



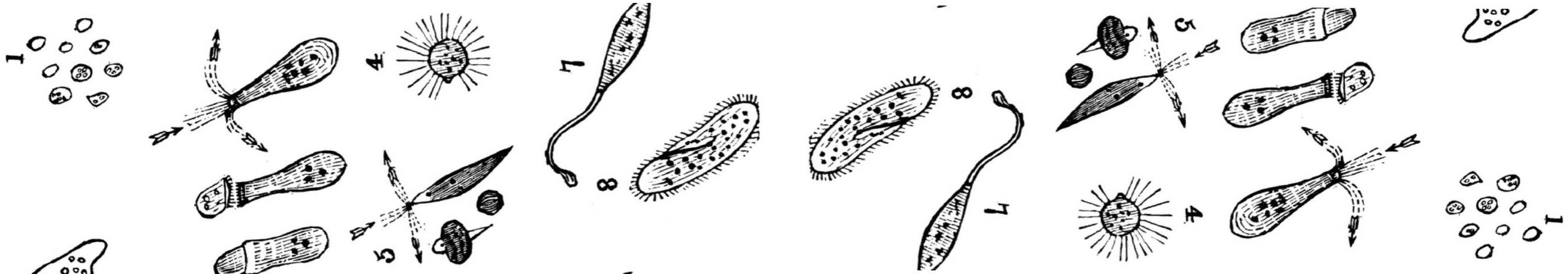
CBIB

CENTER FOR BIOINFORMATICS
& INTEGRATIVE BIOLOGY

Metagenómica

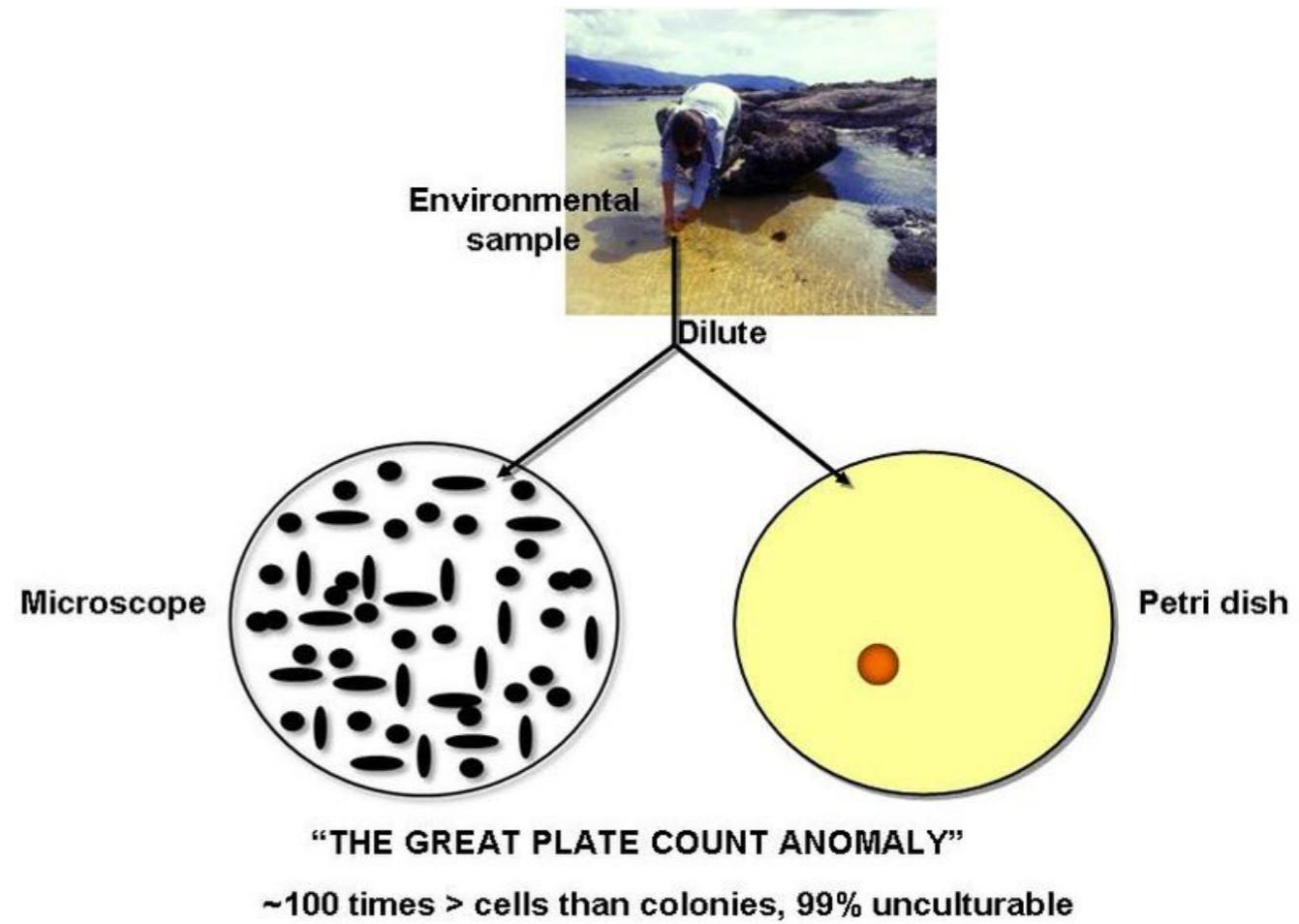
www.castrolab.org
www.cbib.cl

Katterinne N. Mendez, M.Sc.
Eduardo Castro, PhD
Universidad Andrés Bello
30 de octubre de 2018



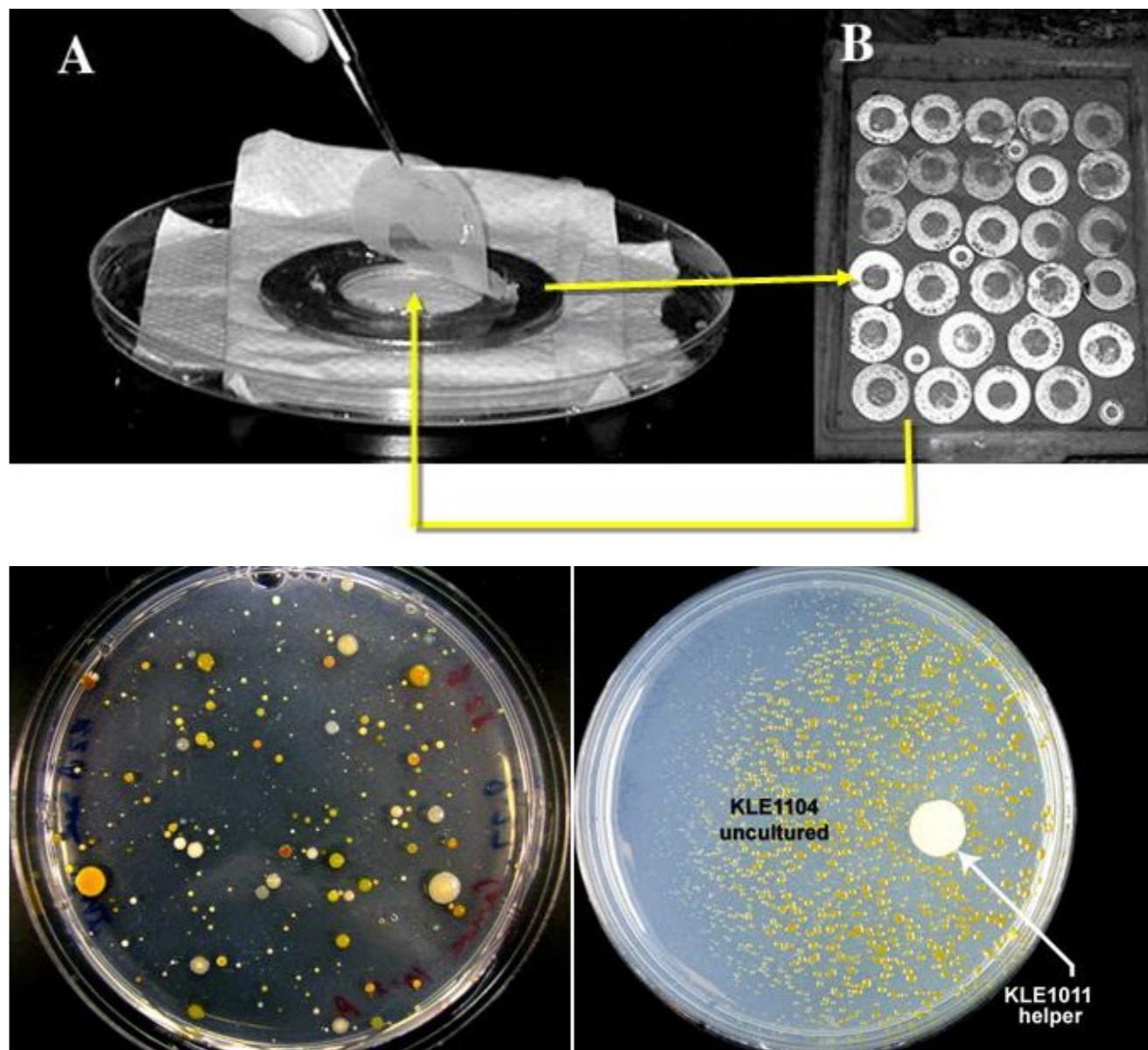
La mayoría de los microorganismos son incultivables

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



La mayoría de los microorganismos son incultivables

- Paradoja: genomas similares, uno crece y otro no. Deberían crecer en el mismo medio. *Bacillus*; *Pseudomonas*.
- Crecer microorganismos en su ambiente natural.
- No solo se trata de requerimientos nutricionales. Hay organismos que crecen solamente en presencia de otra especie proveniente del mismo ambiente.
- “Factores de crecimiento”



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

M. luteus sp.

KLE1104 (close relative of cultivable *M. polysiphoniae*)
KLE1011 (related to *M. luteus*)

Diversidad microbiana

Es posible encontrar microorganismos en la mayoría de los ambientes en la Tierra.



¿Cuántos microorganismos hay?

Table 1 Estimating the magnitude of microbial diversity

Number of bacteriophages on Earth	10^{31}
Number of microbes on Earth	5×10^{30}
Number of stars in the universe	7×10^{21}
Number of microbes in all humans	6×10^{23}
Number of humans	6×10^9
Number of microbial cells in one human gut	10^{14}
Number of human cells in one human	10^{13}
Number of microbial genes in one human gut	3×10^6
Number of genes in the human genome	2.5×10^4
Combined length of all bacteriophages on Earth	10^8 Ly
Diameter of the Milky Way	10^5 Ly

Importancia

- Microorganismos manejan y sostienen toda la vida en el planeta.
- Ciclo del nitrógeno, carbono, oxígeno, 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.



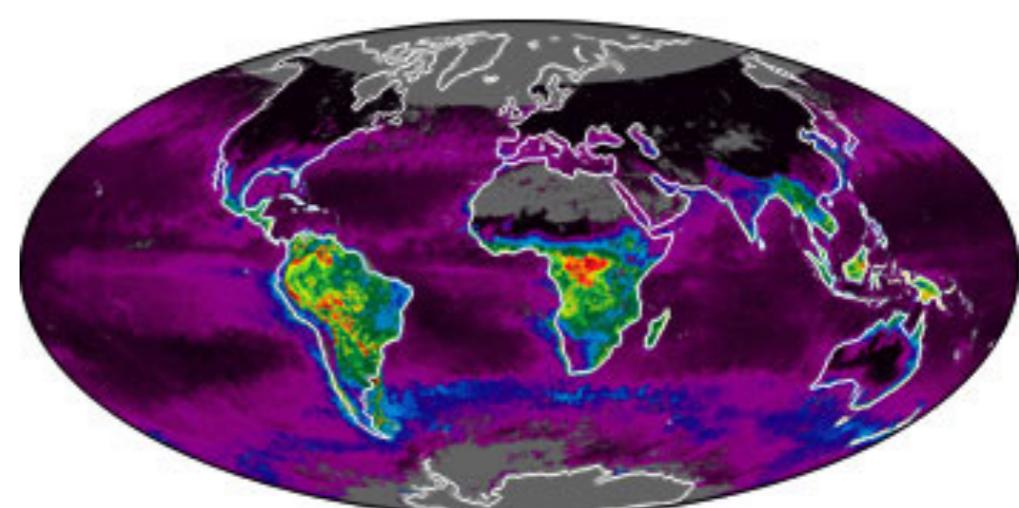
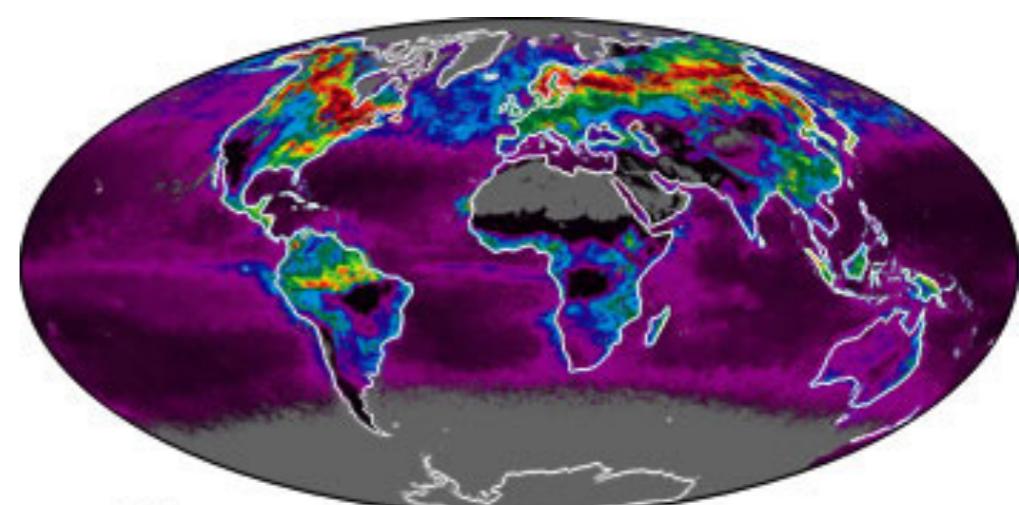
N_2 a compuestos orgánicos
nitrogenados en suelo



Synechococcus
50% del O_2
mundial



Enzimas microbianas degradan
compuestos orgánicos



Net Primary Productivity ($kgC/m^2/year$)
0 1 2 3

Kg de C fijado

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.

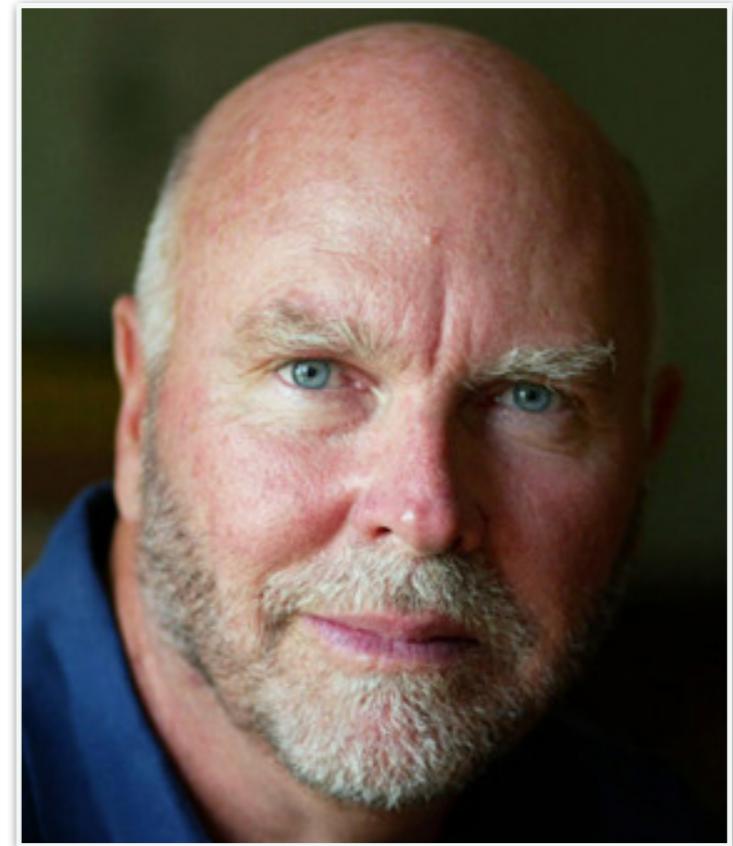
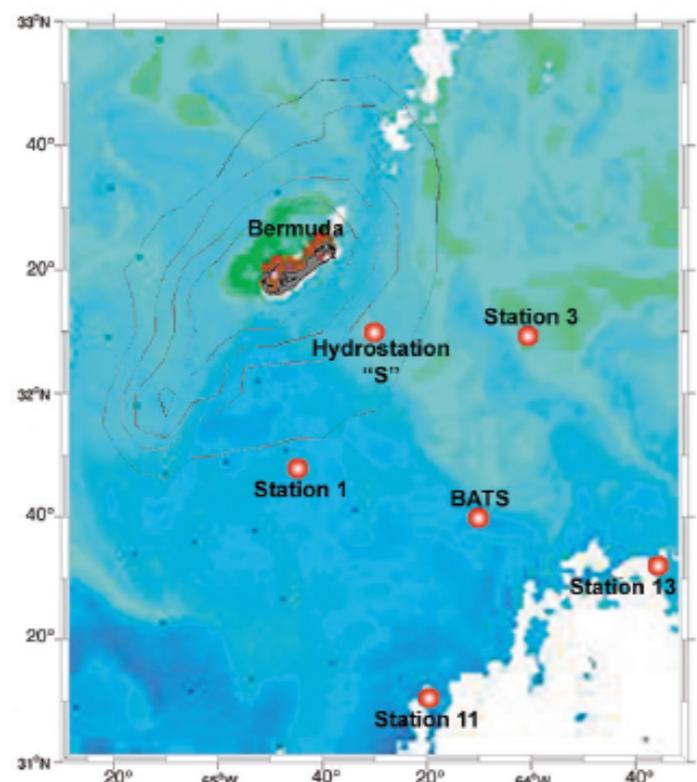
a prefix appearing in loanwords from Greek, with the meanings "after," "along with," "beyond," "among," "behind," and productive in English on the Greek model:
metacarpus; metagenesis.

