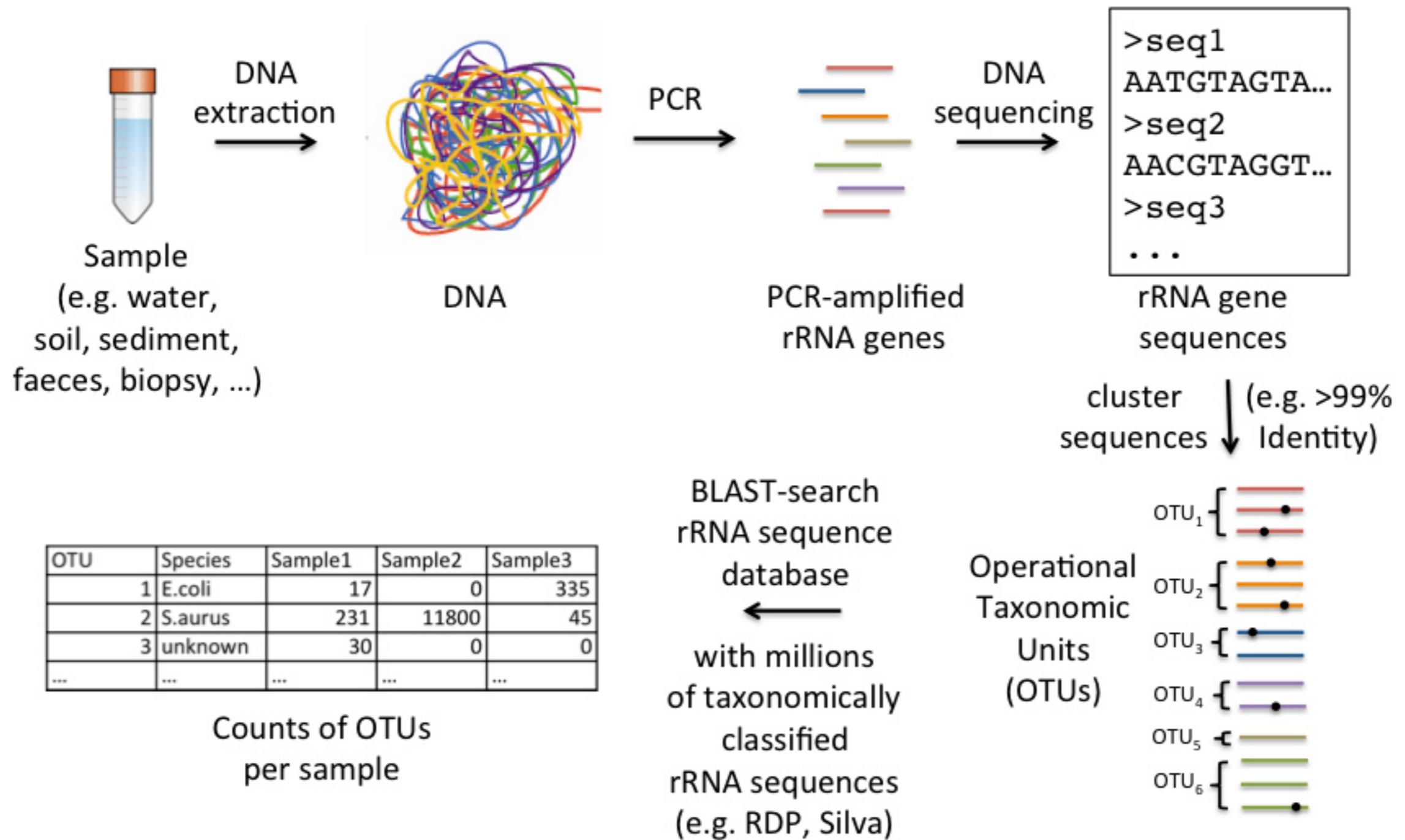
# Secuenciamiento masivo



# Shotgun Sequencing



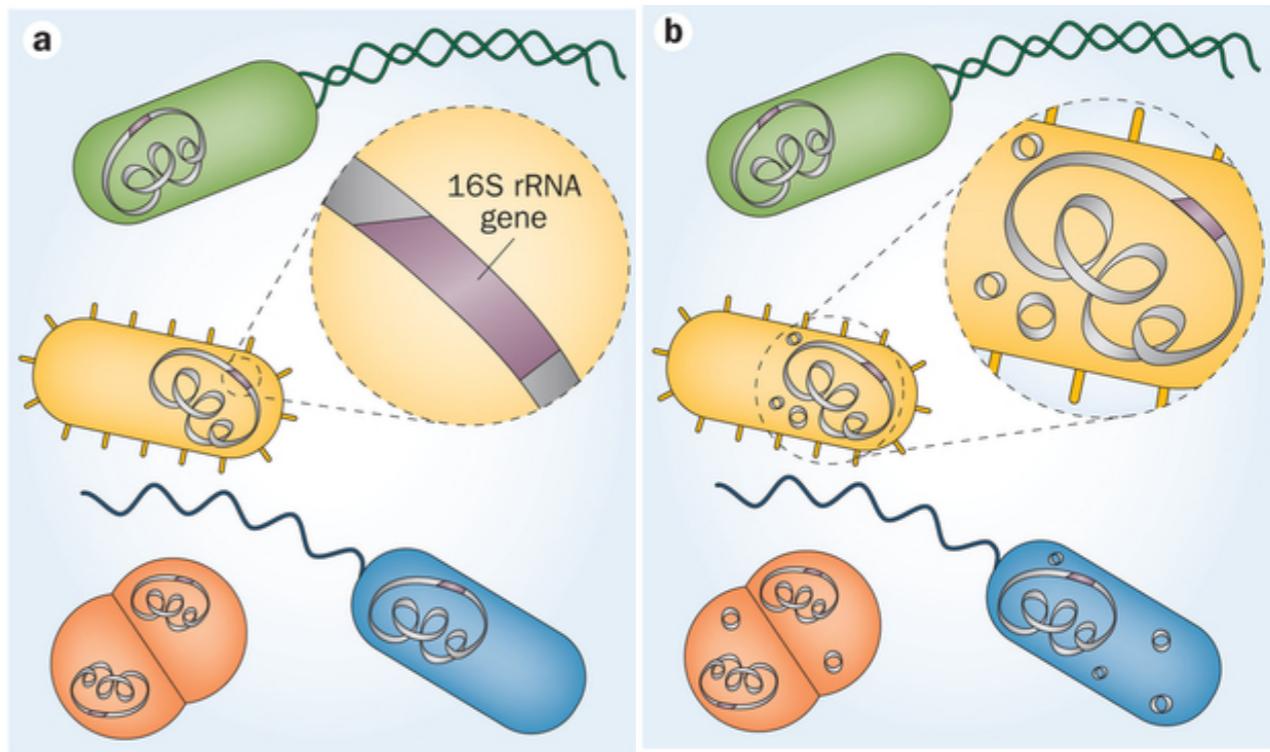
# Amplicon Sequencing



AMPLICON SEQUENCING OF RRNA GENES

# Definiciones

- Exploración por 16S rRNA o metataxonómica
- Metagenómica
- Microbioma



Marchesi and Ravel *Microbiome* (2015) 3:31  
DOI 10.1186/s40168-015-0094-5



EDITORIAL

Open Access

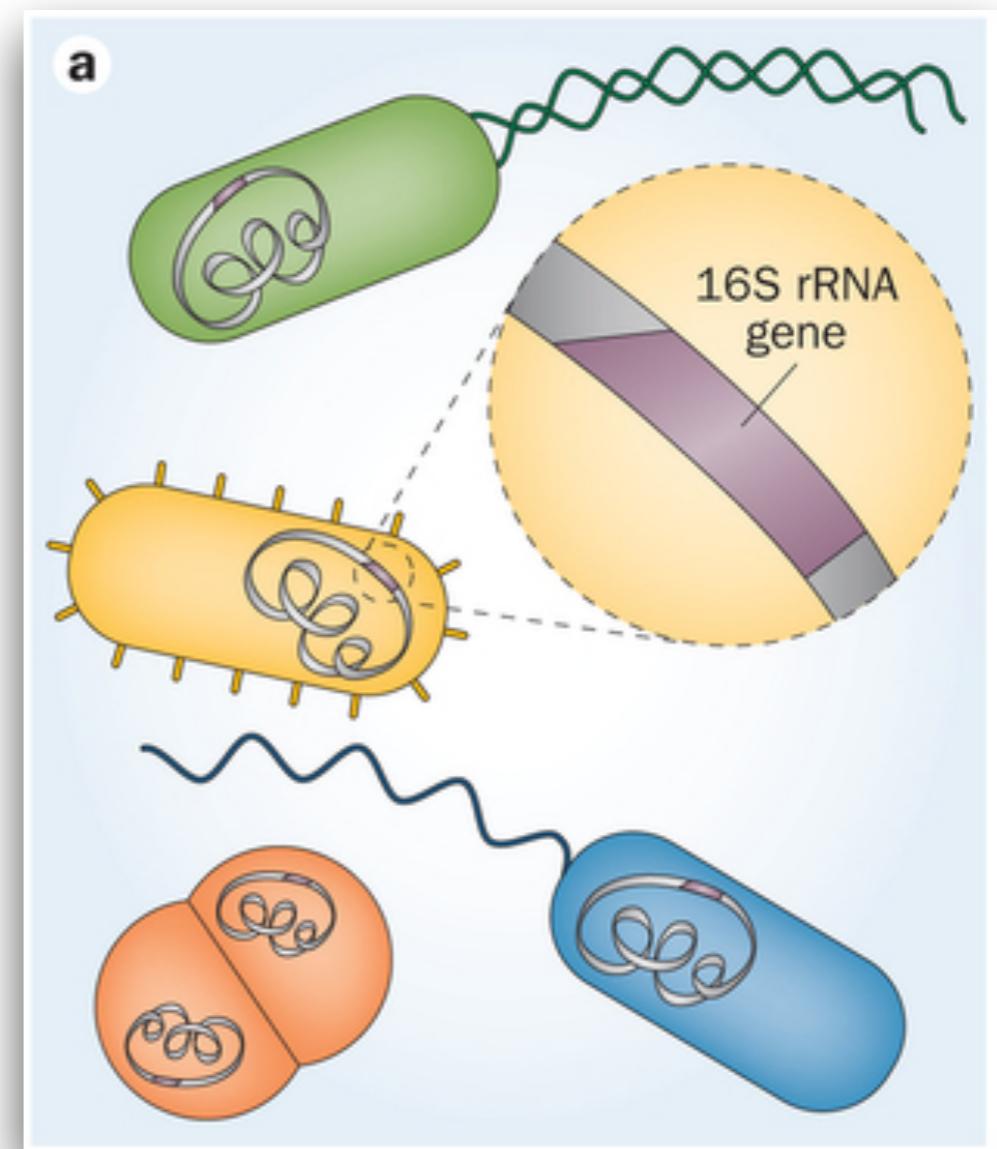
The vocabulary of microbiome research: a proposal



Julian R. Marchesi<sup>1,2</sup> and Jacques Ravel<sup>3,4\*</sup>

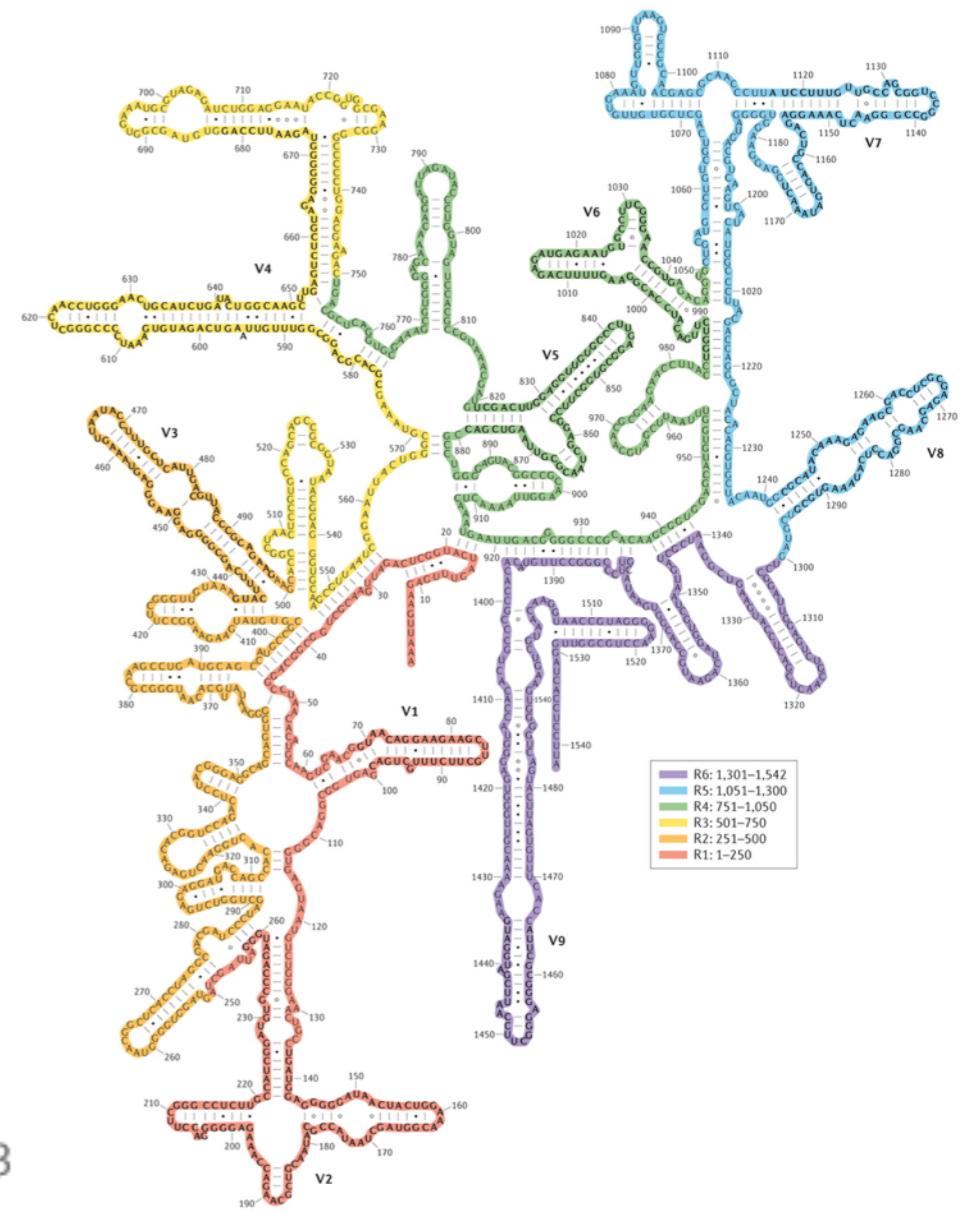
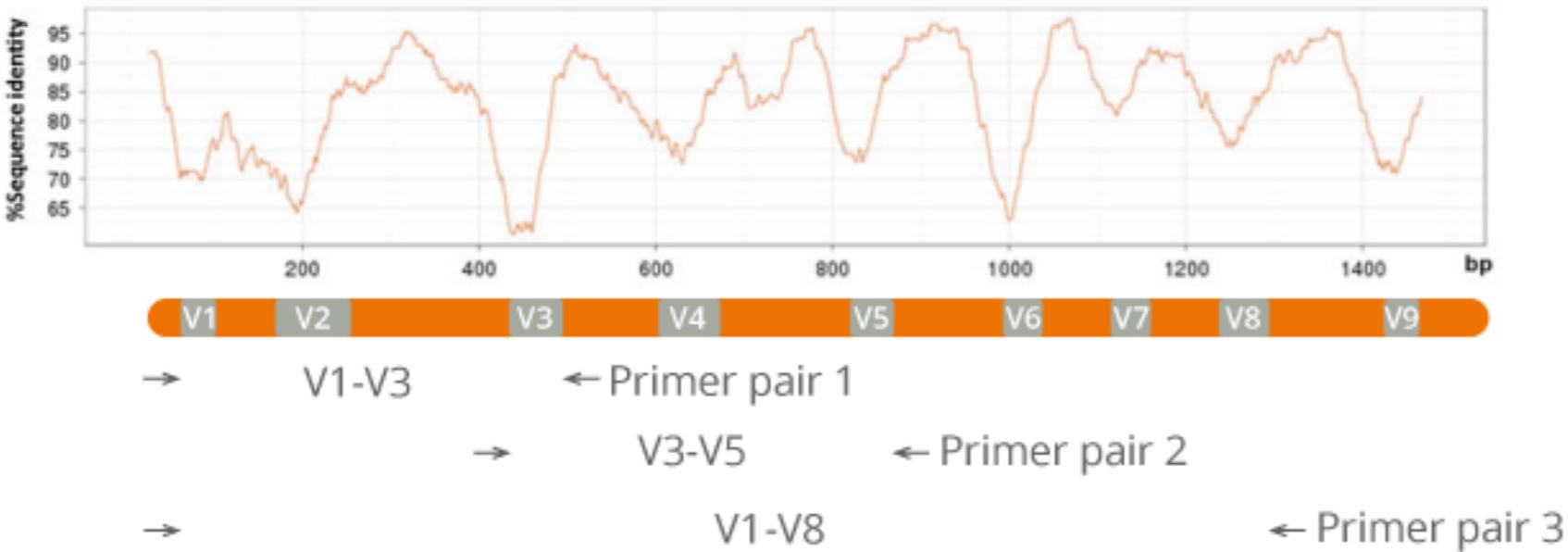
# Definiciones

- Metataxonómica - usar un gen marcador que se pueda asociar con taxonomía - te permite saber qué hay en una muestra - ITS / 16S /18S



# Metataxonómica

Regiones variables y constantes componen el gen 16S rRNA



# Metataxonómica

**silva**  
high quality ribosomal RNA databases

Home SILVAngs Browser Search Aligner Download Documentation Projects FISH & Probes Shop Jobs Contact

**Background**

The SILVA databases are developed and maintained by the [Microbial Genomics and Bioinformatics Research Group](#) in Bremen, Germany, in cooperation with the [Department of Microbiology](#) at the Technical University Munich and the company [Ribocon GmbH](#).

SILVA is an interdisciplinary project of biologists and computer scientists to provide:

- fully aligned and up to date small (16S/18S, SSU) and large (23S/28S, LSU) subunit ribosomal RNA "Parc" databases on the webpage as well as ARB files
- preconfigured subsets of only high quality, full-length sequences as ARB & FASTA files ([SSU/LSU Ref](#))
- extensive browse & search functionalities for sequence retrieval
- a clear rating system for all steps of data processing with emphasis on sequence and alignment quality
- full compatibility to the software package ARB & the latest official alignments released by the ARB project
- compatibility to many common programs like Phylip or Paup via direct Fasta export or the ARB program
- substantial support related to all aspects of data sets, sequence analysis and probe/primer design by our partner [Ribocon GmbH](#).

**Motivation**

Sequencing the ribosomal RNA (rRNA) genes is currently the method of choice for phylogenetic reconstruction and nucleic acid based detection and quantification of microbial diversity. The ARB software suite with its corresponding rRNA databases has been accepted by researchers worldwide as their standard tool for large scale ribosomal RNA analysis. More than 20 years of development have already been invested to extend and maintain the system. To provide high quality and comprehensive rRNA databases comprising *Bacteria*, *Archaea* and *Eukarya* the SILVA (from Latin *silva*, forest) system has been implemented in 2007. It is designed as an automatic software pipeline for sequence retrieval, quality assignment and the alignment of nucleic acid sequences based on the latest comprehensive ARB alignments.