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

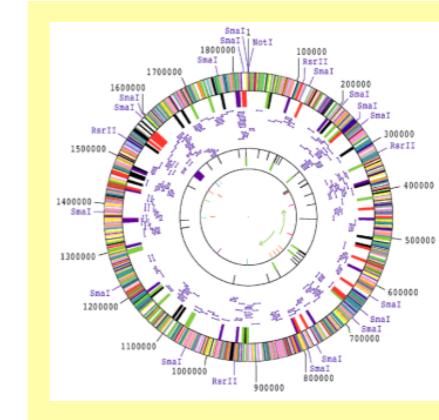
J. Craig Venter^{1,*}, Karin Remington¹, John F. Heidelberg³, Aaron L. Halpern², Doug Rusch², Jonathan A. Eisen³, Dongying Wu³, Ian Paulsen³, Karen E. Nelson³, William Nelson³, Derrick E. Fouts³, Samuel Levy², Anthony H. Knap⁶, Michael W. Lomas⁶, Ken Nealson⁵, Owen White³, Jeremy Peterson³, Jeff Hoffman¹, Rachel Parsons⁶, Holly Baden-Tillson¹, Cynthia Pfannkoch¹, Yu-Hui Rogers⁴, Hamilton O. Smith¹

+ 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

...the “Environmental Genomic” approach...



in situ

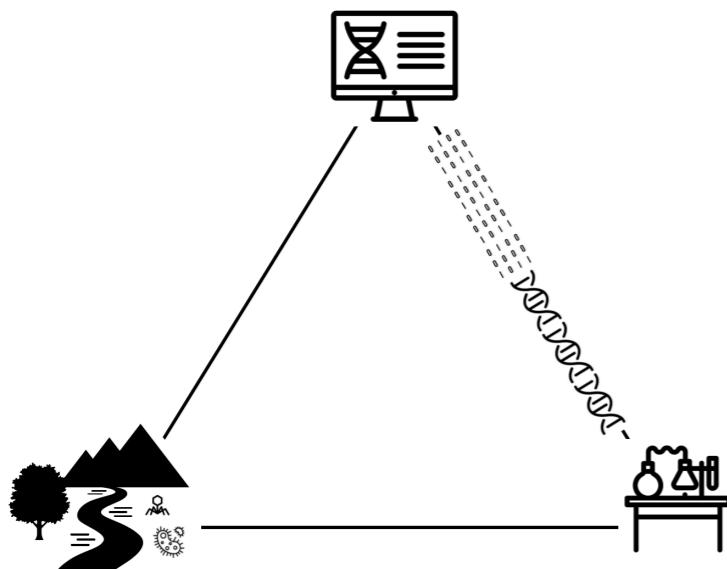
- Field collection
- Geochemistry
- Sample analysis
- Sample preservation

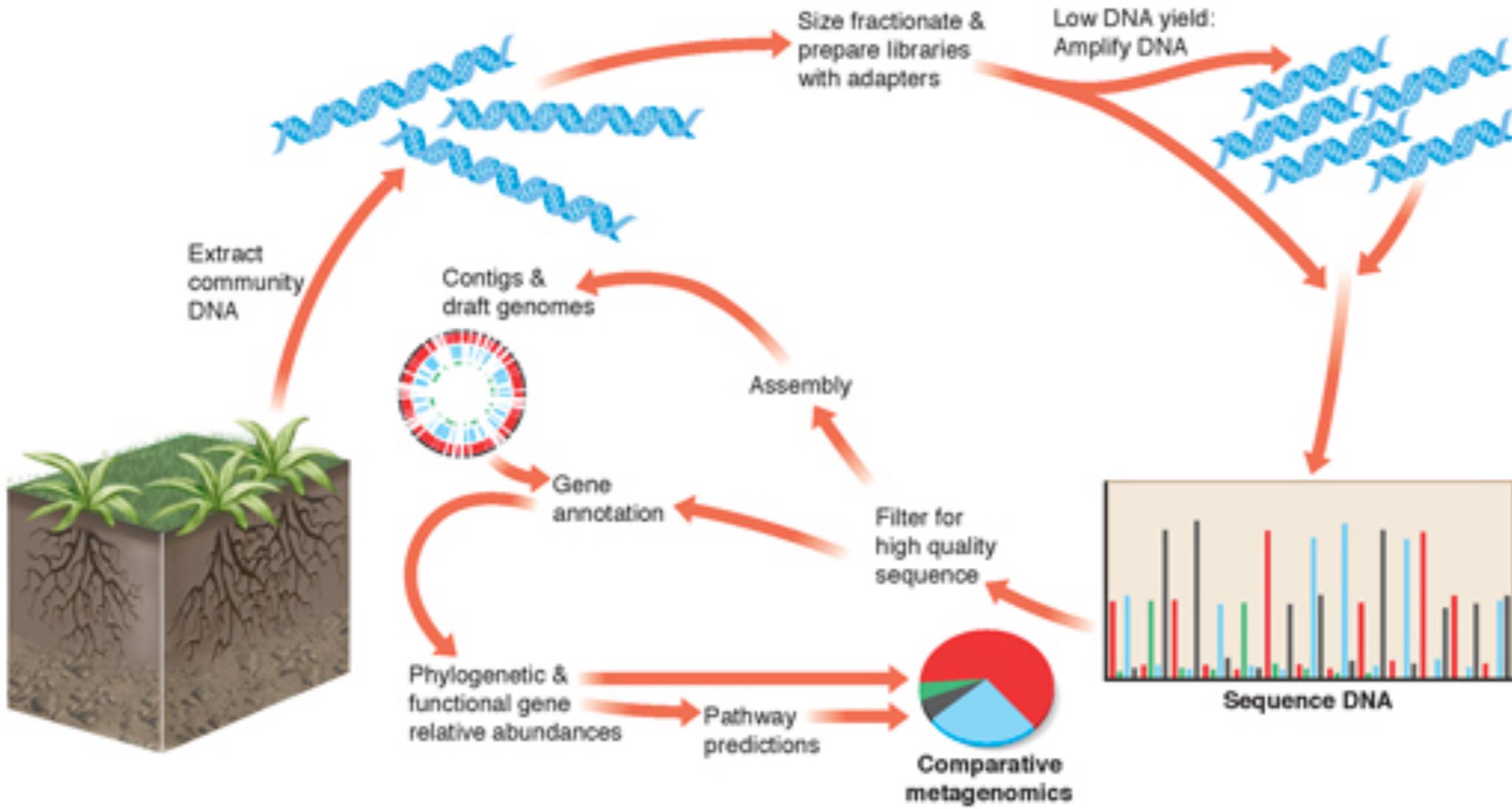
ex vivo

- DNA extraction
- PCR amplification
- Library construction
- RNA extraction
- cDNA synthesis
- DNA sequencing
- Protein extraction
- LC/MS-MS analysis

in silico

- Sequence processing
- Quality control filtering
- Sequence assembly
- Phylogenetic binning
- Functional annotation
- Comparative genomics
- Metabolism/physiology
- mRNA/protein expression



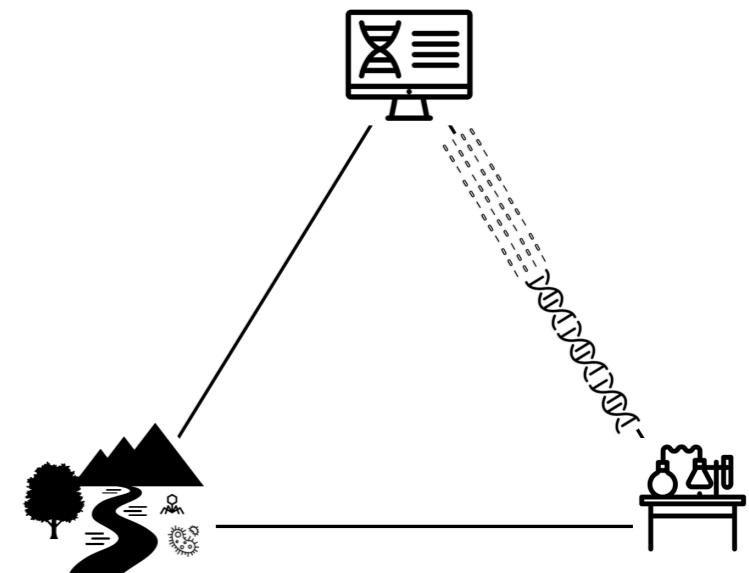


MICROBE

Issue: July 2011, Dr. Jansson

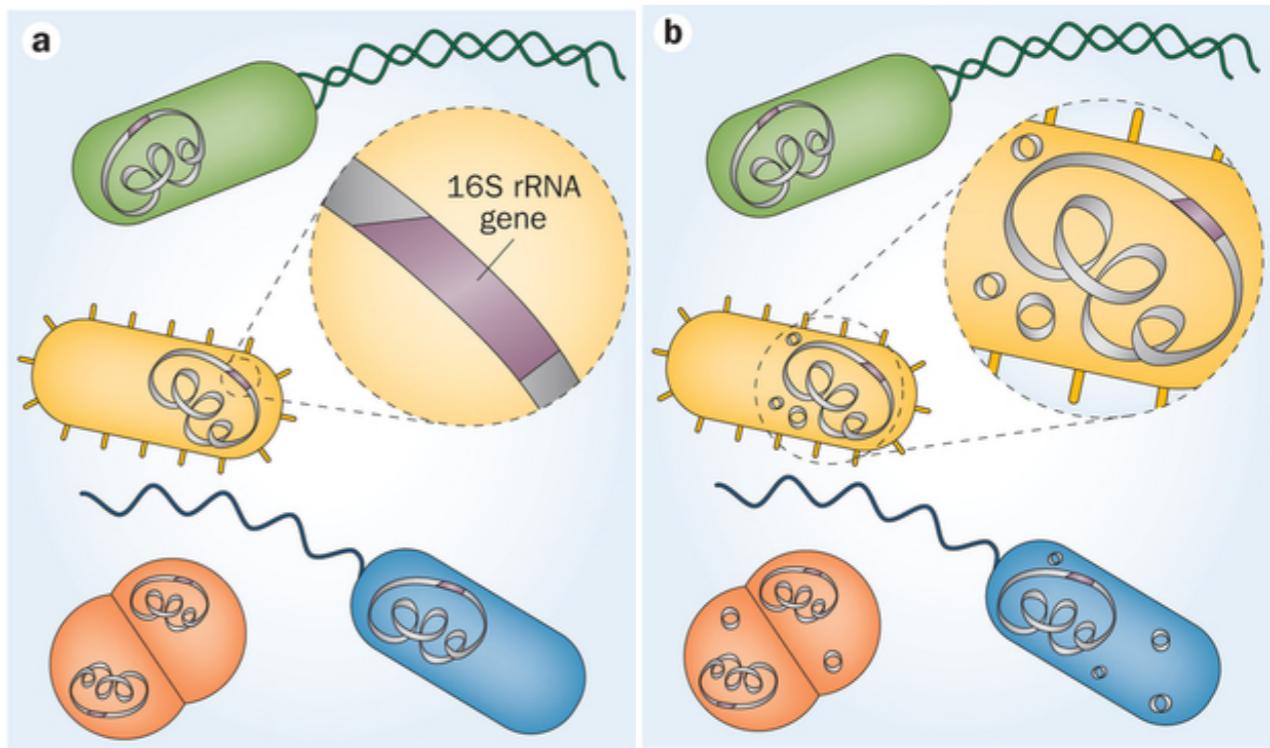
Penumbra Design, Inc. 06/09/11

Fig.#01



Definiciones

- Exploración por gen marcador o metataxonómica.
- Metagenómica.
- Microbioma.
- Metatranscriptómica.



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



Open Access

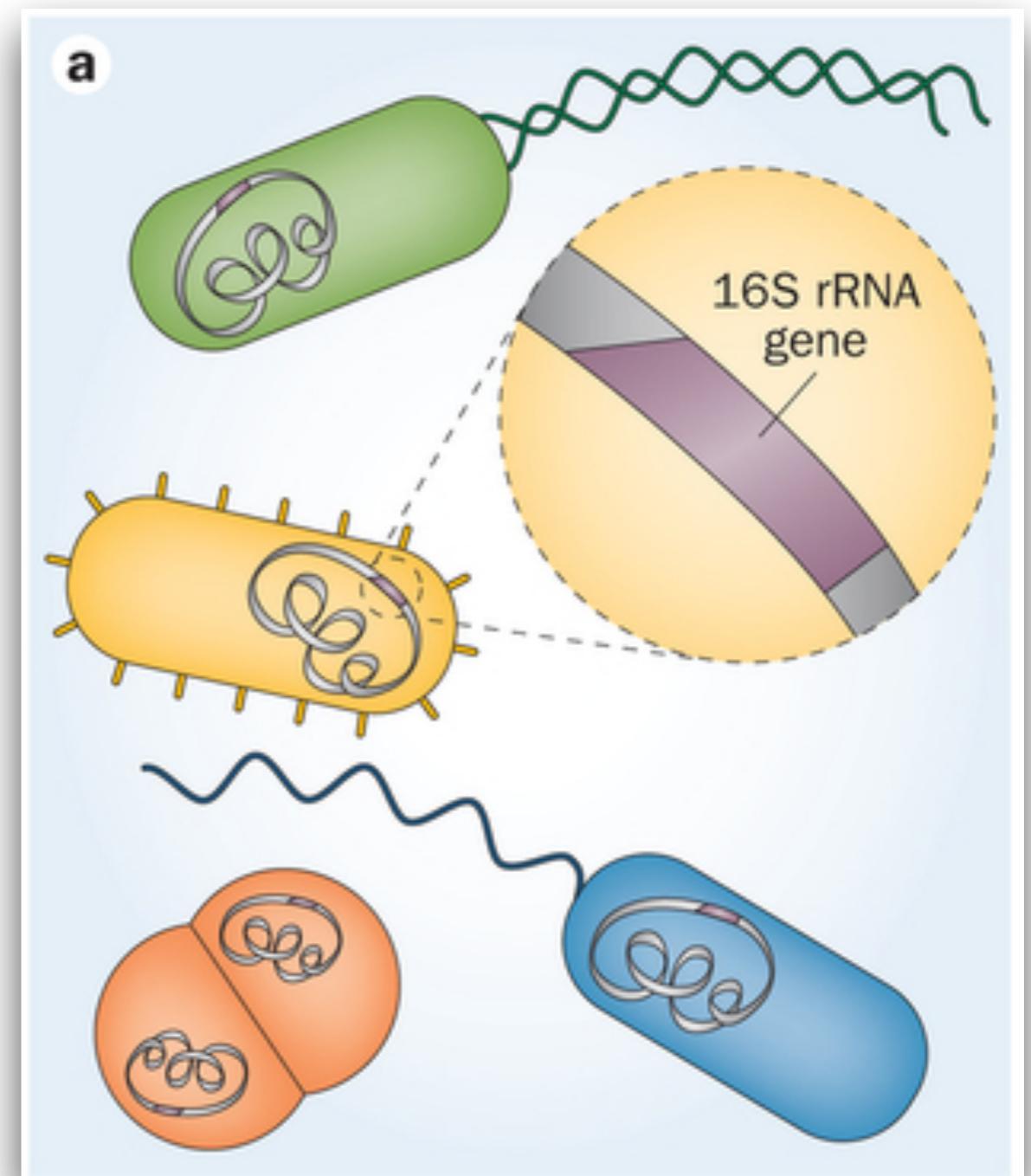
The vocabulary of microbiome research: a proposal

Julian R. Marchesi^{1,2} and Jacques Ravel^{3,4*}



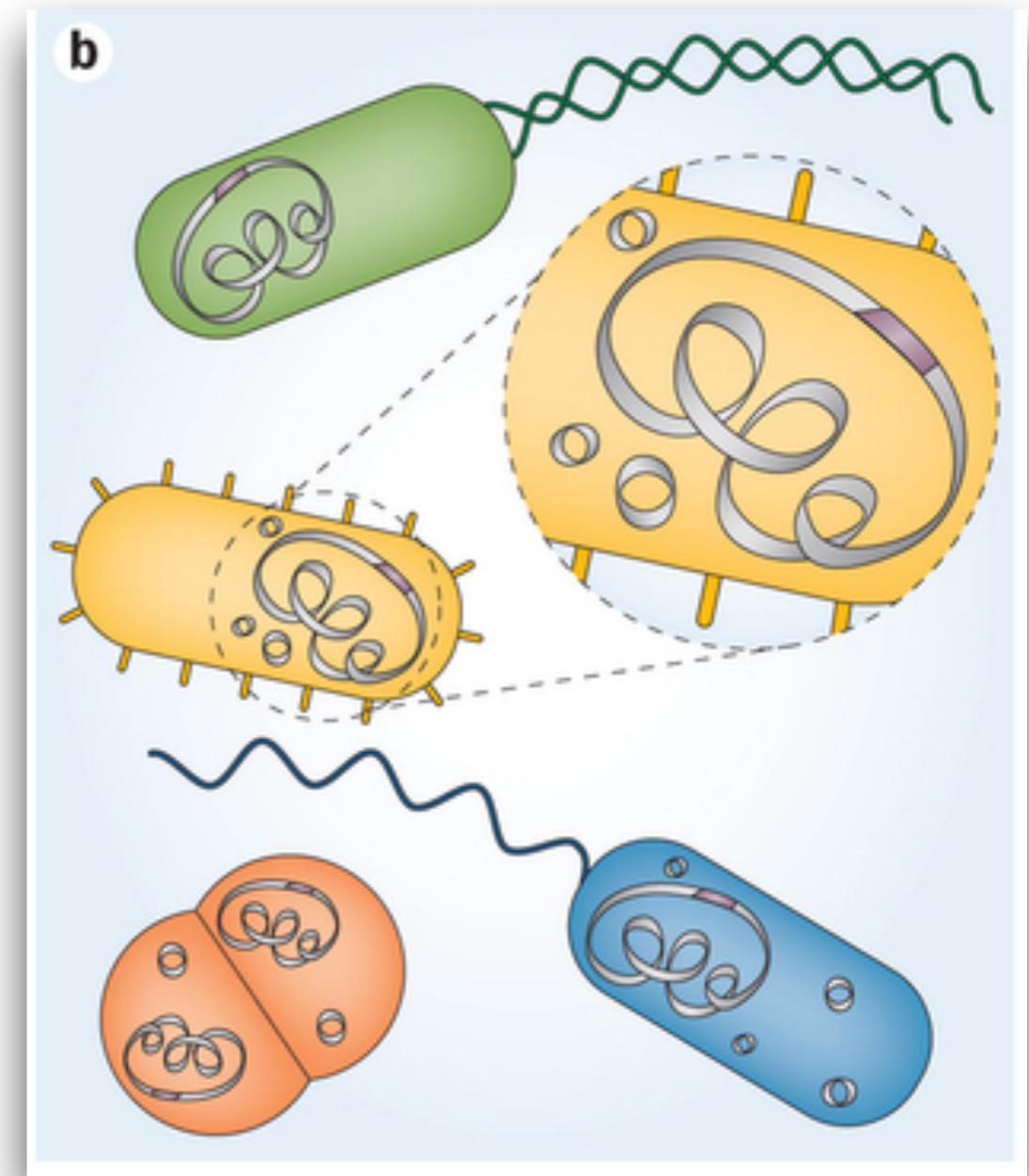
Metataxonómica

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



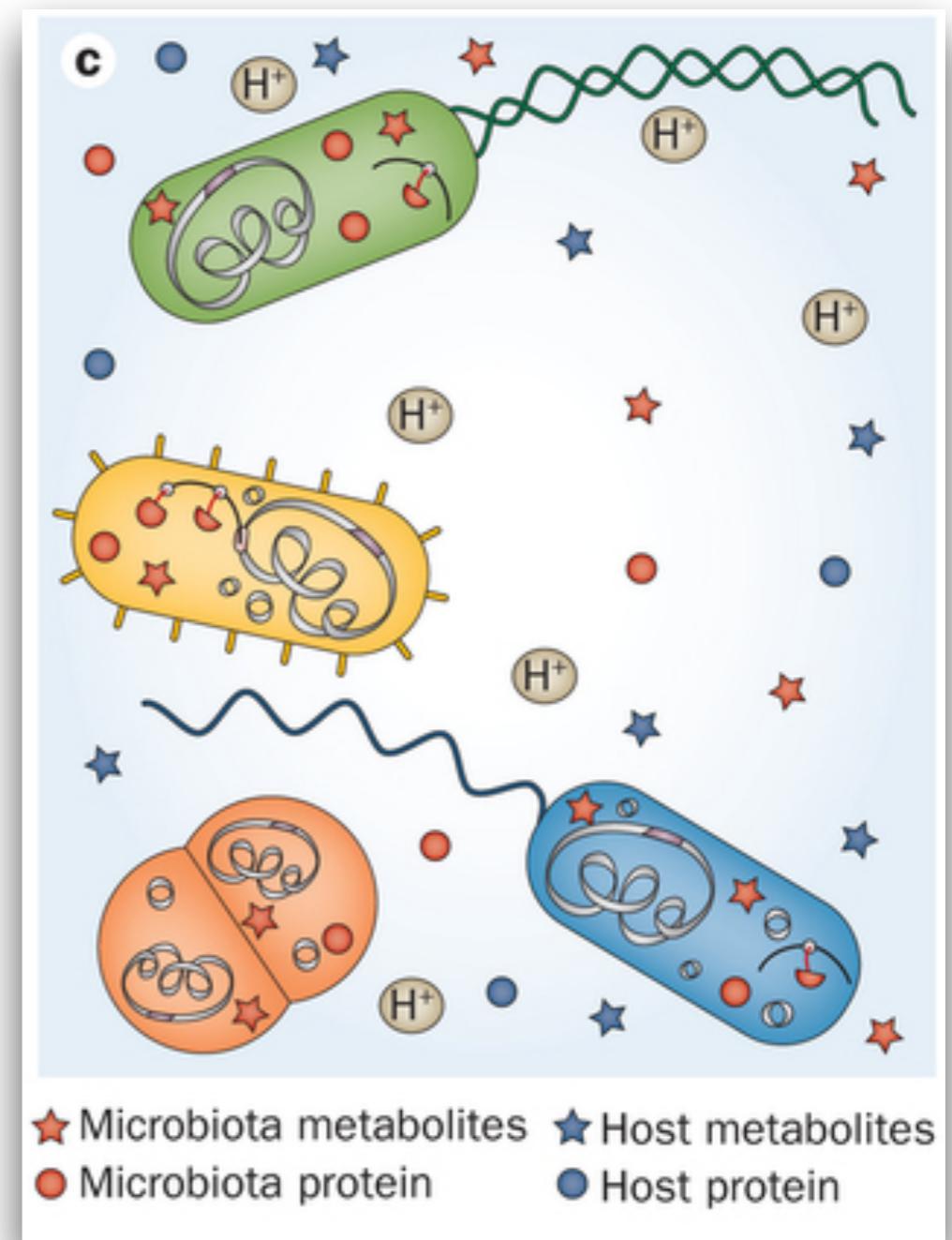
Metagenómica

- Colección de genes y genomas de un ecosistema - más que la taxonomía, también la función - qué hay y qué (potencialmente) hacen.
- Secuenciar todo el DNA (cromosomal, plasmidial, etc.).
- Generar un perfil de miembros del metagenoma.
- Qué hay y en qué proporción.



Microbioma

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

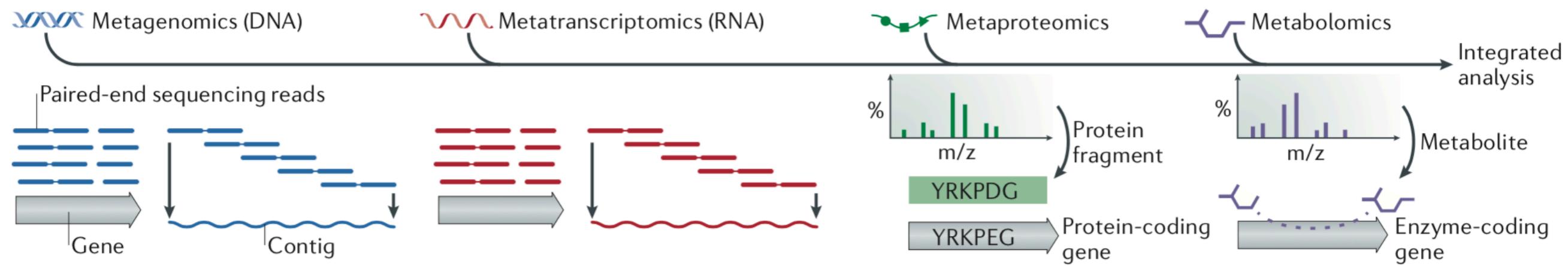


Metatranscriptómica

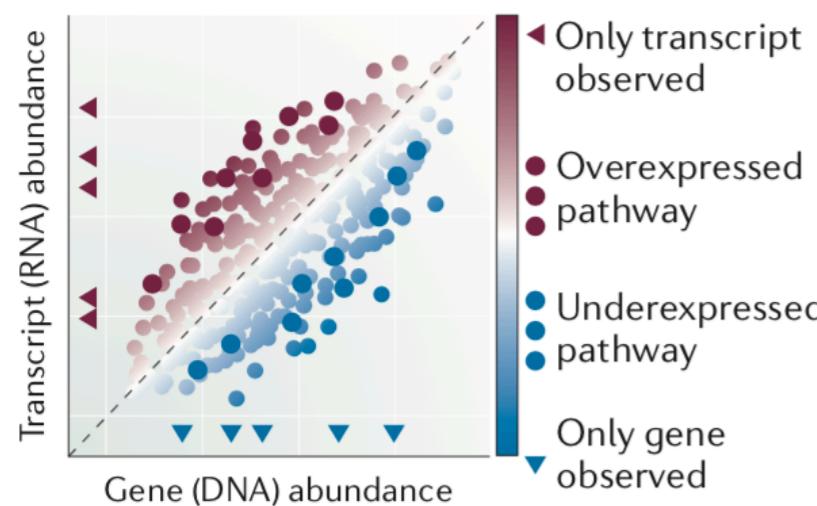
- Secuenciar todo el RNA.
- Generar un perfil de expresión de genes en la comunidad microbiana.
- Qué genes se expresan y en qué medida.

Análisis integrados “omics”

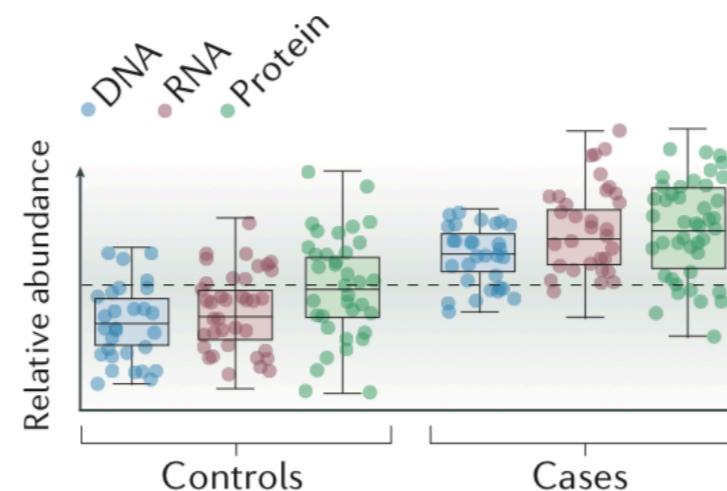
a Multi-omics data types



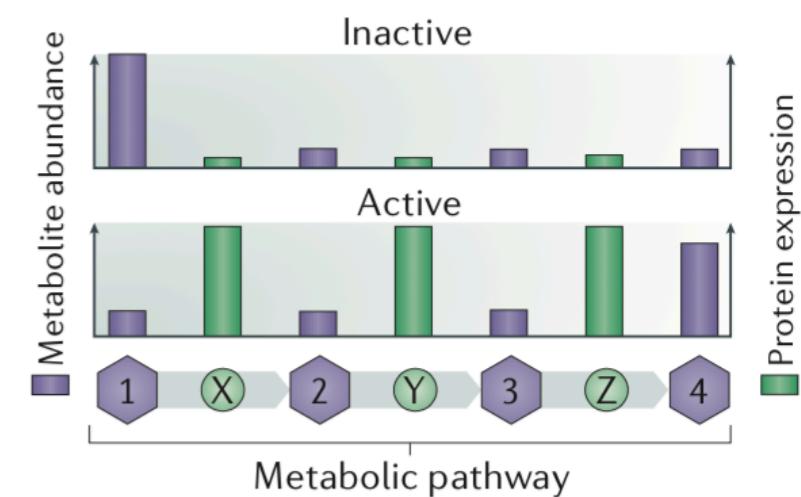
b Normalization



c Strengthening hypotheses

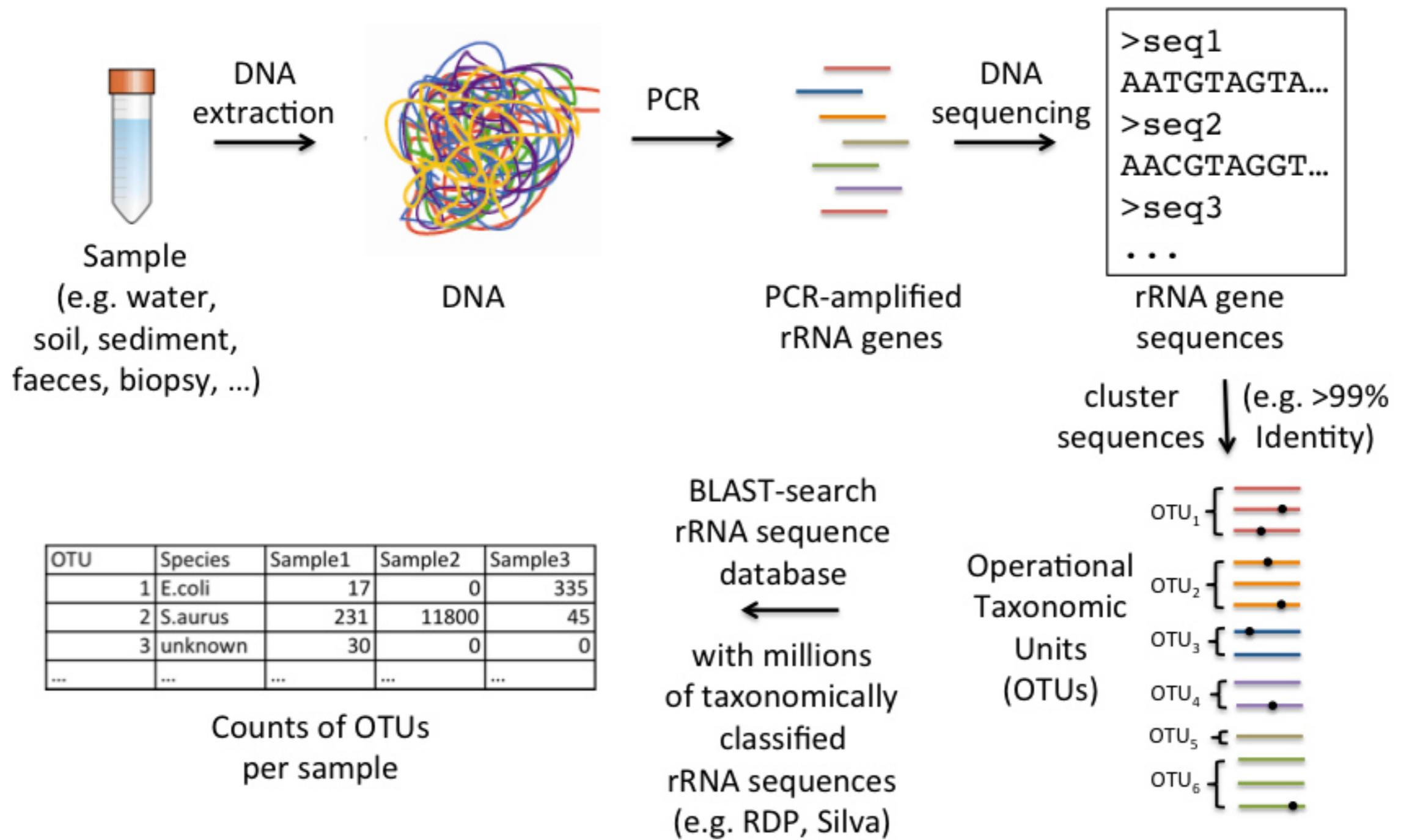


d Descriptive modelling



Metataxonómica ≠ Metagenómica
16S rRNA (amplicones)

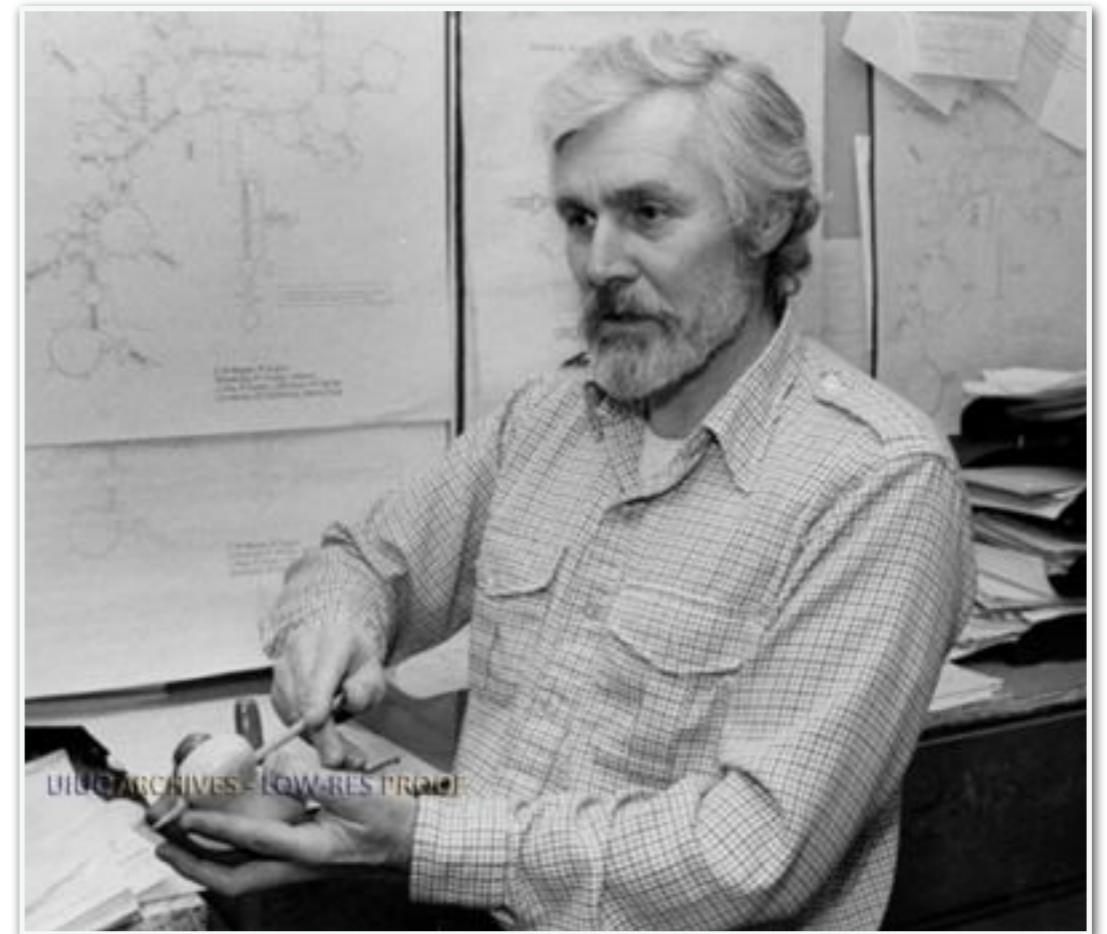
Amplicon Sequencing



AMPLICON SEQUENCING OF RRNA GENES

16S rRNA

- Un gen de copia única o múltiple.
- Históricamente usado para identificar bacterias y arqueas (género).
- Universal (no se transfiere horizontalmente). ¿Por qué universal?