**SINA: The new SILVA (Web)Aligner**

We developed a new aligner called SINA (SILVA INcremental Aligner) that is able to accurately align hundred thousands of sequences based on a curated SEED alignment. In a first step the aligner determines the next related sequences using an optimized Suffix Tree server. To find the optimal alignment for a new sequence up to 40 reference sequences are taken into account. While running, the system simulates the manual refinement process to optimize the result.

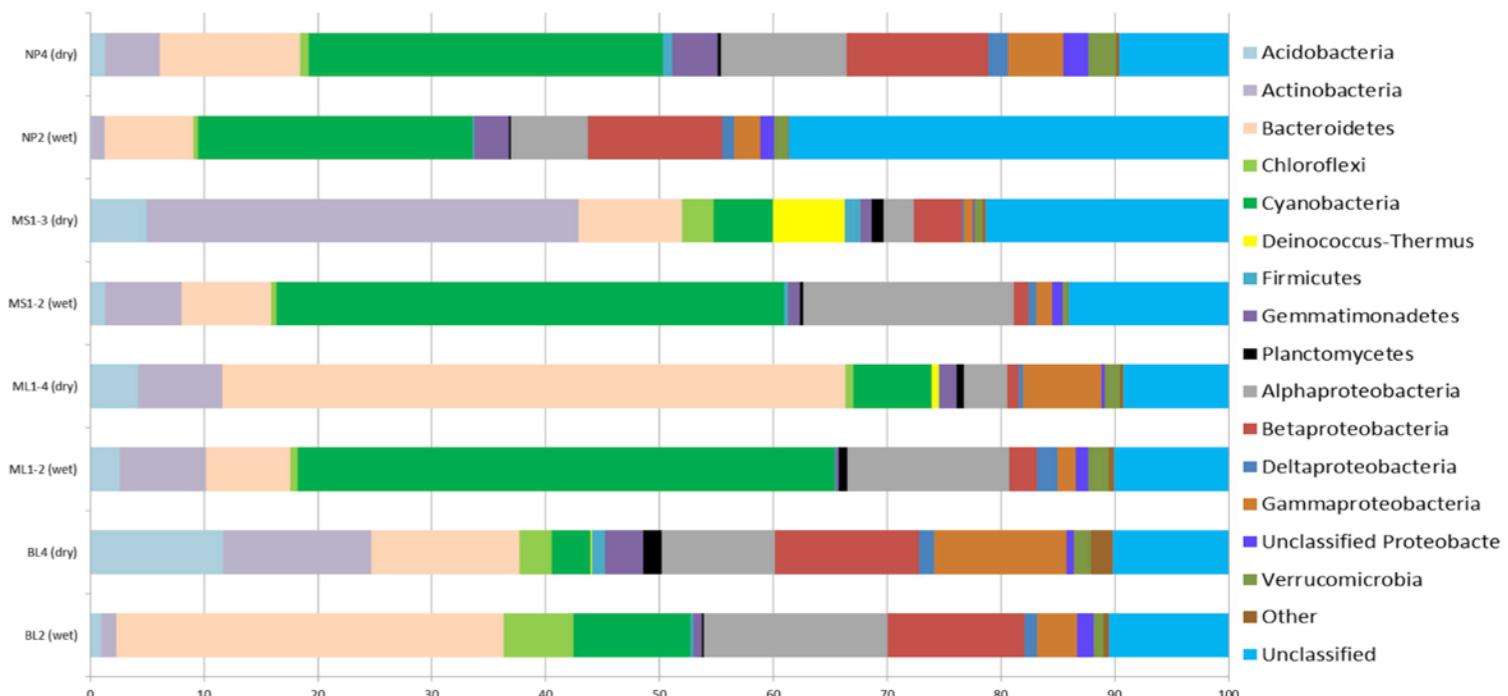
**Features of SINA:**

- Process and quality values are added to each sequence indicating e.g. the alignment quality
- Only minimum manual revision of the output alignment is required (e.g. no base-spreading at the ends)
- Improved alignment quality due to advanced alignment technology compared to e.g. the ARB Aligner ("Fastaligner")

SINA is also available [online](#) for small scale projects. More information about SINA as well as the corresponding publication can be found [here](#).

**Release information & Database history**

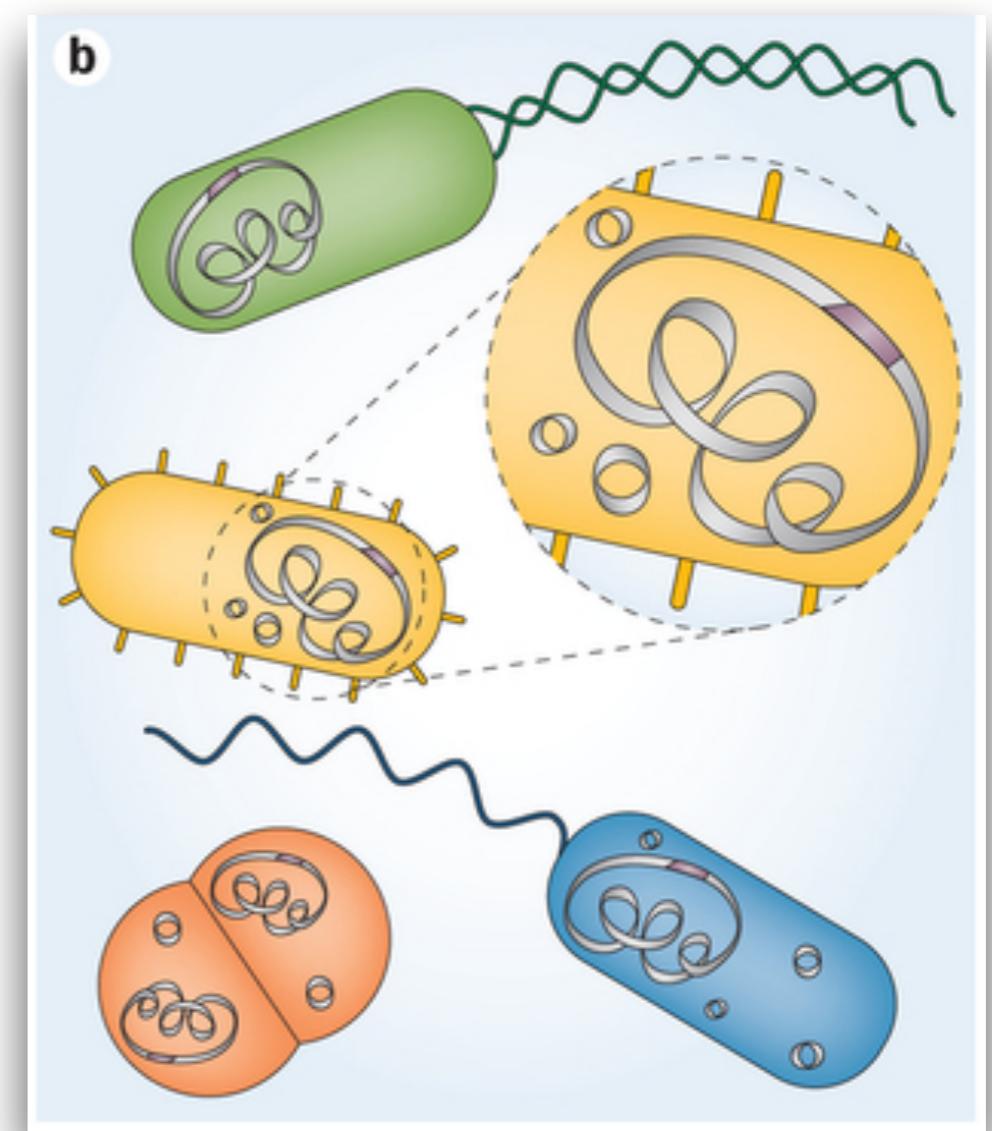
Version 89 of our datasets have been made available in February 2007. Version 123 was released in July 2015 and increased the number of available SSU/LSU sequences to over 5,300,000. Detailed information about the content of the databases and statistics can be found [here...](#)