Carl Woese
1928-2012

Teoría de los tres dominios

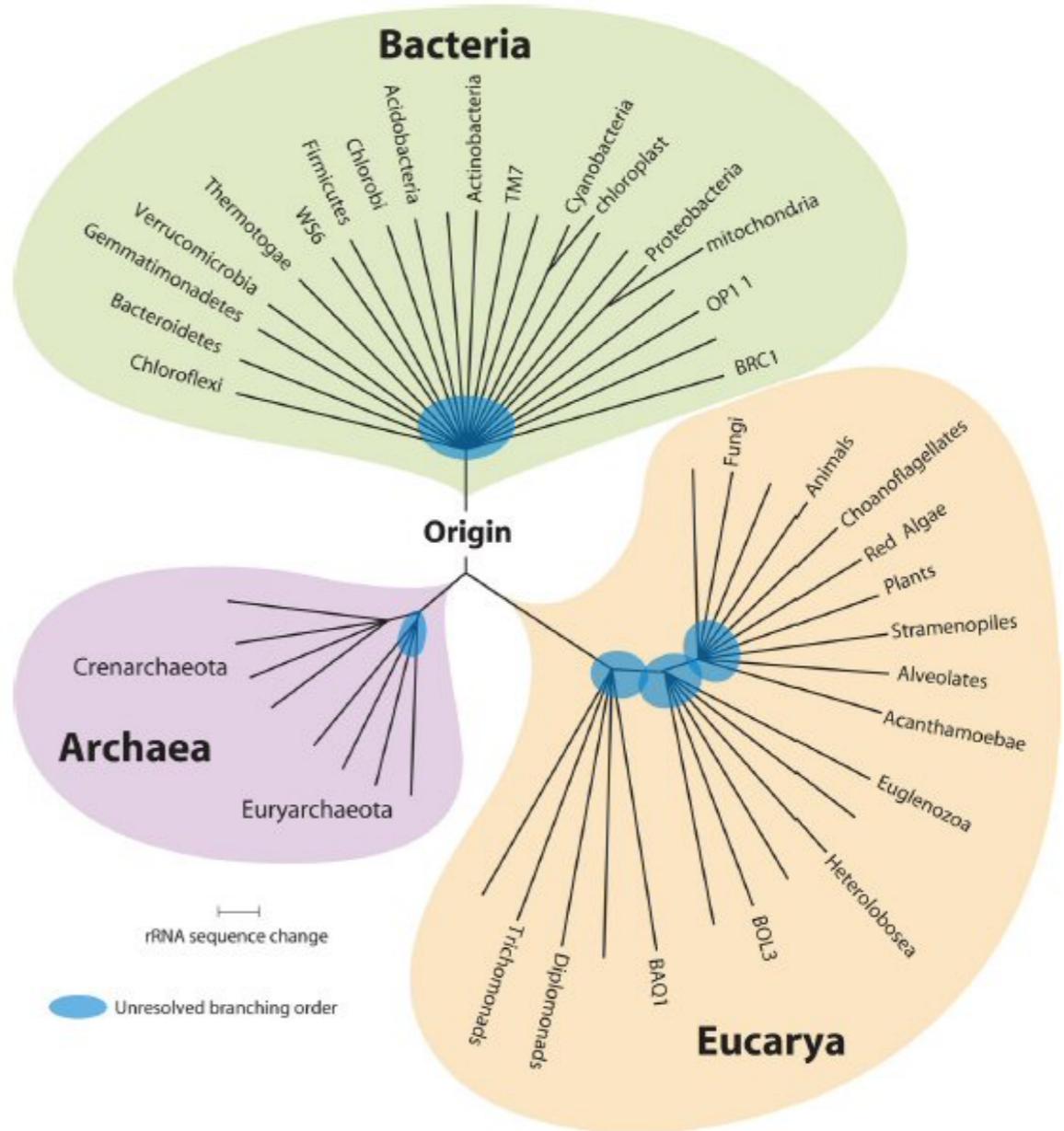
- En 1977 a partir de filogenias de 16/18S
- Arqueas eran consideradas parte de Bacterias

Phylogenetic structure of the prokaryotic domain: The primary kingdoms

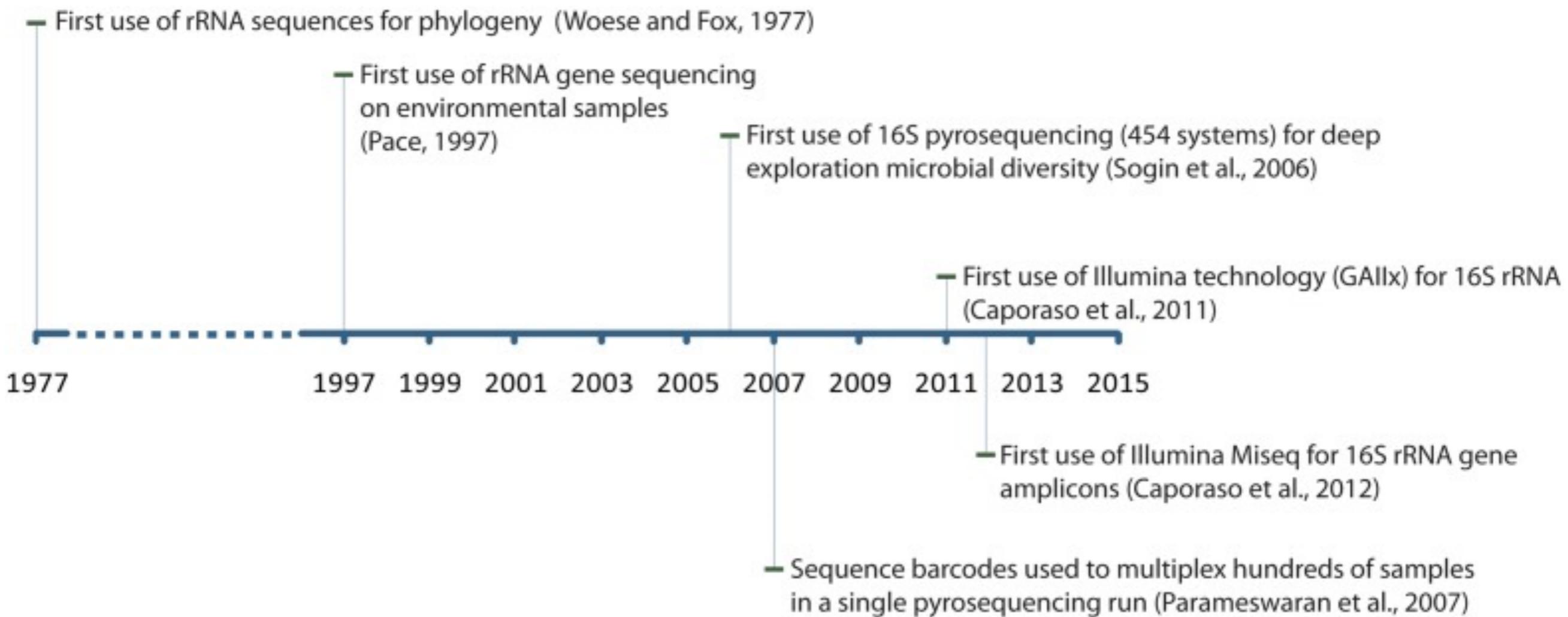
Carl R. Woese and George E. Fox*

Abstract

A phylogenetic analysis based upon ribosomal RNA sequence characterization reveals that living systems represent one of three aboriginal lines of descent: (i) the eubacteria, comprising all typical bacteria; (ii) the archaeabacteria, containing methanogenic bacteria; and (iii) the urkaryotes, now represented in the cytoplasmic component of eukaryotic cells.

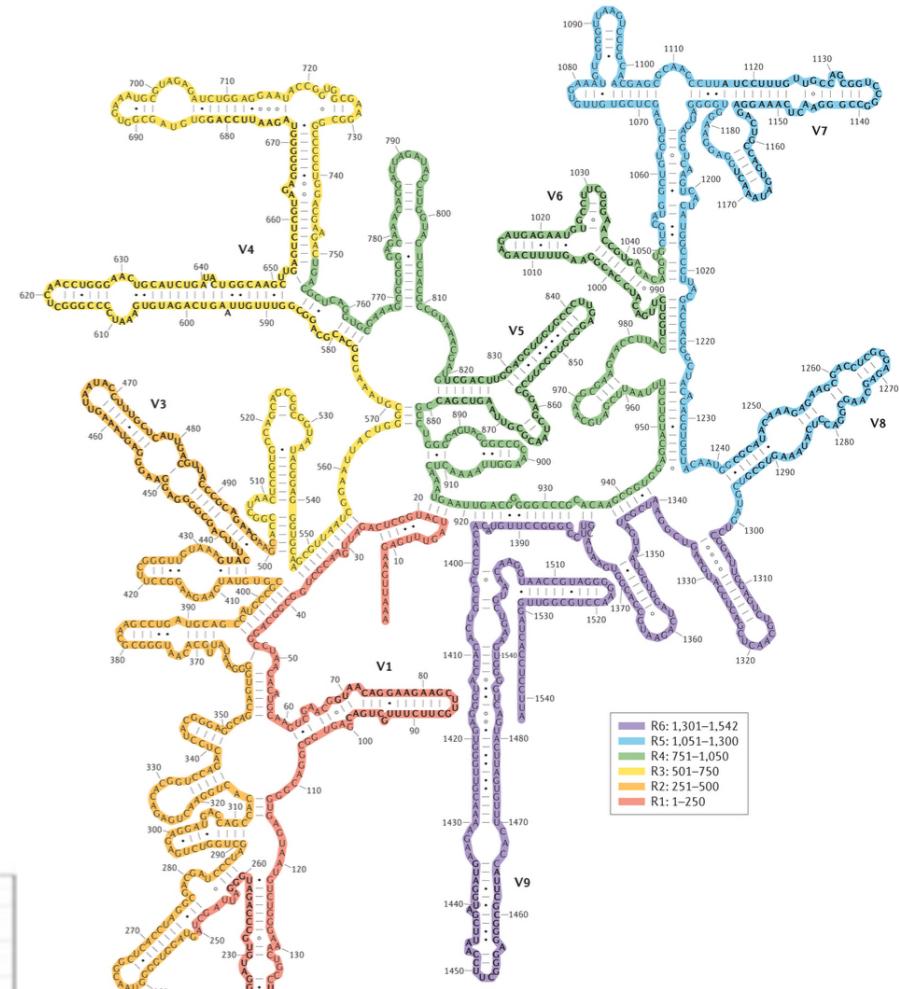
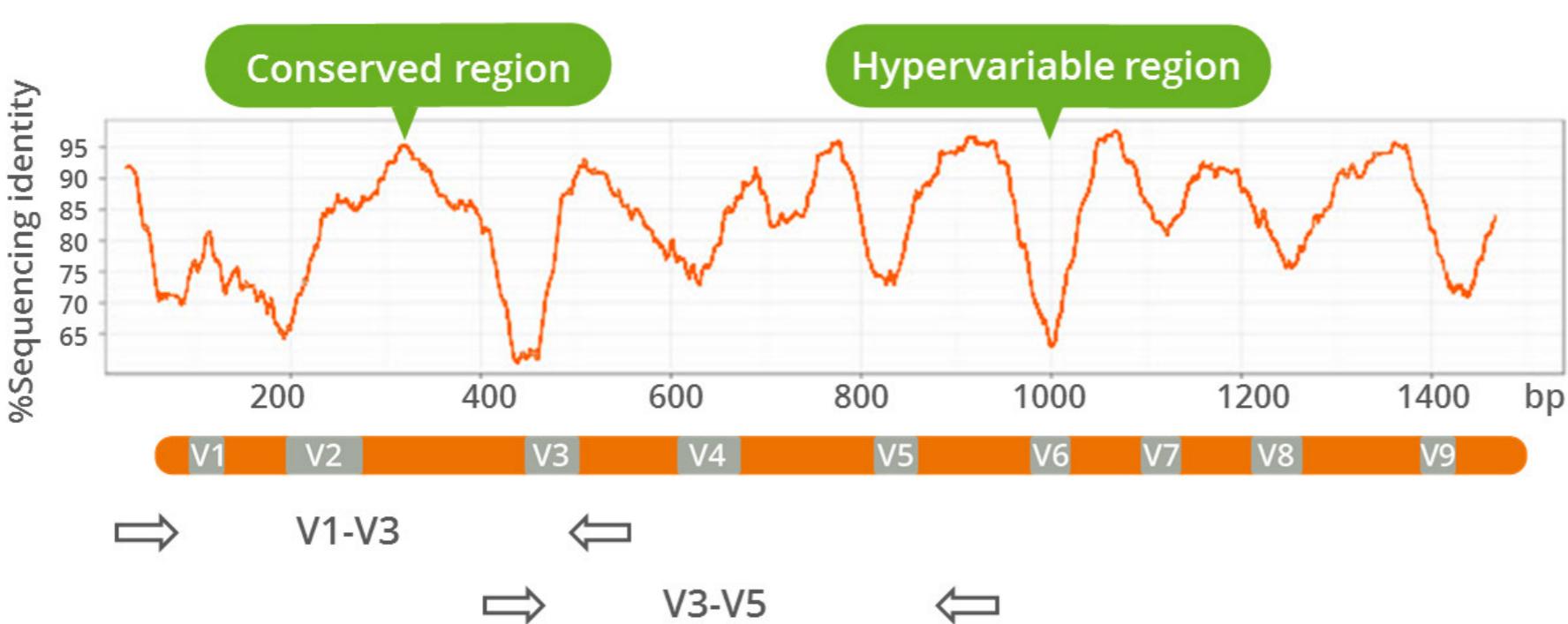


16S rRNA



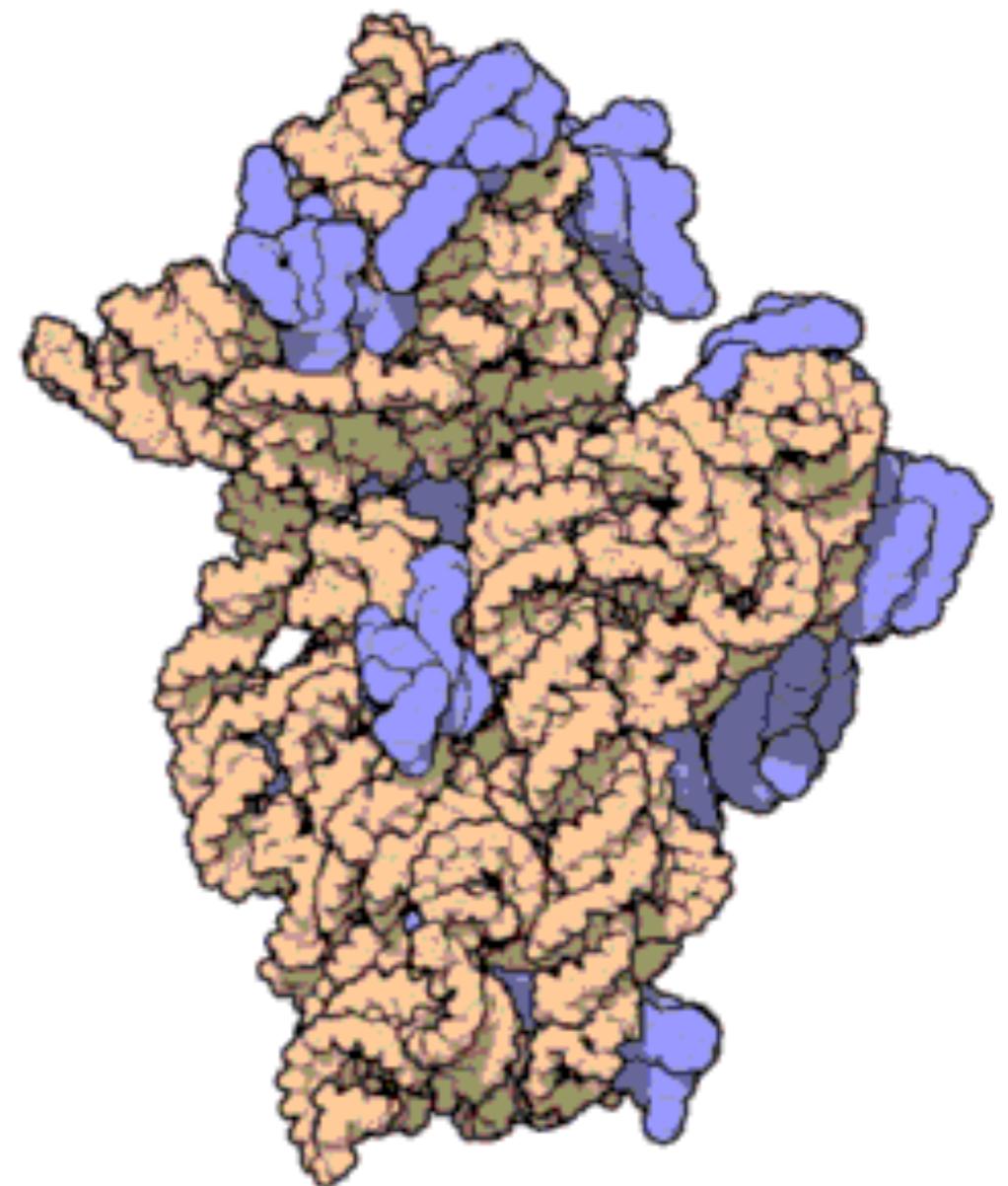
Estructura y ventajas del 16S

- Regiones variables y constantes.
- Bases de datos comprehensivas.