Bases de datos → Perfil taxonómico

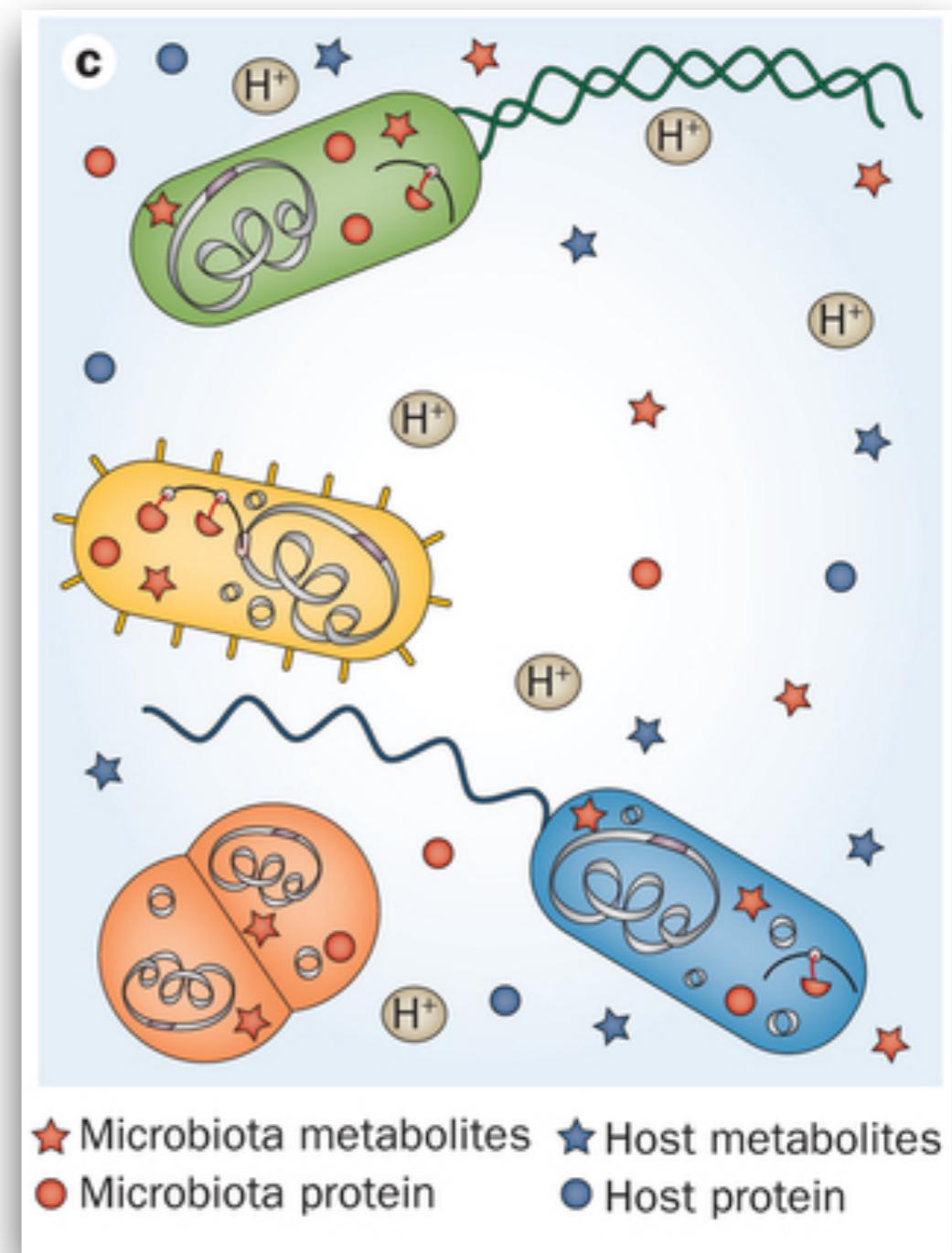
# Definiciones

- Metagenoma - colección de genes y genomas de un ecosistema - más que la taxonomía, también la función - qué hay y qué (potencialmente) hacen



# Definiciones

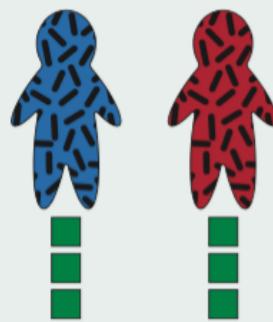
- Microbioma - similar a metagenómica - aplicado casi exclusivamente a estudios en humanos.
- También como extensión de bioma (factores bióticos y abióticos)



# Diseño Experimental

## a Fixed sequencing budget

High sequencing depth reveals rare features within each sample



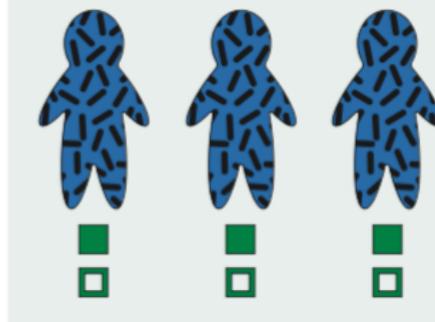
Reduced depth enables larger sample sizes (greater statistical power)



Time courses within communities reveal changes in response to stimuli and other dynamical properties



Combined DNA and RNA sequencing reveals differences between community functional potential and functional activity

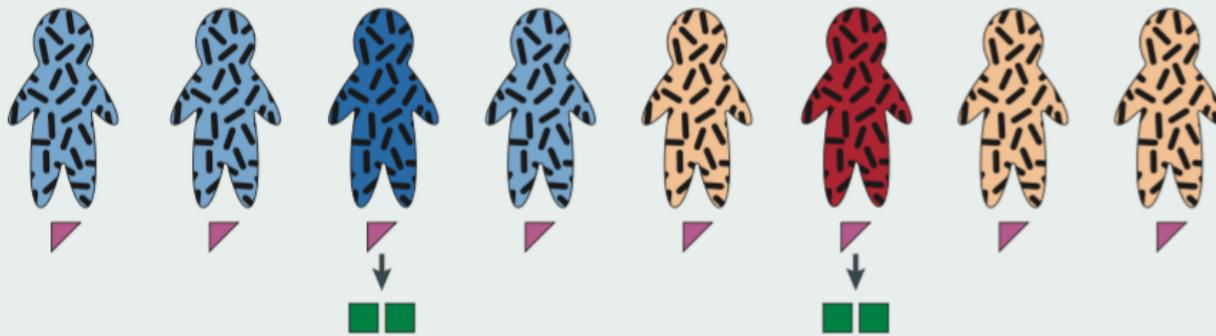


■ One 'unit' of WMS sequencing

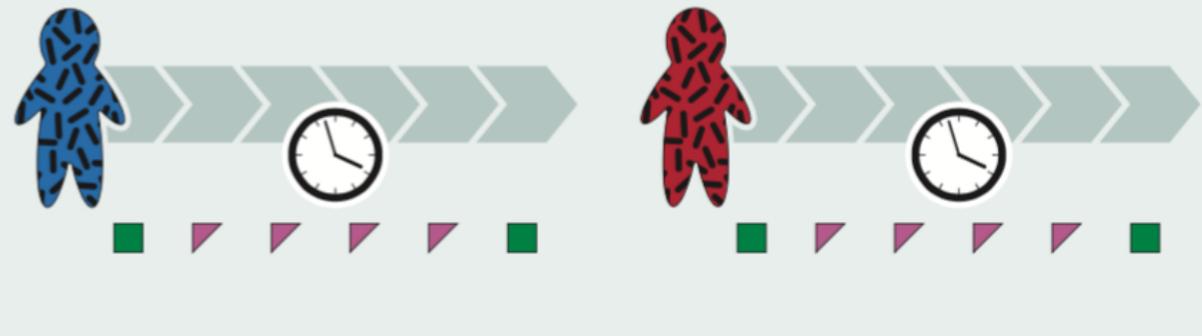
■ DNA ■ RNA

## b Combining WMS and amplicon sequencing

In a tiered study, many samples are initially surveyed by amplicon sequencing; later, a subset of representative or extreme samples are explored in greater detail by WMS sequencing



In time-course studies, amplicon sequencing can be applied to survey a large number of internal time points, while WMS sequencing can be used to dissect a subset of time points (e.g. the first and last) in greater detail