Estructura del 16S

- Parte de la subunidad 30S que se une a la secuencia Shine-Dalgarno (RBS en procariotas).



¿Qué consideraciones prácticas debemos tener para un estudio de comunidades microbianas?

- Poder estadístico.
- Almacenamiento y transporte de muestras.
- Partidores.
- Cantidad de templado.



Poder estadístico

- ¿Cuántas muestras necesito para detectar verdaderos positivos con X probabilidad?

Alpha = 1%	Reads							
Subjects	500	1000	2500	5000	10000	20000	50000	1000000
10	8.57%	9.56%	10.06%	10.98%	10.51%	10.50%	10.62%	10.17%
15	15.88%	17.42%	18.91%	19.55%	19.85%	19.29%	19.32%	20.10%
25	36.36%	38.81%	41.65%	41.65%	42.91%	42.93%	42.66%	43.54%
50	81.81%	85.60%	87.38%	88.16%	87.50%	87.98%	88.30%	88.59%

Alpha = 5%	Reads							
Subjects	500	1000	2500	5000	10000	20000	50000	1000000
10	23.60%	24.60%	26.30%	22.80%	24.50%	28.20%	25.50%	25.70%
15	32.90%	38.70%	38.60%	40.10%	40.00%	39.10%	37.90%	43.00%
25	61.40%	63.50%	63.90%	65.60%	66.40%	64.90%	66.90%	67.10%
50	93.20%	94.80%	96.50%	95.30%	96.50%	95.40%	96.60%	97.40%

doi:10.1371/journal.pone.0052078.t004

Poder estadístico

- ¿Cuántas muestras necesito para detectar verdaderos positivos con X probabilidad?

Bioinformatics. 2015 Aug 1;31(15):2461-8. doi: 10.1093/bioinformatics/btv183. Epub 2015 Mar 29.

Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA.

Kelly BJ¹, Gross R¹, Bittinger K², Sherrill-Mix S², Lewis JD³, Collman RG¹, Bushman FD², Li H³.

A web application for sample size and power calculation in case-control microbiome studies 

Federico Mattiello , Bie Verbist, Karoline Faust, Jeroen Raes, William D. Shannon, Luc Bijnens, Olivier Thas

Bioinformatics (2016) 32 (13): 2038-2040. DOI: <https://doi.org/10.1093/bioinformatics/btw099>

Published: 19 February 2016 Article history ▾

La manera de almacenar las muestras afecta el perfil taxonómico

- 72 h, diversidad alfa y beta cambian con respecto a muestra control.
- Algunos autores reportan efectos poco significativos.

Article | OPEN

Sample storage conditions significantly influence faecal microbiome profiles

Jocelyn M Choo, Lex EX Leong & Geraint B Rogers 

Article | OPEN

Common methods for fecal sample storage in field studies yield consistent signatures of individual identity in microbiome sequencing data

Ran Blekhman , Karen Tang, Elizabeth A. Archie, Luis B. Barreiro, Zachary P. Johnson, Mark E. Wilson, Jordan Kohn, Michael L. Yuan, Laurence Gesquiere, Laura E. Grieneisen & Jenny Tung 

La manera de almacenar las muestras afecta el perfil taxonómico

- 8 semanas, 95% etanol, FTA, omnigene. Gut presentan diferencias comparables a réplicas técnicas.
- T ambiente x 24 h y con RNALater, diferencias inconsistentes entre individuos. Congelar/ descongelar lo mismo.
- A 20, 4, -20, y -80 no se observan cambios en suelo.

Research Article | Applied and Environmental Science

Preservation Methods Differ in Fecal Microbiome Stability, Affecting Suitability for Field Studies

Se Jin Song, Amnon Amir, Jessica L. Metcalf, Katherine R. Amato, Zhenjiang Zech Xu, Greg Humphrey, Rob Knight

M. Denise Dearing, Editor

DOI: 10.1128/mSystems.00021-16

Storage conditions of intestinal microbiota matter in metagenomic analysis

Silvia Cardona, Anat Eck, Montserrat Cassellas, Milagros Gallart, Carmen Alatrue, Joel Dore, Fernando Azpiroz, Joaquim Roca, Francisco Guarner and Chaysavanh Manichanh 

BMC Microbiology 2012 12:158 | DOI: 10.1186/1471-2180-12-158 | © Cardona et al.; licensee BioMed Central Ltd. 2012

Received: 6 March 2012 | Accepted: 20 July 2012 | Published: 30 July 2012

RESEARCH LETTER

Effect of storage conditions on the assessment of bacterial community structure in soil and human-associated samples

Christian L. Lauber¹, Nicholas Zhou², Jeffrey I. Gordon³, Rob Knight⁴ & Noah Fierer^{1,2}

¹Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA; ²Department of Ecology and Biology, University of Colorado, Boulder, CO, USA; ³Center for Genome Sciences, Washington University School of Medicine, St. Louis,

⁴Department of Chemistry and Biochemistry, University of Colorado, Boulder, CO, USA

Misma muestra, distinta extracción de DNA, distinto perfil taxonómico

- Consenso en que el método de extracción afecta la composición microbiana resultante.

Original Research

DNA extraction protocols cause differences in 16S rRNA amplicon sequencing efficiency but not in community profile composition or structure

Benjamin E. R. Rubin , Jon G. Sanders, Jarrad Hampton-Marcell, Sarah M. Owens, Jack A. Gilbert, Corrie S. Moreau

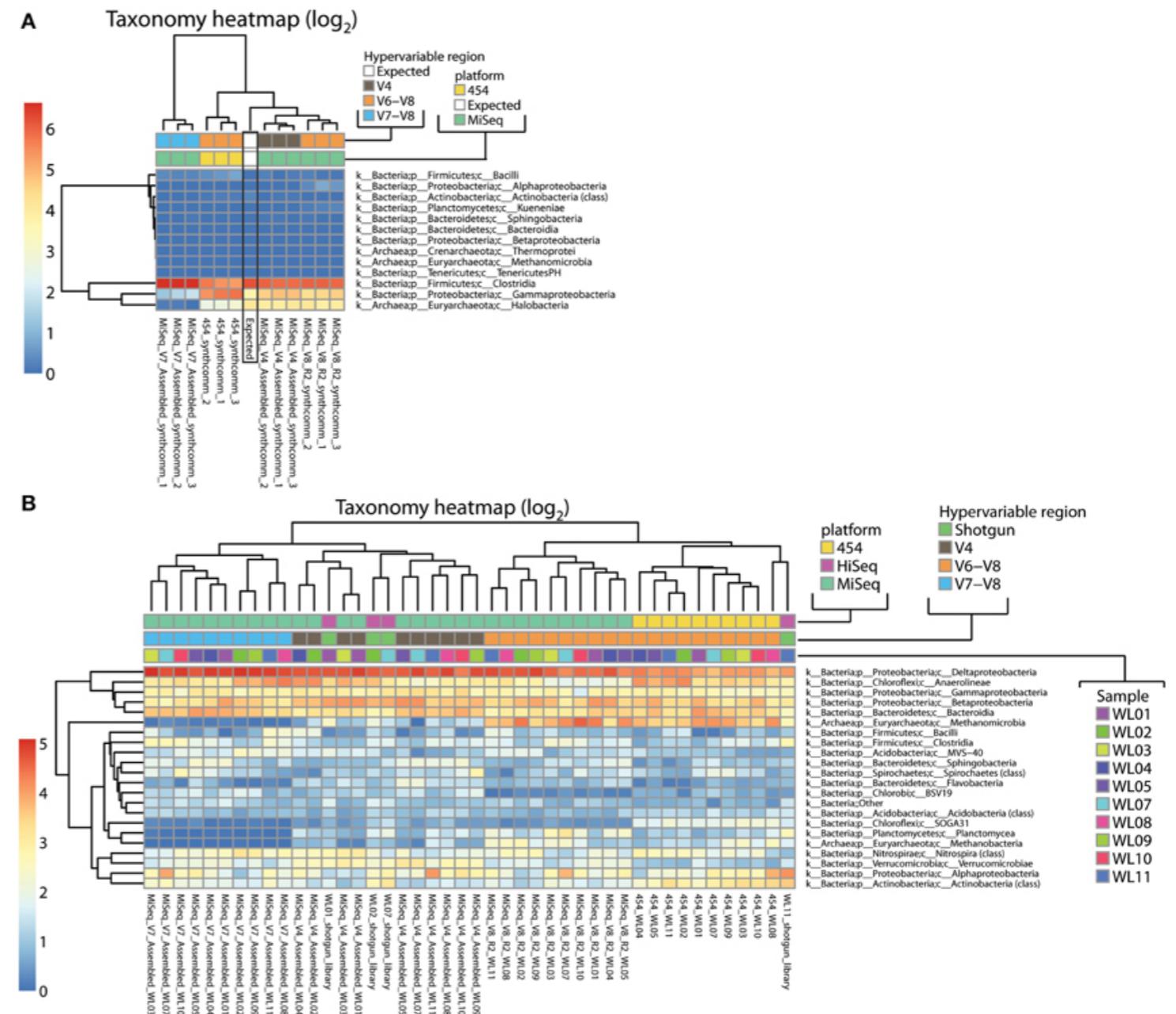
The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies

J Paul Brooks , David J Edwards, Michael D Harwich Jr, Maria C Rivera, Jennifer M Fettweis, Myrna G Serrano, Robert A Reris, Nihar U Sheth, Bernice Huang, Philippe Girerd, Vaginal Microbiome Consortium (additional members), Jerome F Strauss III, Kimberly K Jefferson and Gregory A Buck

BMC Microbiology 2015 15:66 | DOI: 10.1186/s12866-015-0351-6 | © Brooks et al.; licensee BioMed Central. 2015
Received: 17 September 2014 | Accepted: 16 January 2015 | Published: 21 March 2015

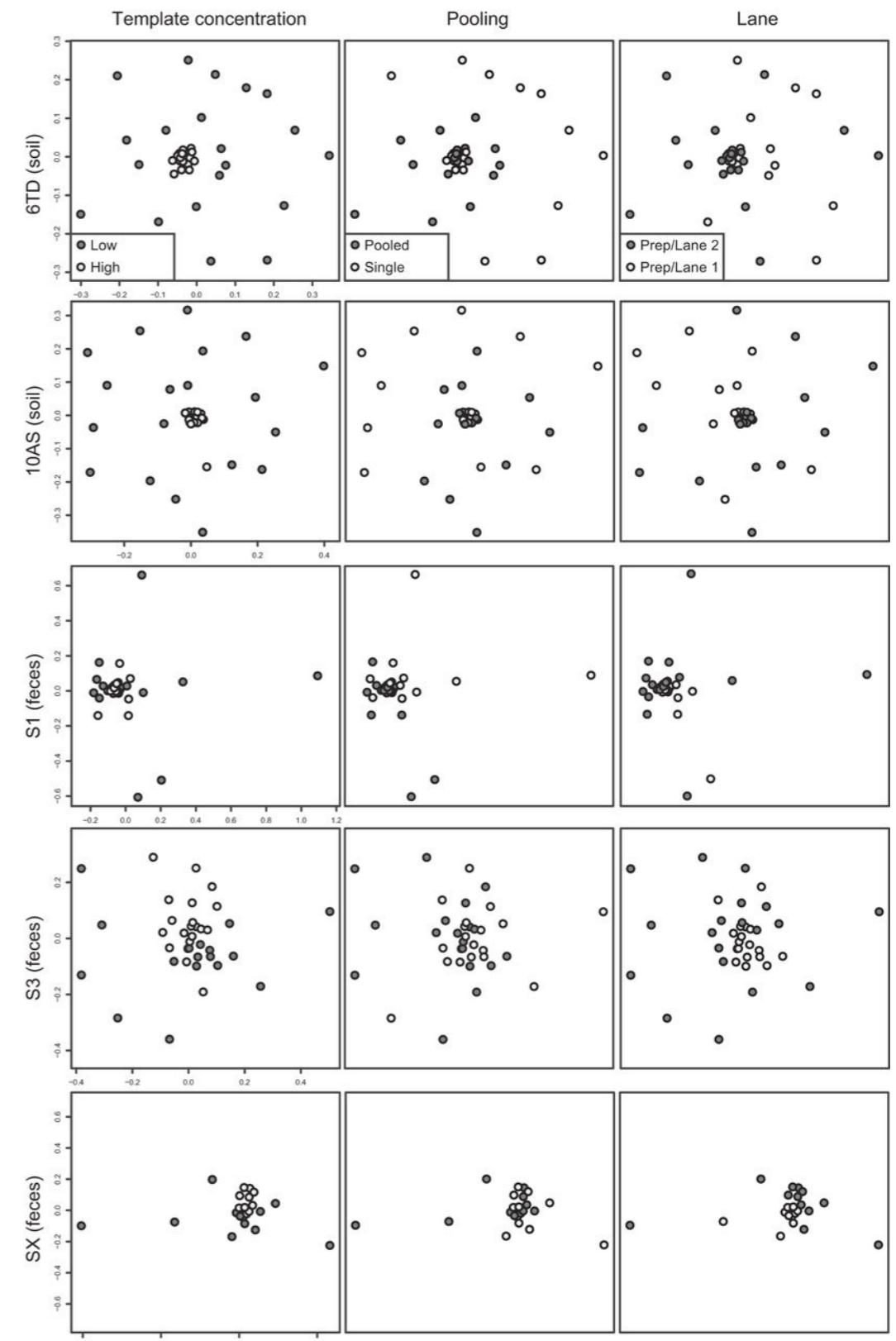
Partidores tienen un efecto mayor en perfil taxonómico

- Distintos partidores tienen preferencia por distintos grupos taxonómicos.
- Taxonomía cambia si primers cambian.
- Plataforma de secuenciación tiene mínimo efecto cuando se usan los mismos partidores.



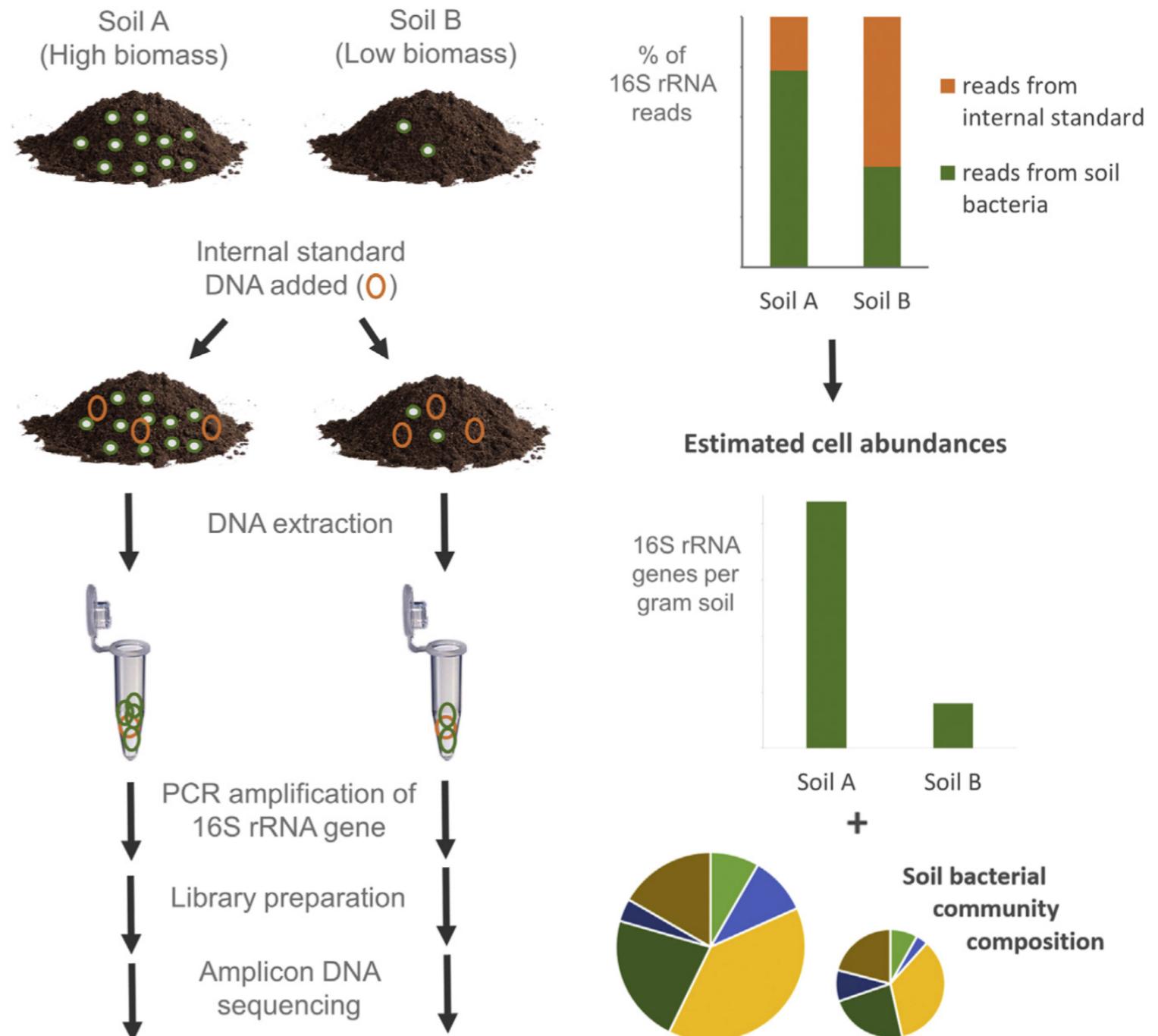
Cantidad de templado podría afectar la precisión de la identificación taxonómica

- Distintas concentraciones de templado introducen variabilidad en la composición de microbiota.
- Altas concentraciones de templado (i.e., 5-10 ng) incrementan la precisión de la identificación taxonómica, en comparación con bajas concentraciones (i.e., 0.1 ng).



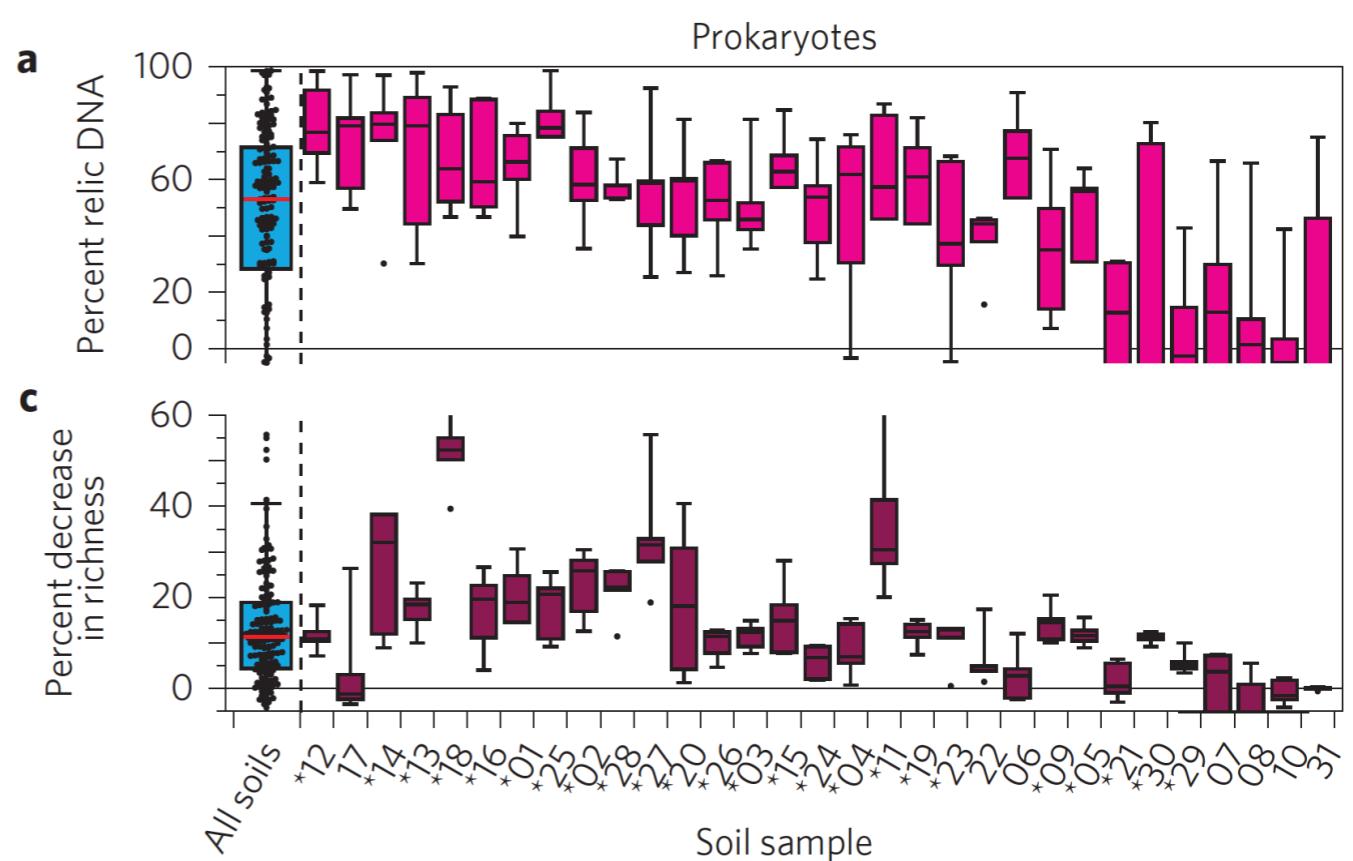
Siempre cuantificamos de manera relativa, a menos que...

- Cuantificación es siempre relativa a menos que se introduzcan controles internos.



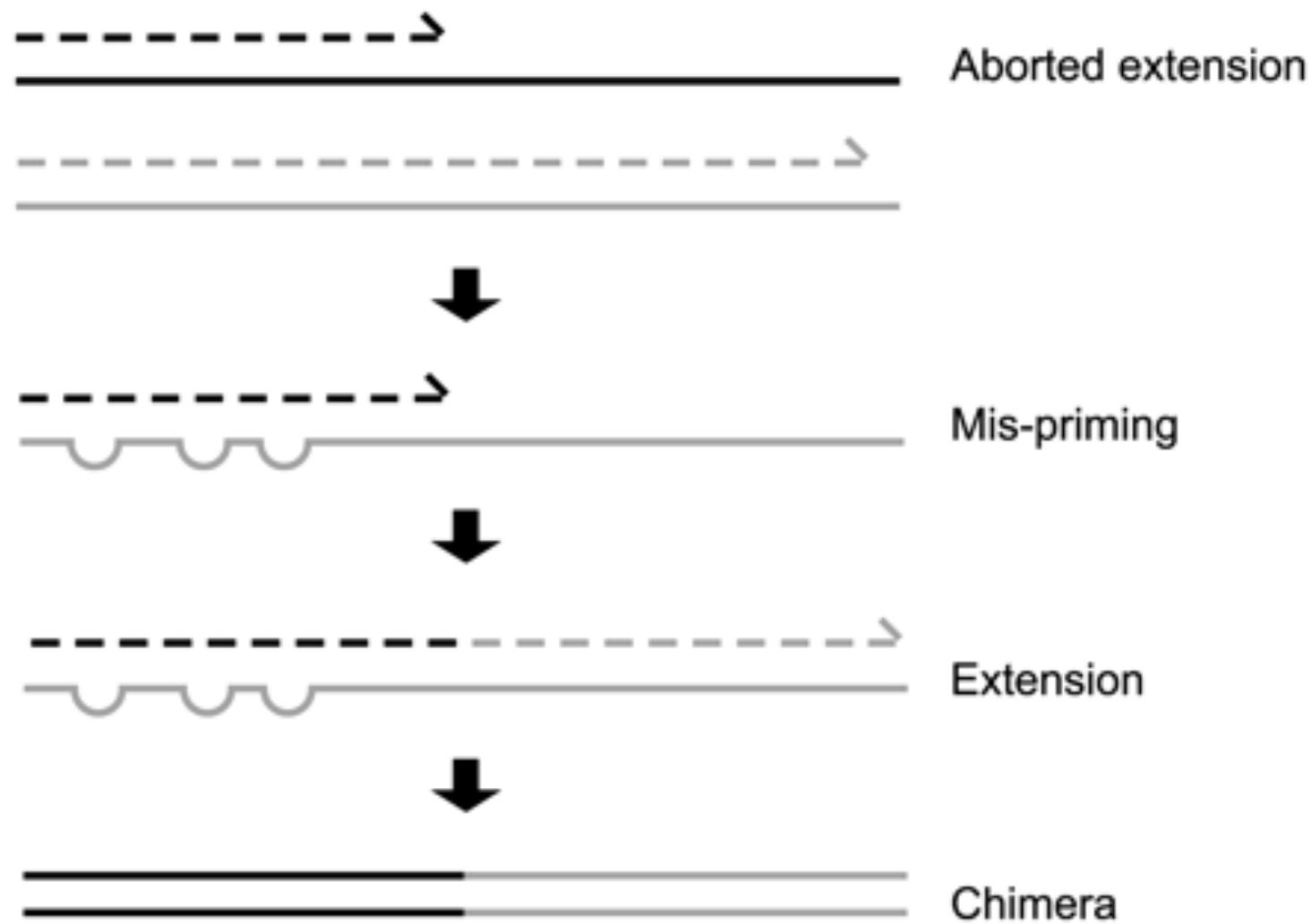
Secuenciamos el DNA de células vivas y muertas, DNA extra celular, etc.

- Hasta un 40% del DNA puede venir de células no intactas o extracelular.
- Infla estimados de diversidad alfa.
- Estima mal abundancia taxonómica.



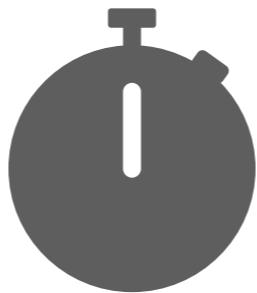
Errores en la PCR aumentan la diversidad

- Infla estimados de diversidad.
- Afecta filogenias.
- Organismos nuevos falsos.
- Más prevalente en organismos menos abundantes (hasta 70%).



> Pausa <

10 minutos

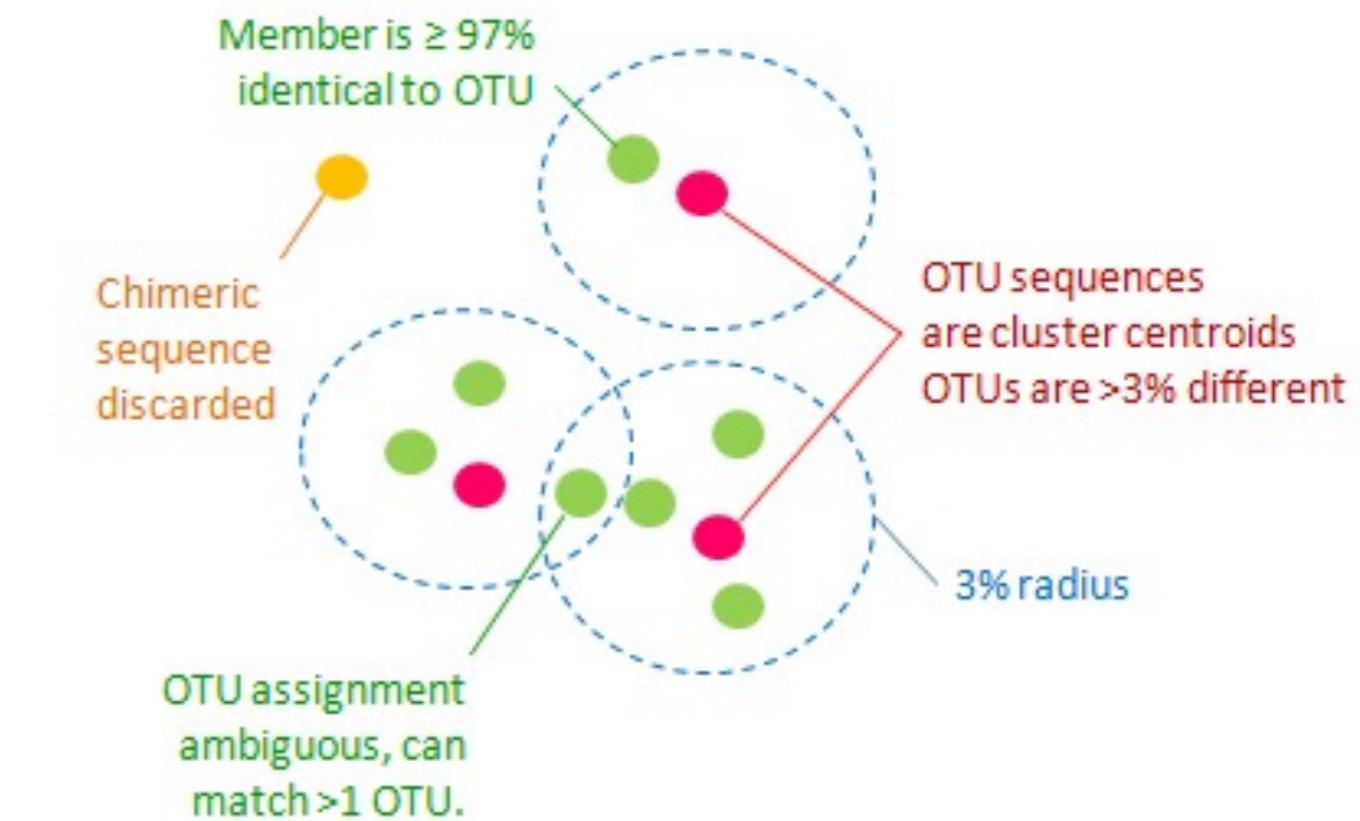


¿Cómo analizamos
muestras de 16S rRNA?

Después de la
secuenciación...

Operational Taxonomic Units

- Unidades taxonómicas operacionales.
- ¿Por qué no especies?
- Mismo 16S, diferente genoma.
- Misma especie, diferente 16S.
- OTUs son clusters de secuencias que divergen como máximo a X% de identidad (3% normalmente).



[Clin Microbiol Rev.](#) 2004 Oct; 17(4): 840–862.
doi: [10.1128/CMR.17.4.840-862.2004](https://doi.org/10.1128/CMR.17.4.840-862.2004)

PMCID: PMC523561

Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases

Jill E. Clarridge, III*

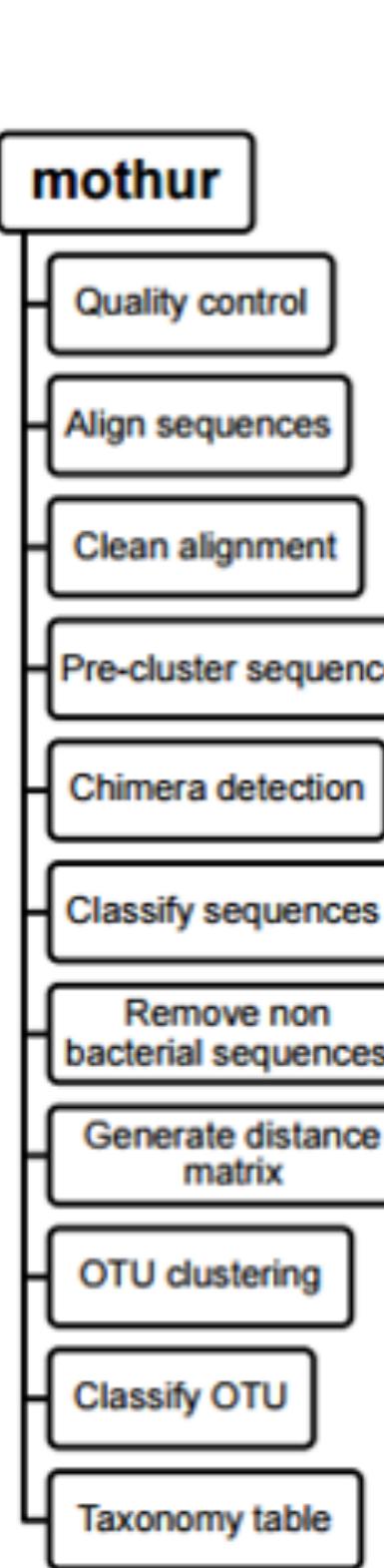
[Author information ▶](#) [Copyright and License information ▶](#)

[Appl Environ Microbiol.](#) 2004 Aug; 70(8): 4831–4839.
doi: [10.1128/AEM.70.8.4831-4839.2004](https://doi.org/10.1128/AEM.70.8.4831-4839.2004)

Ecological Significance of Microdiversity: Identical 16S rRNA Gene Sequences Can Be Found in Bacteria with Highly Divergent Genomes and Ecophysiolgies

Elke Jaspers† and Jörg Overmann *

PMCID: PMC492463



mothur

- Clustering por similitud de secuencia
97% identidad.
- Formar OTUs.
- Tomar representante.
- Clasificar reads.

Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities

[PD Schloss, SL Westcott, T Ryabin... - Applied and ..., 2009 - Am Soc Microbiol](#)

ABSTRACT mothur aims to be a comprehensive software package that allows users to use a single piece of software to analyze community sequence data. It builds upon previous tools to provide a flexible and powerful software package for analyzing sequencing data. As a case study, we used mothur to trim, screen, and align sequences; calculate distances; assign sequences to operational taxonomic units; and describe the α and β diversity of ...

[Cited by 6741](#) [Related articles](#) [All 20 versions](#) [Import into EndNote](#) [Save](#) [More](#)



www.mothur.org

¿Cómo clasificamos las reads?

- Alineamiento estructural en contra de una base de datos curada.
- SILVA más popular - más de 20 años - más de 2,000 citas.
- > 600 mil secuencias.

SILVA SSU / LSU 128 - full release

	SSU Parc	SSU Ref	SSU Ref NR 99	LSU Parc	LSU Ref
Minimal length	300	1200/900	1200/900	300	1900
Quality filtering	basic	strong	strong	basic	strong
Guide Tree	no	no	yes	no	yes
Release date	28.09.16	28.09.16	28.09.16	28.09.16	28.09.16
Aligned rRNA sequences	5,616,941	1,922,213	645,151	735,238	154,297

[HTML] The **SILVA ribosomal RNA gene** database project: improved data processing and web-based tools

C Quast, E Pruesse, P Yilmaz, J Gerken... - Nucleic acids ..., 2013 - academic.oup.com

Abstract **SILVA** (from Latin **silva**, forest, <http://www.arb-silva.de>) is a comprehensive web resource for up to date, quality-controlled databases of aligned **ribosomal RNA (rRNA) gene** sequences from the Bacteria, Archaea and Eukaryota domains and supplementary online

Cited by 2107 Related articles All 17 versions Import into EndNote Save More

Operational Taxonomic Units

RStudio Source Editor

tax_table(physeq) * Filter

	Kingdom	Phylum	Class	Order	Family	Genus
Otu08098	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu17259	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales_unclassified	NA
Otu07427	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu08947	Bacteria	Firmicutes	Firmicutes_unclassified	Firmicutes_unclassified	Firmicutes_unclassified	NA
Otu22896	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu28542	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales_unclassified	NA
Otu06520	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae	NA
Otu21167	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	Hallella
Otu21163	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu16680	Bacteria	Proteobacteria	Epsilonproteobacteria	Campylobacterales	Helicobacteraceae	Helicobacter
Otu05526	Bacteria	Chlamydiae	Chlamydiae	Chlamydiales	Chlamydiateae	Chlamydia
Otu24307	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales_unclassified	NA
Otu19152	Bacteria	Firmicutes	Clostridia	Clostridiales	Ruminococcaceae	Pseudoflavorifractor
Otu20183	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Marvinbryantia
Otu20140	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Bacteroidales_unclassified	NA
Otu30064	Bacteria	Bacteroidetes	Bacteroidia	Bacteroidales	Prevotellaceae	NA
Otu32335	Bacteria	Proteobacteria	Proteobacteria_unclassified	Proteobacteria_unclassified	Proteobacteria_unclassified	NA
Otu18835	Bacteria	Firmicutes	Clostridia	Clostridiales	Clostridiales_unclassified	NA
Otu30638	Bacteria	Bacteria_unclassified	Bacteria_unclassified	Bacteria_unclassified	Bacteria_unclassified	NA
Otu18095	Bacteria	Bacteroidetes	Bacteroidetes_unclassified	Bacteroidetes_unclassified	Bacteroidetes_unclassified	NA
Otu20049	Bacteria	Bacteroidetes	Bacteroidetes_unclassified	Bacteroidetes_unclassified	Bacteroidetes_unclassified	NA
Otu30558	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu31046	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA
Otu32321	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	Clostridium_XIVb
Otu30917	Bacteria	Firmicutes	Clostridia	Clostridiales	Lachnospiraceae	NA

Showing 1 to 25 of 32,361 entries

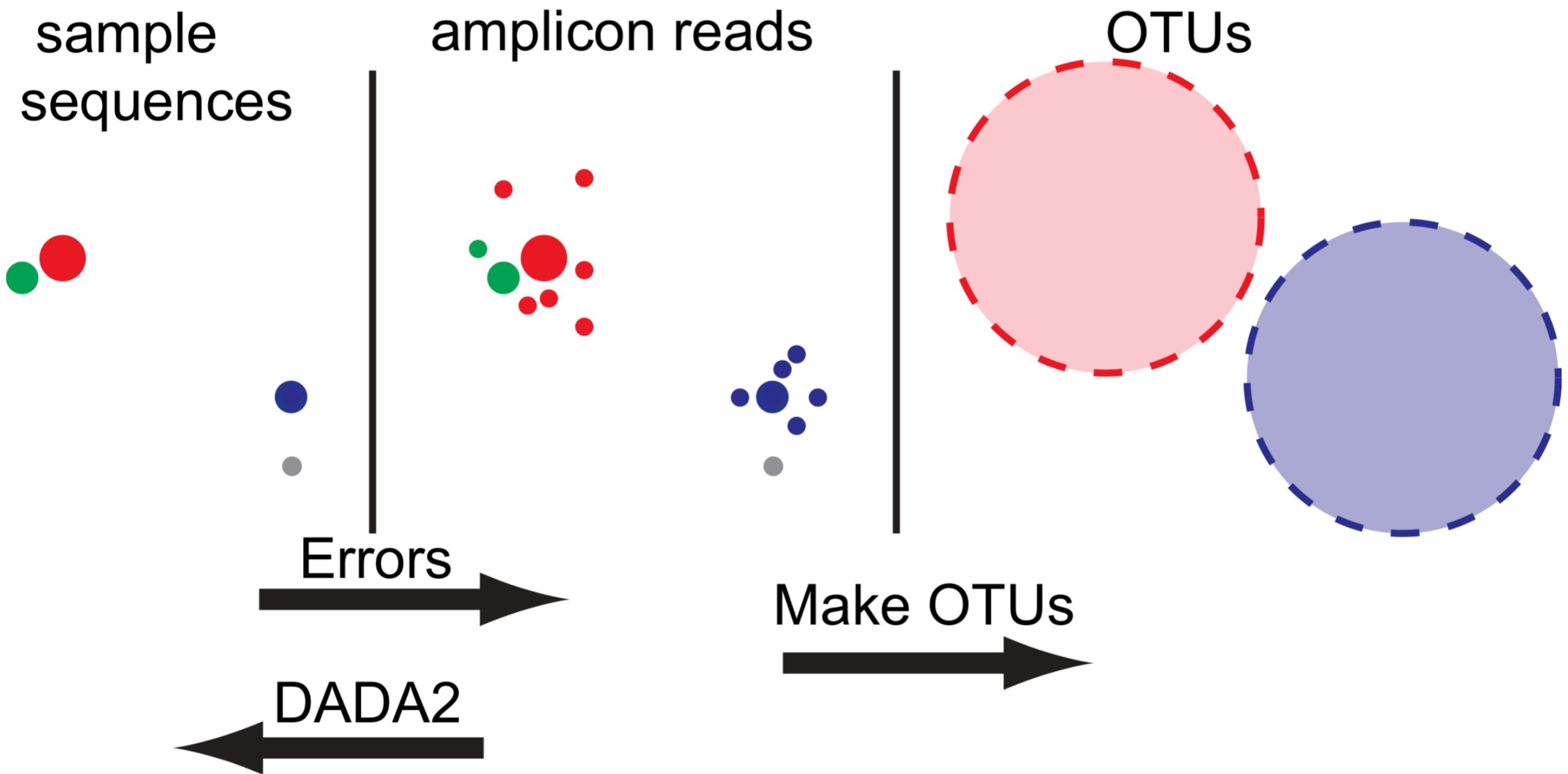
Usar OTUs es limitante

- Una OTU en un análisis no corresponde a las mismas secuencias ni taxonomía que en otro análisis (OTUs no son comparables).
- OTUs tienen baja resolución, normalmente hasta género.
- Sobreestiman el número de microorganismos en una muestra.
- Estima mal la abundancia de cada microorganismo.

¿Por que no usar las secuencias directamente?

- Al principio Illumina producía muchos errores. Ahora no.
- El problema es distinguir errores en secuenciación de variación biológica.

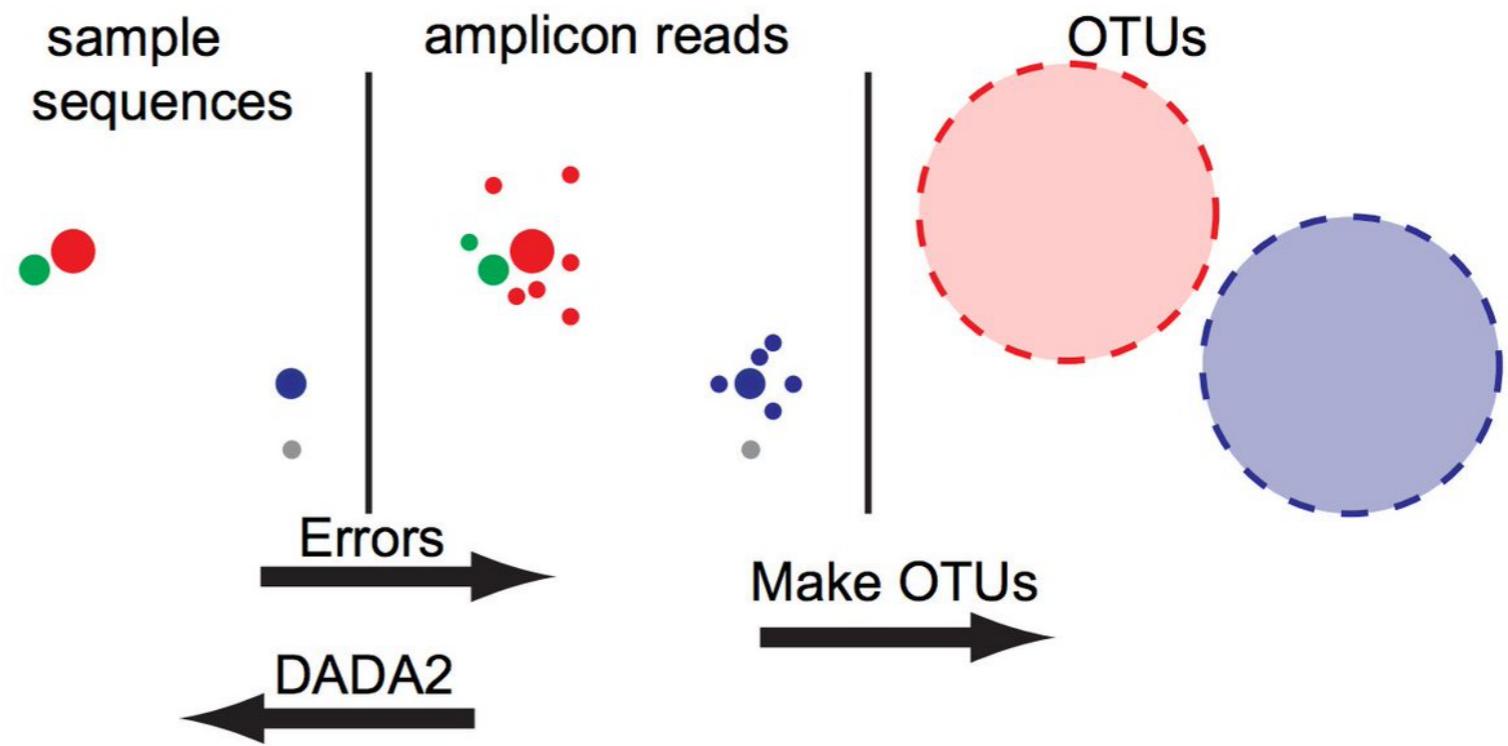
Pipeline DADA2



Resolución a nivel de nucleótido

Amplicon Sequence Variants (ASVs)

- Mayor resolución taxonómica.
- ASVs son consistentes.



Callahan, et al. Nature Methods, 2016.

Modelo de errores de DADA2

s: ATTAACGAGATTATAACCAGAGTACGAATA...
| |

r: AT**C**AACGAGATTATAAC**A**AGAGTACGAATA...

$$p(r|s) = \prod_{i=1}^L p(r(i)|s(i), q_r(i), Z)$$

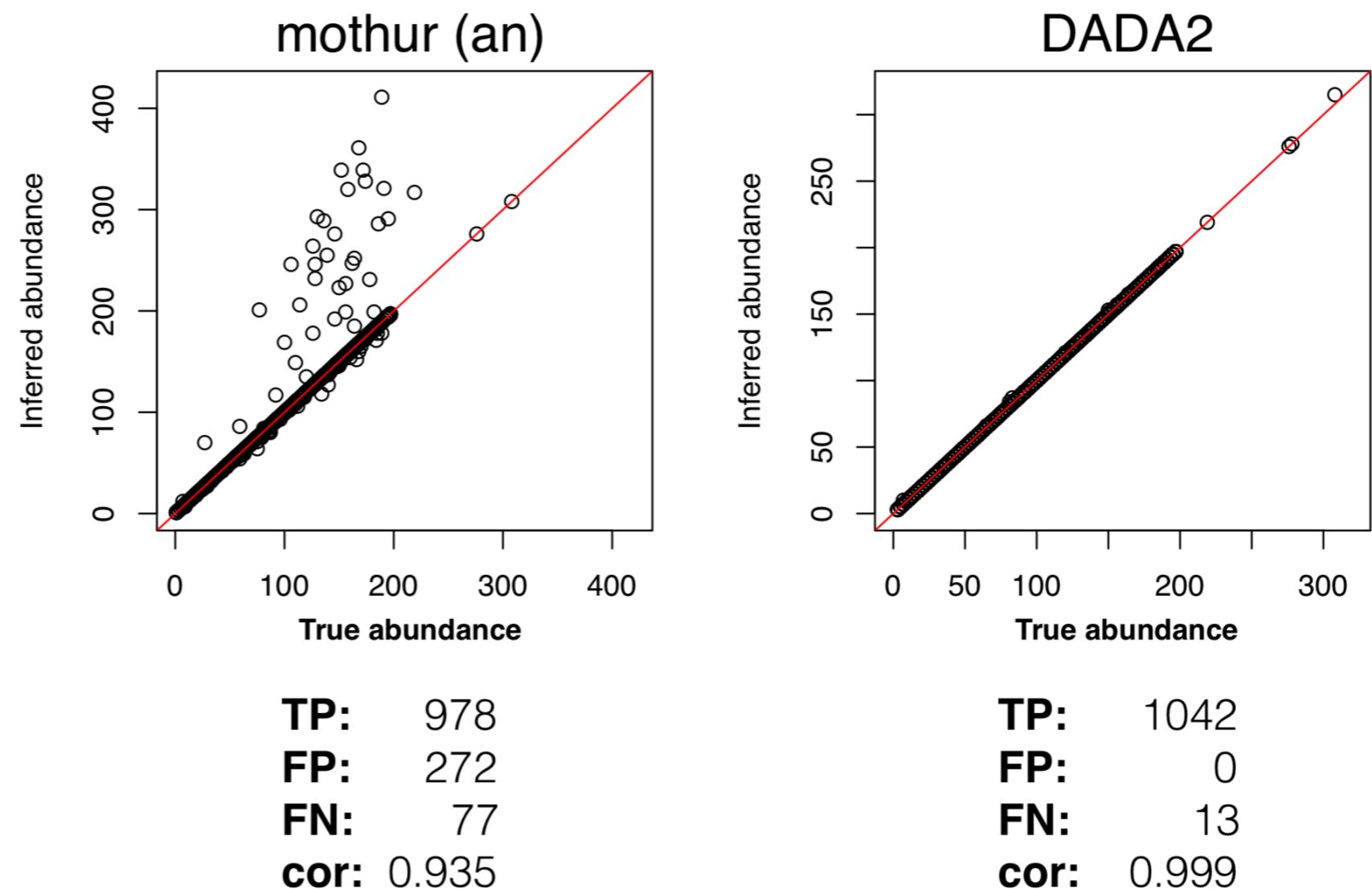
Error rates depend on....

- Substitution (eg. A->C)
- Quality score (eg. Q=30)
- Batch effect (eg. run)

Using more data!