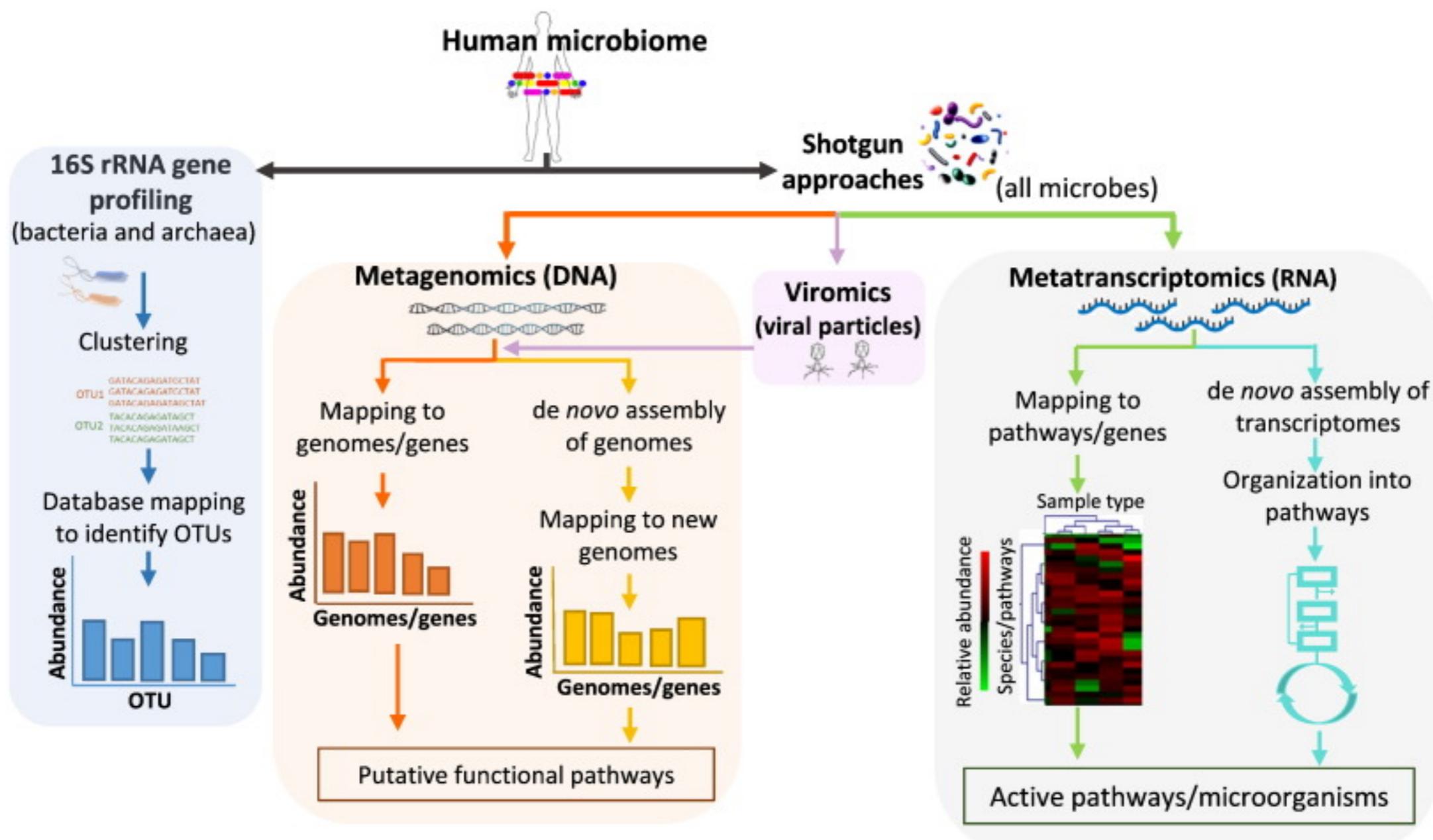
■ One 'unit' of amplicon sequencing

■ = ■ ■ ■ ■ ■

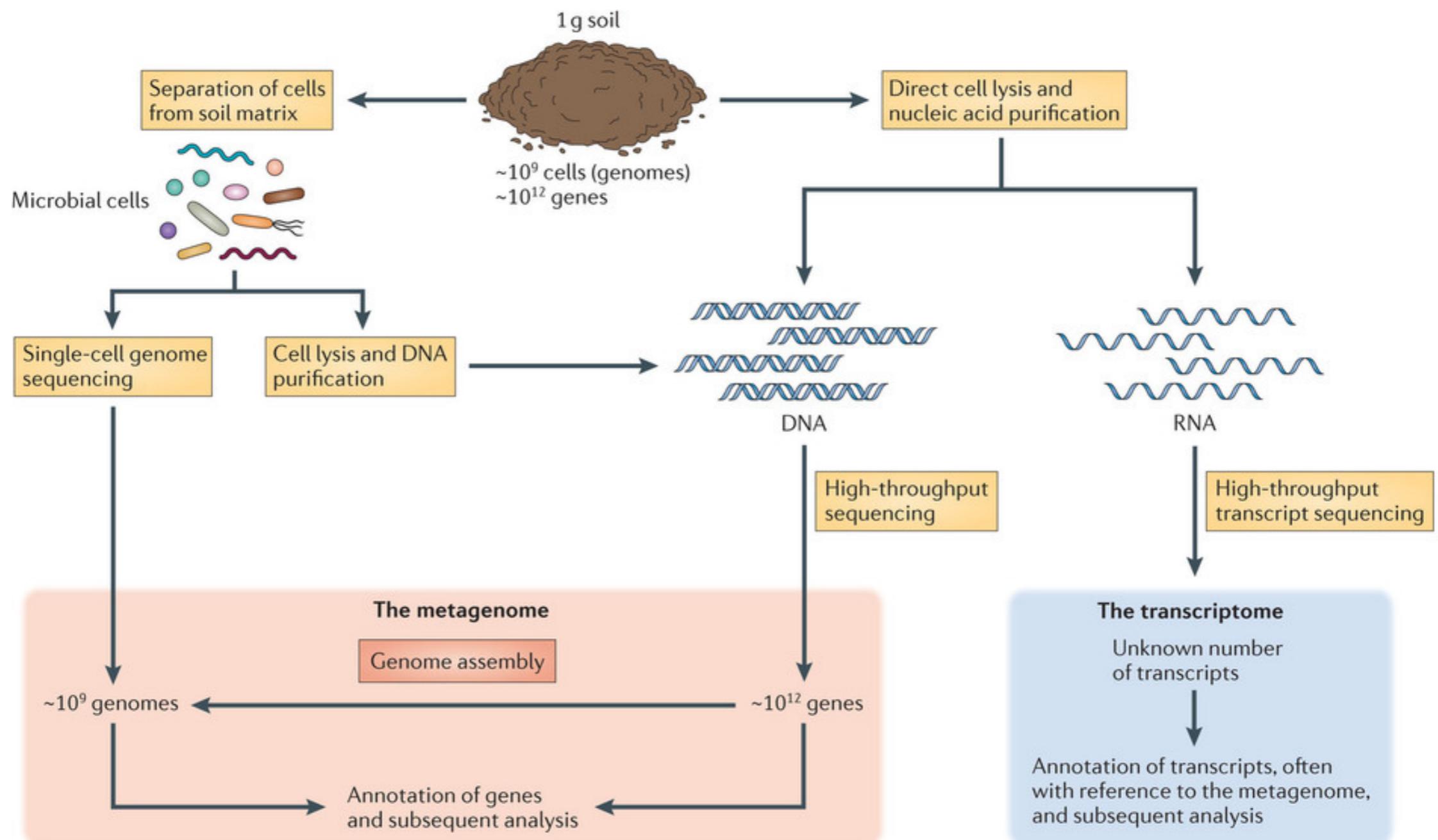
# Taxonomic and Functional Profiling of microbiota

- Taxonomic Profiling - qué hay
- Functional Profiling - qué hacen o potencialmente hacen

# Taxonomic and Functional Profiling of microbiota



# Taxonomic and Functional Profiling of microbiota



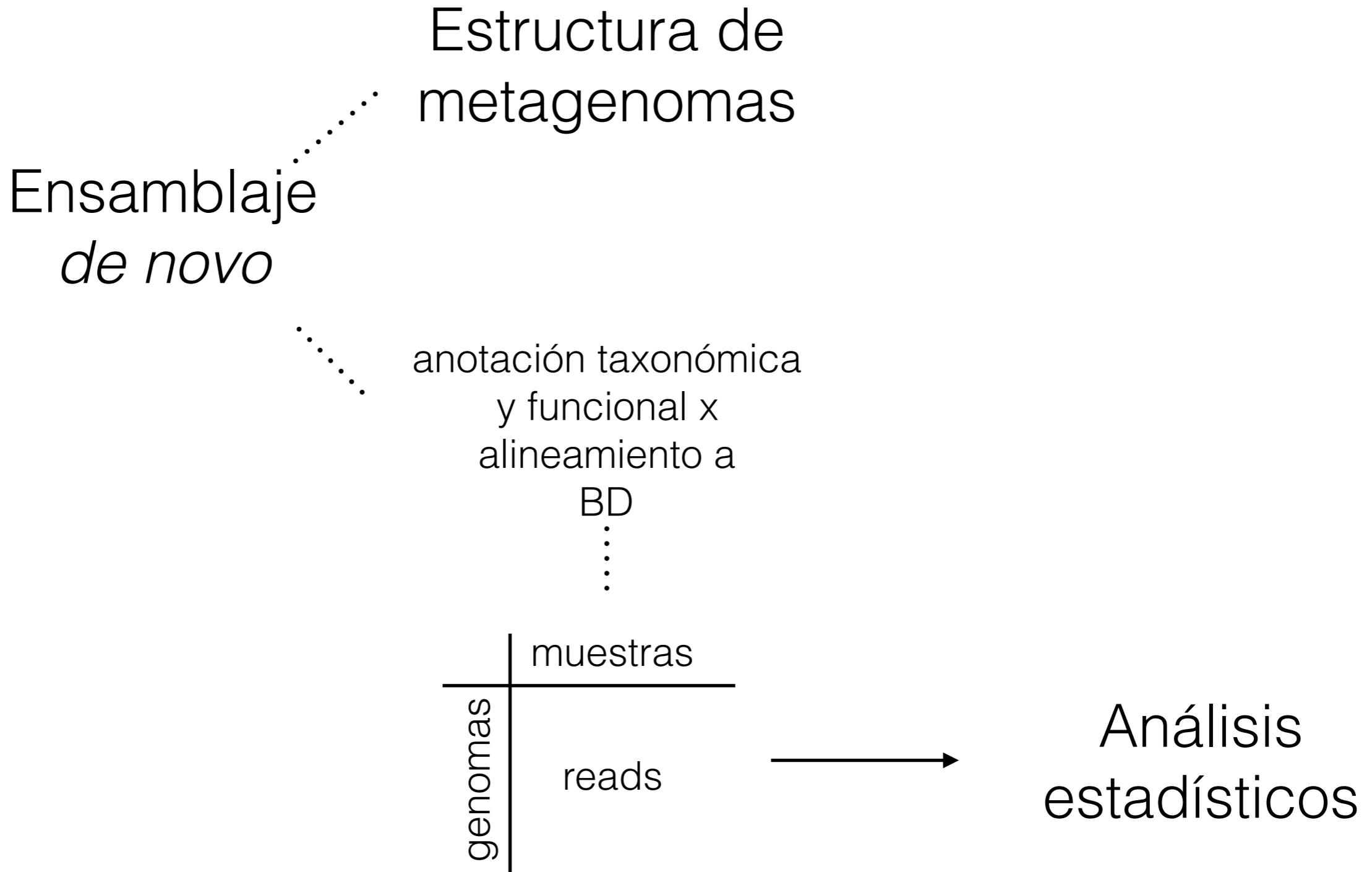
# Análisis de metagenomas

# Dos estrategias de análisis

- Basado en ensamblaje *de novo*
- Basado en mapeo en contra de referencias

# Estrategias analíticas para datos de metagenomas

# Estrategias analíticas para datos de metagenomas

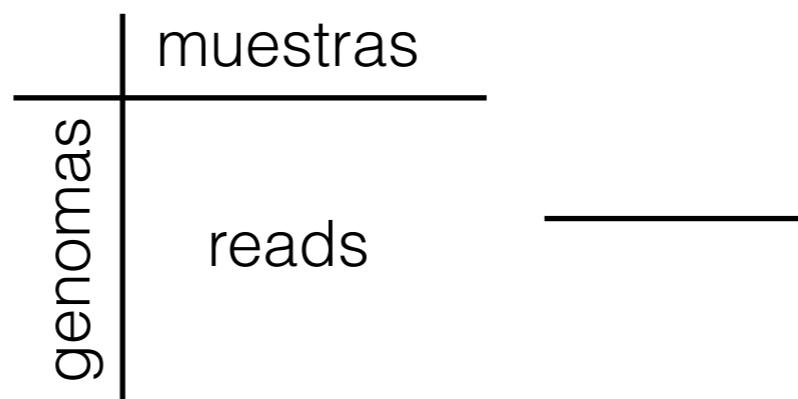


# Estrategias analíticas para datos de metagenomas

Estructura de metagenomas

anotación taxonómica  
y funcional x  
alineamiento a  
BD

⋮



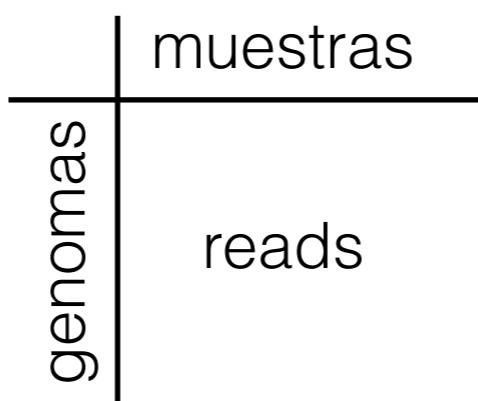
Análisis  
estadísticos

# Estrategias analíticas para datos de metagenomas

Estructura de metagenomas

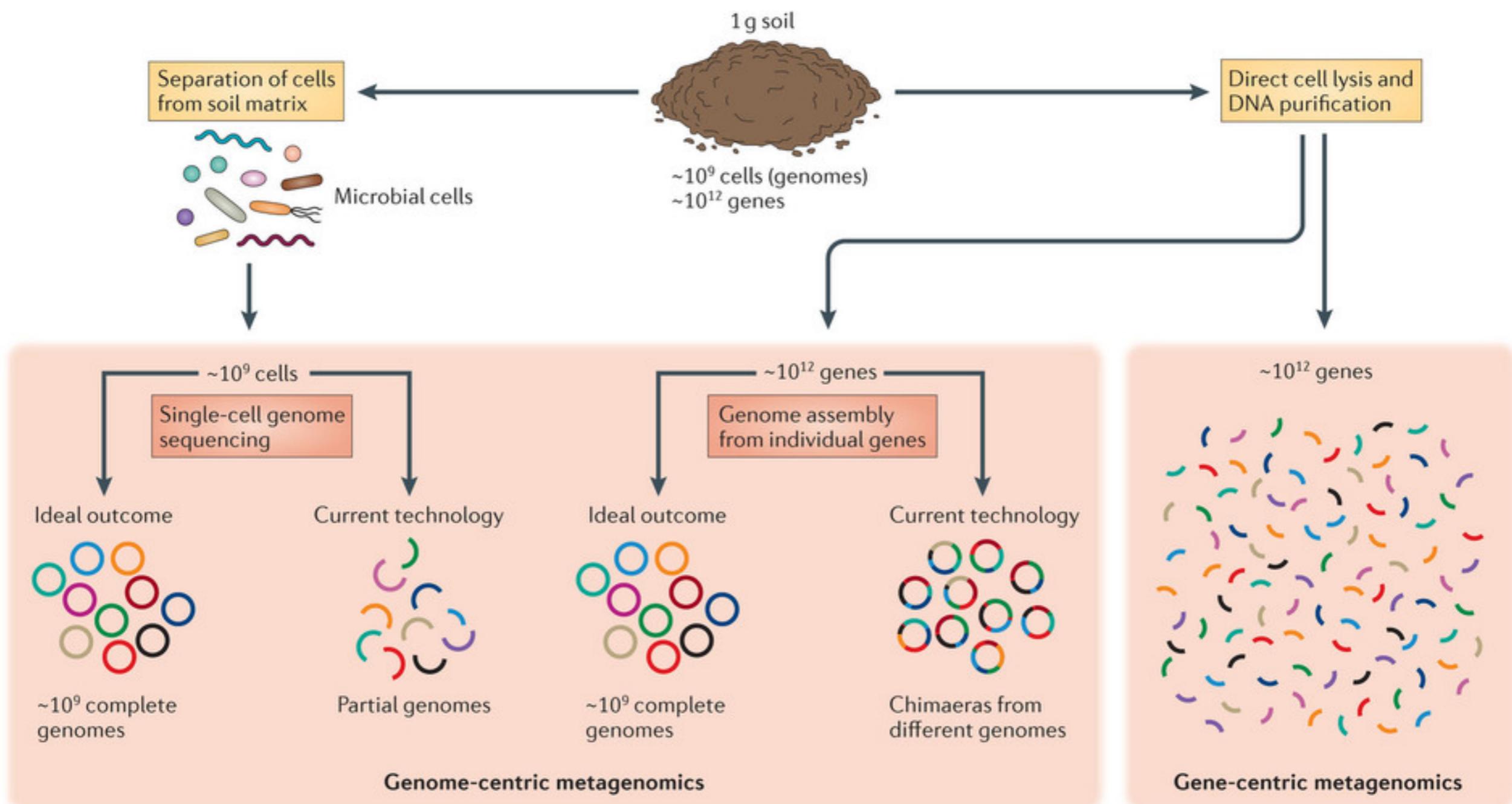
datos sin ensamblar

anotación taxonómica  
y funcional x  
alineamiento a  
BD



Análisis estadísticos

# Análisis basado en genes y basado en genomas



# Metagenómica *de novo*

- Ensamblar reads de manera similar a genomas individuales
- Problema - generación de contigs químéricos

**Chimeric contigs** are contigs that combine sequences from more than one genome. Short sequence reads from two different genomes can be incorrectly assembled into one contig due to a short region of similar sequence.

- ¿Qué hace diferentes a los genes de distintos microorganismos?

# Binning metagenómico

- Genomas de distintos microorganismos tienen:
  - ★ Distinto uso de codones - código genético
  - ★ Distinta composición nucleotídica - distribución de K-mers; frecuencia de tetranucleótidos, etc.
  - ★ Distinto *coverage*

# Binning metagenómico

PhD student

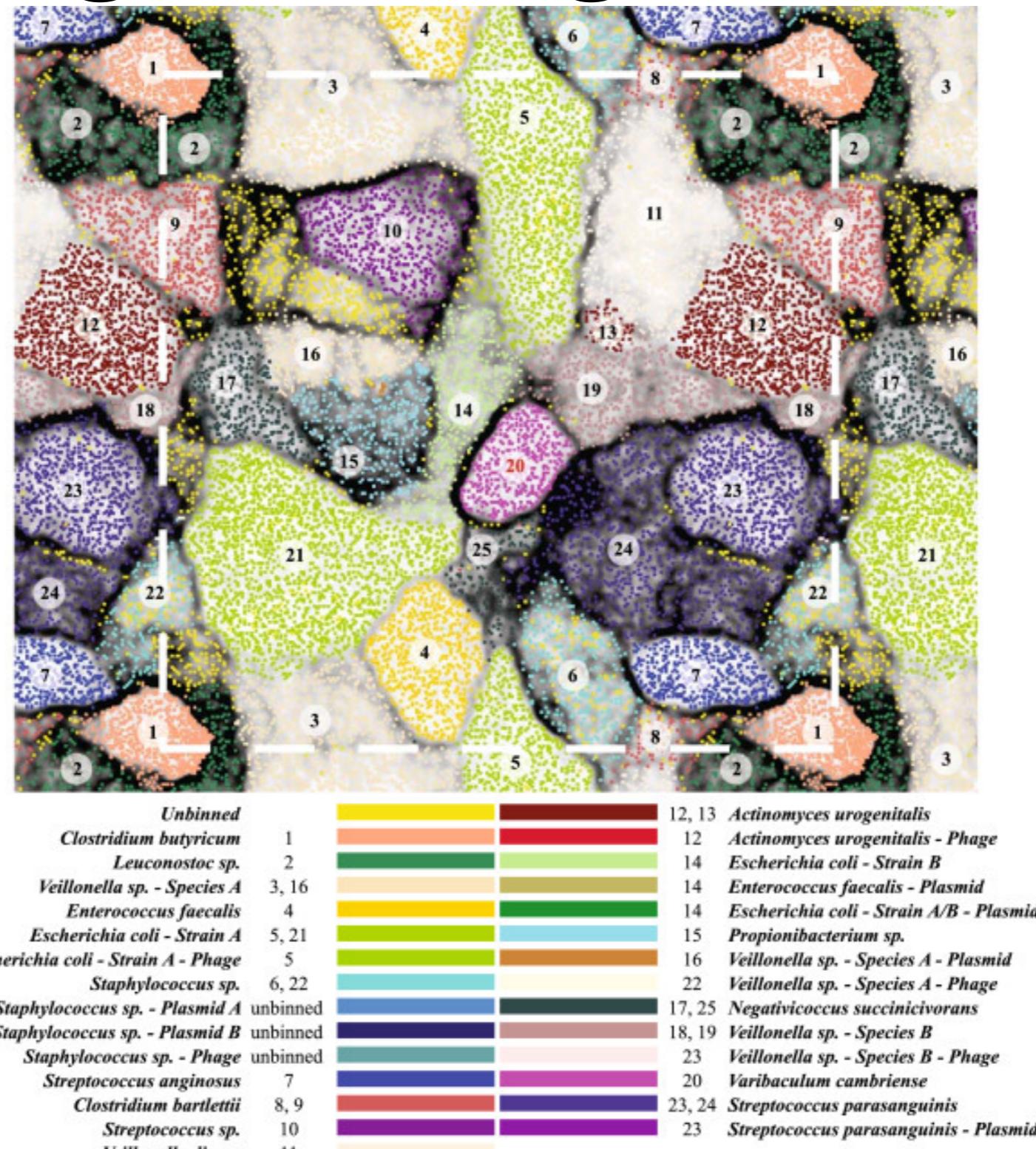


Complex sample

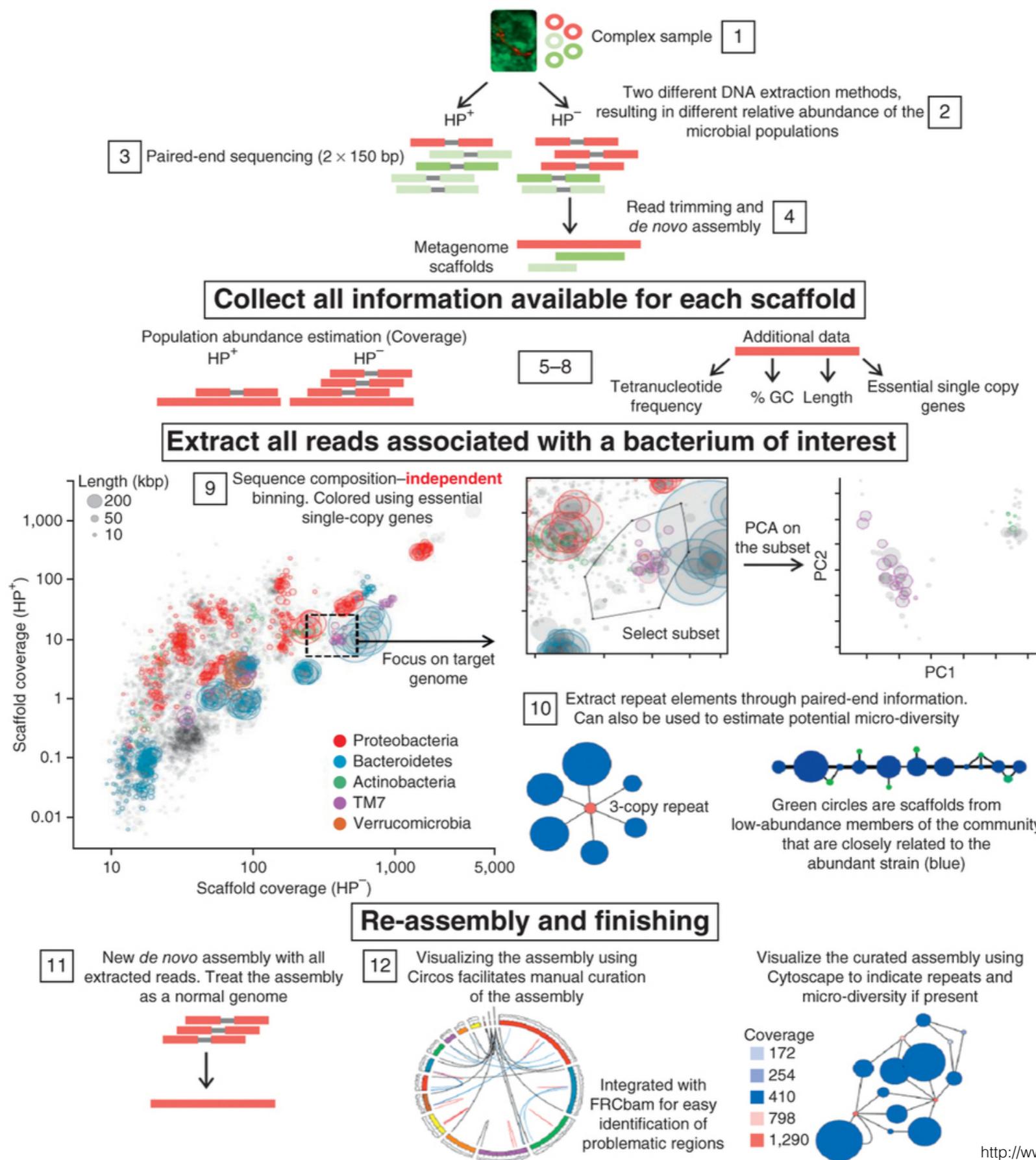
"Binning"



# Binning metagenómico



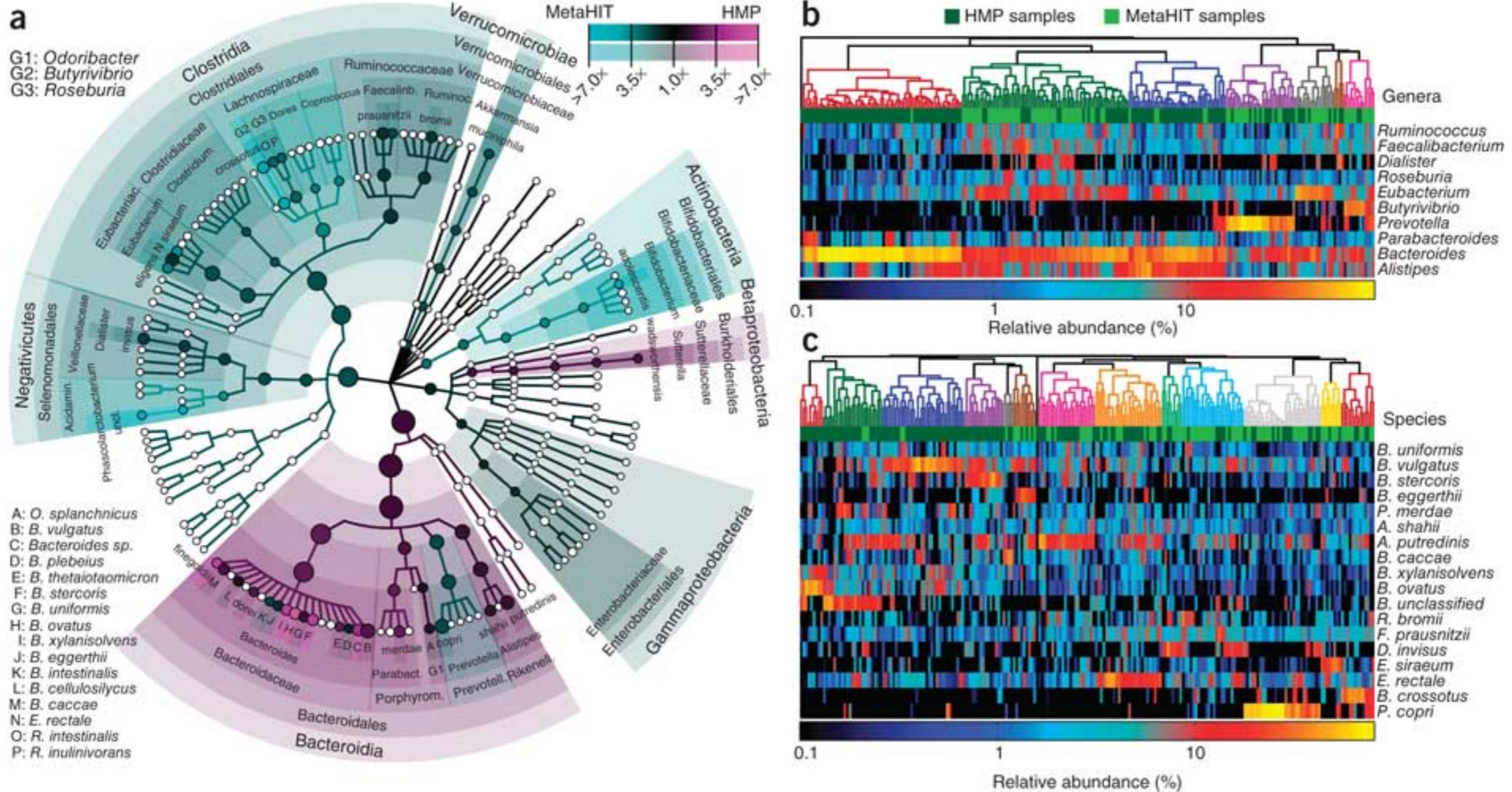
# Extracción de genomas desde metagenomas



# Metagenómica por referencias

- Usar RefSeq u otra base de datos de genomas
- Muy útiles para muestras de humanos, + referencias!
- No muy útil para muestras ambientales, por qué?
- Programas clásicos: PathoScope, MetaPhLan

# Metagenómica por referencias - MetaPhlan

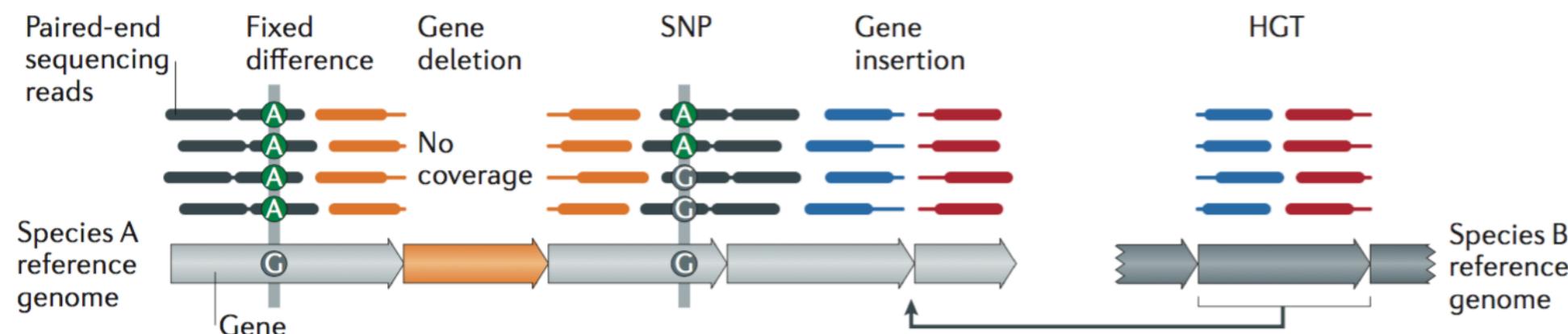


# ¿Cuál es el producto de un análisis metagenómico?

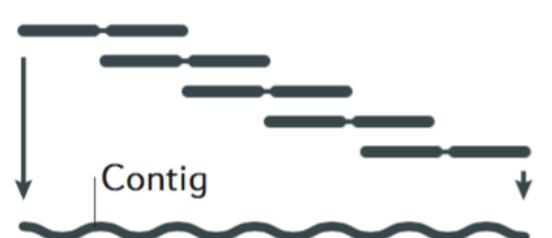
- Una tabla de abundancia de cada especie en la muestra - qué hay y en qué frecuencia
- ¿Qué podemos hacer con eso? —> usando modelos estadísticos se puede averiguar qué microorganismos es importante, cómo están relacionados entre sí, cómo y por qué varian en el tiempo, etc.
- Estadística en metagenómica —> otro curso por sí solo

# Detección de variabilidad genética

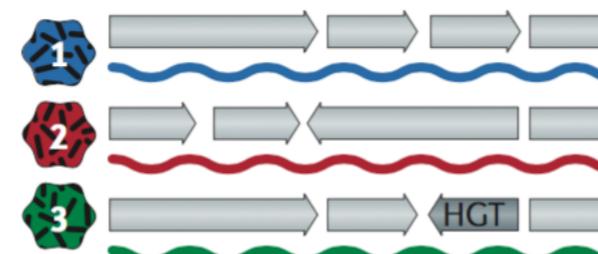
## a Detecting strain variation



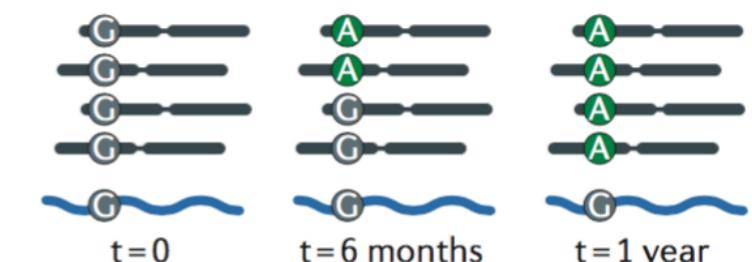
## b Using metagenomic assembly



Assembly merges overlapping reads into larger genomic fragments called contigs

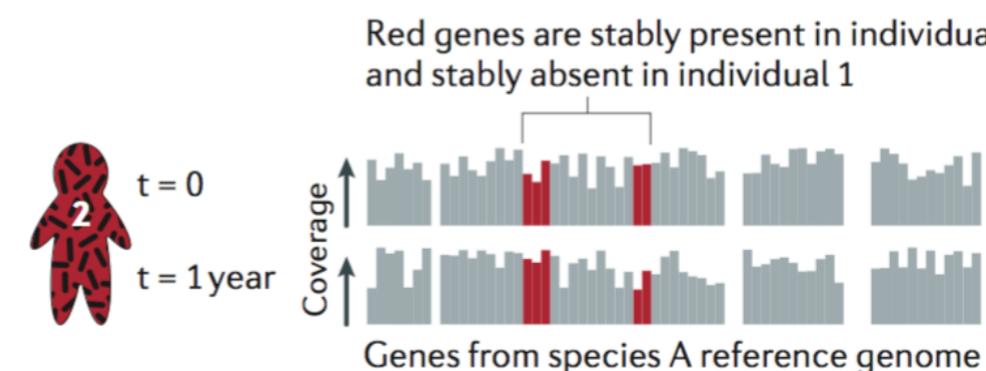
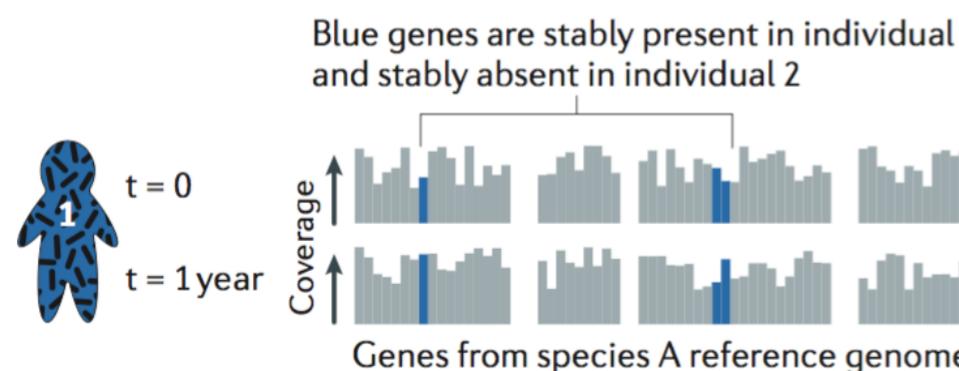


Contigs from three environmental isolates of species A reveal novel genomic arrangements and HGT events



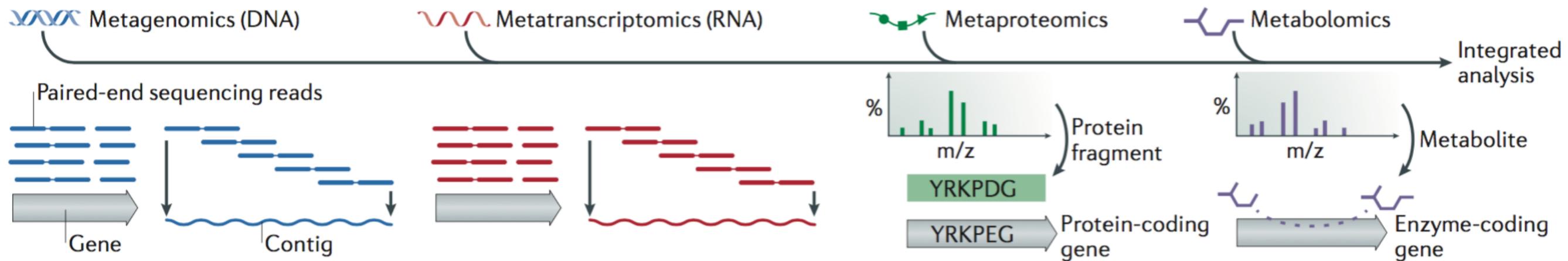
Assembly enables strain-level analysis in the absence of a reference genome

## c Longitudinal analysis

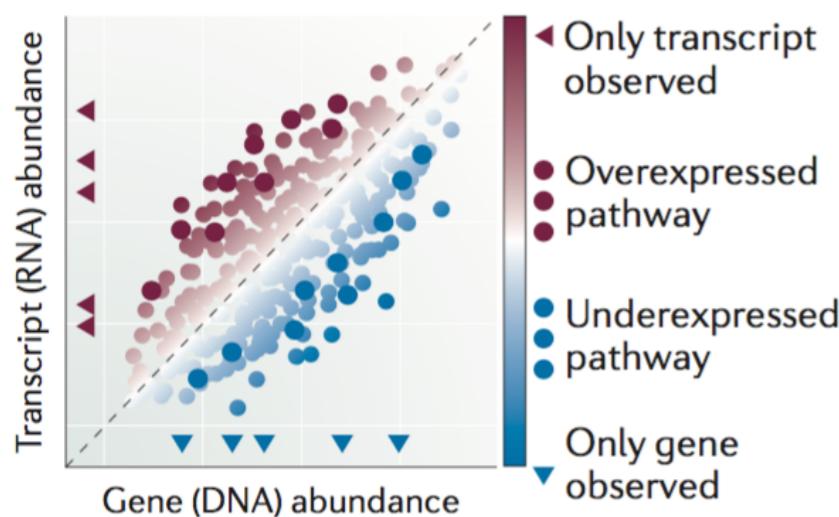


# Integración de datos multiómicos

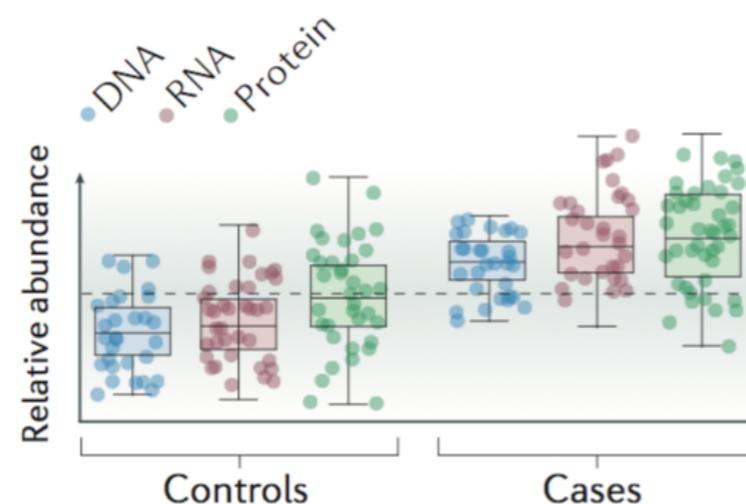
## a Multi-omics data types



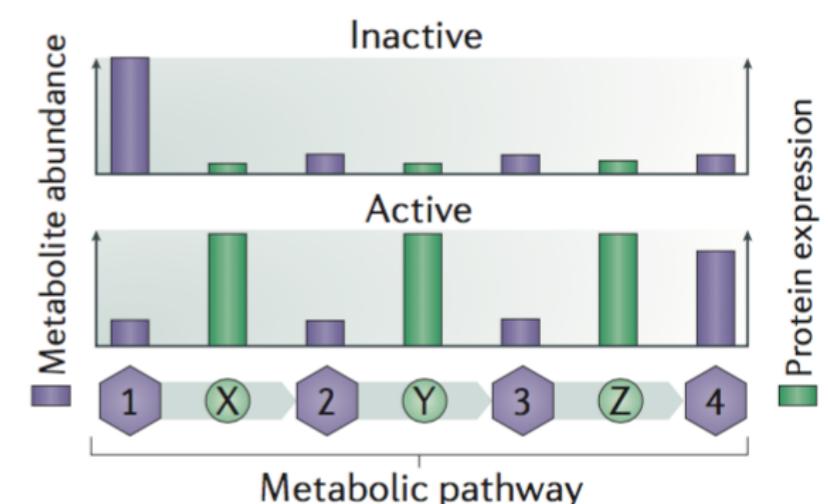
## b Normalization



## c Strengthening hypotheses



## d Descriptive modelling



Occblj

Meerwunder und selzame Thier/ wie sie in den Ostanächtigen Ländern im Meer  
und auf dem Lande gefundas werden.

Occbr