Resolución y exactitud

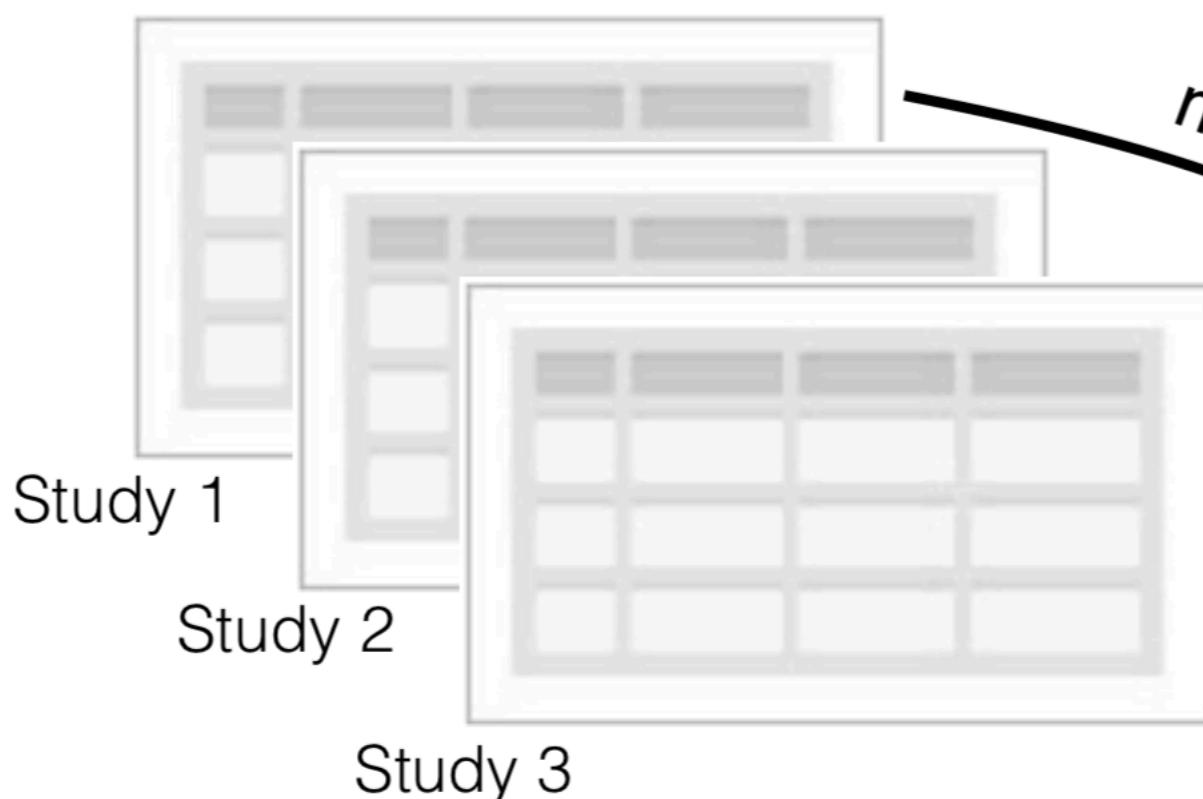
- Predicción de abundancia relativa en DADA2 (ASV) es más exacta que con mothur (OTUs).



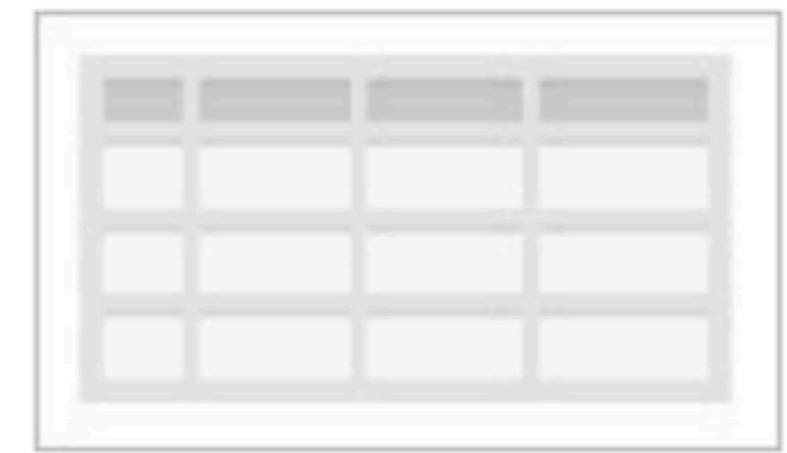
Data: Kopylova, et al. mSystems, 2016.

No hay necesidad de reprocesar los datos en conjunto

Sequence Tables

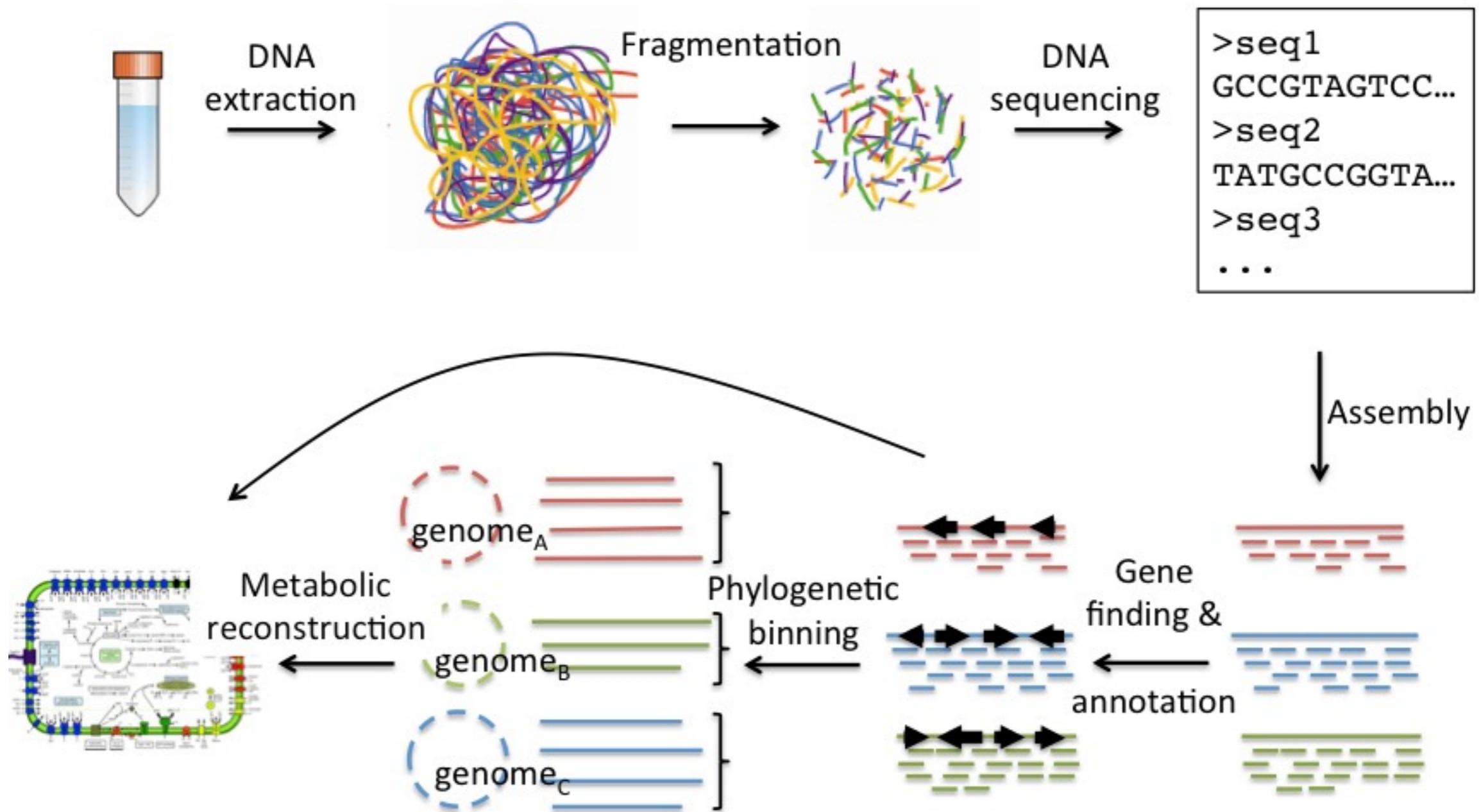


Cross-study comparison

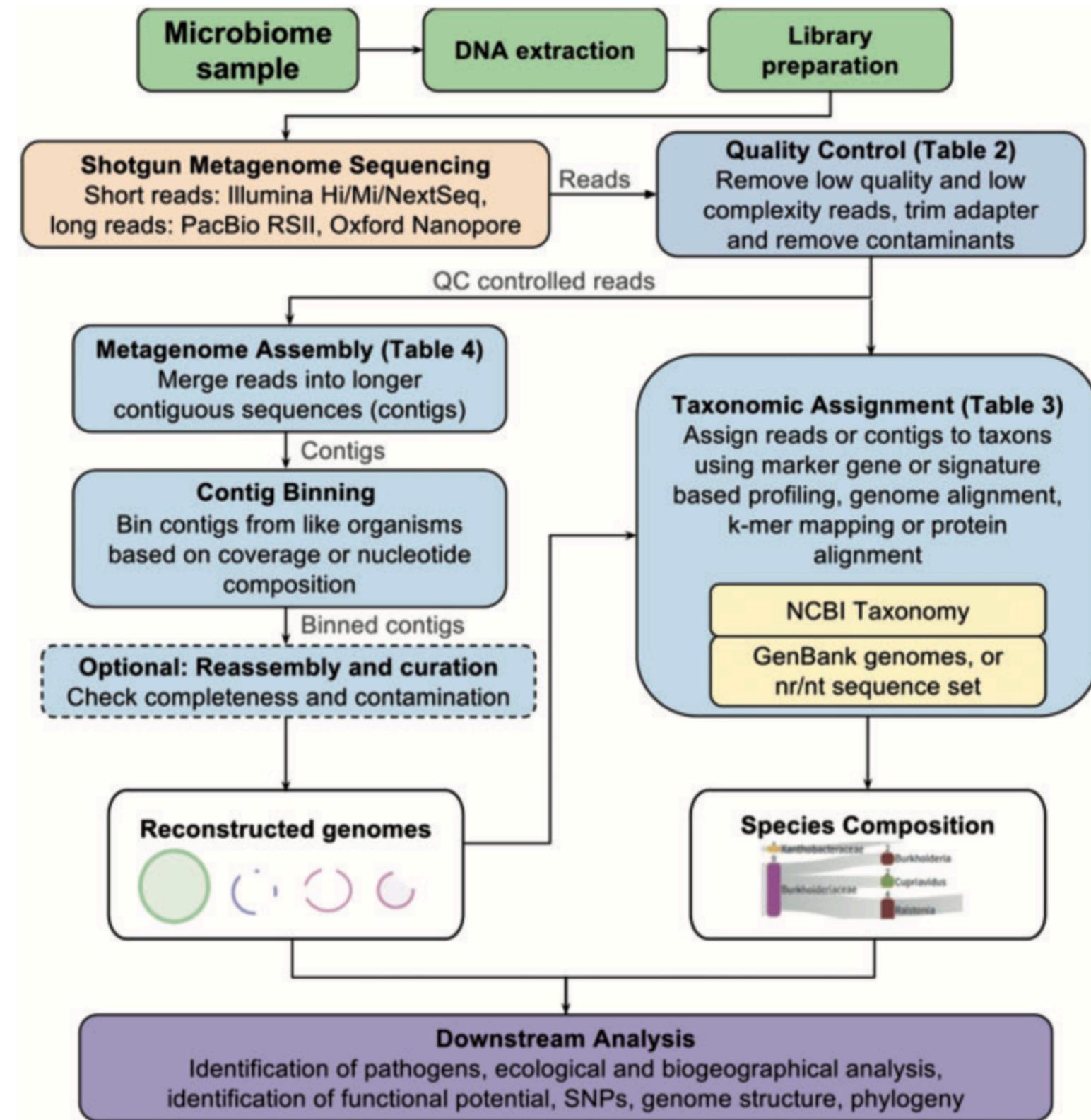


Metataxonómica ≠ Metagenómica
16S rRNA (amplicones)

Shotgun Sequencing



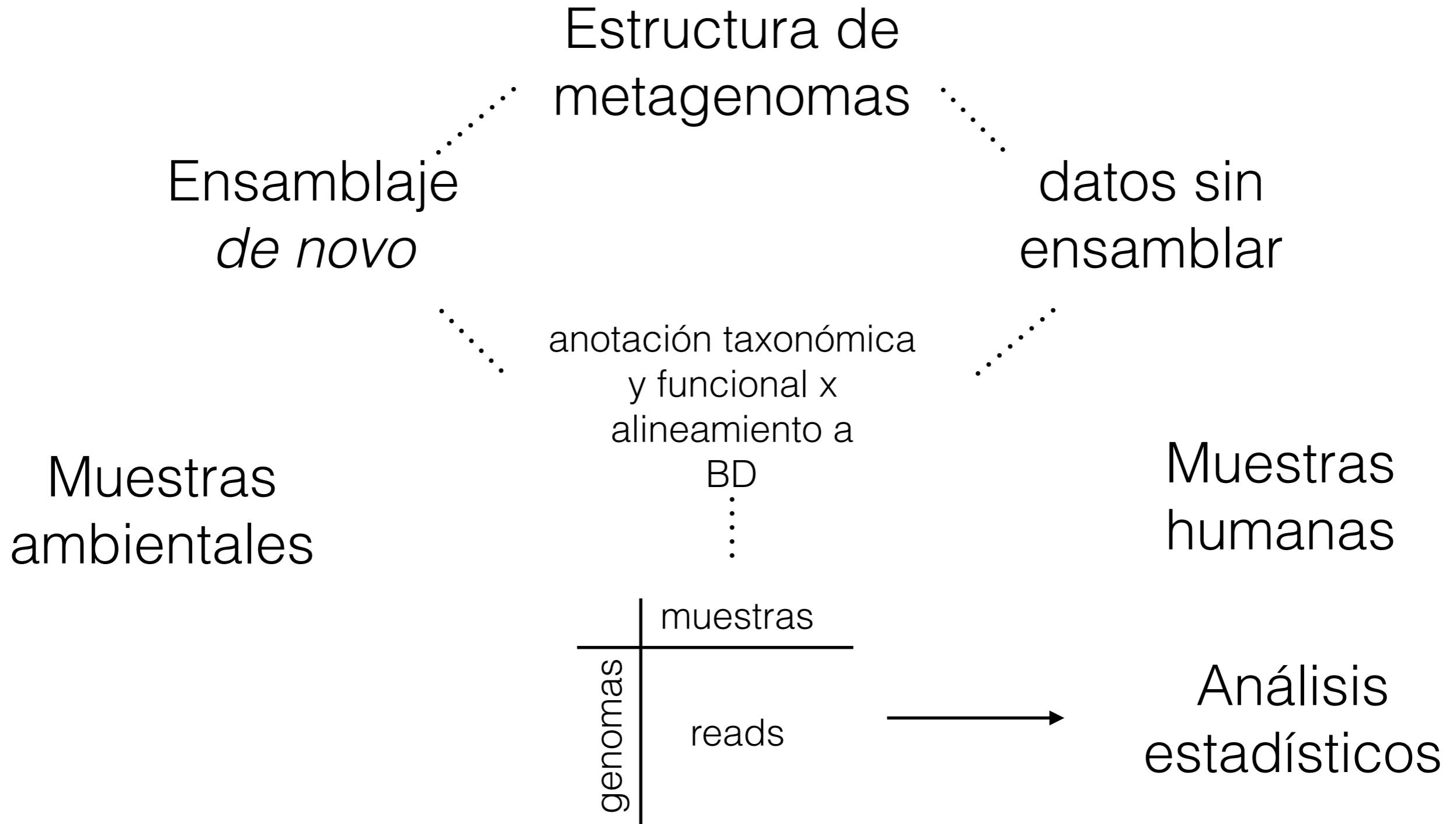
Análisis de datos metagenómicos



Dos estrategias de análisis

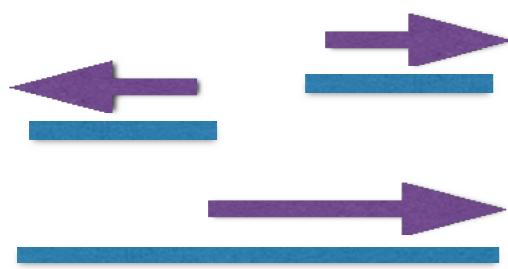
- Basado en ensamblaje *de novo* (*contig-based/assembly-based*).
- Basado en mapeo en contra de referencias (*read-based*).

Estrategias analíticas para datos de metagenomas



Two approaches in metagenomic studies

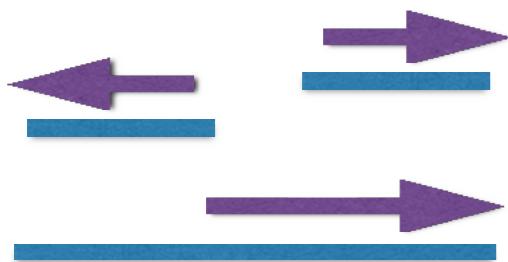
Read-based



- Taxonomic and functional community profile
- Allows quick comparative analysis
- Provides a broad picture of the community

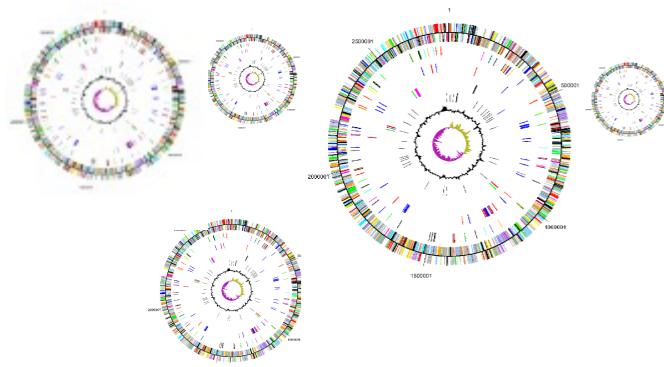
Two approaches in metagenomic studies

Read-based



- Taxonomic and functional community profile
- Allows quick comparative analysis
- Provides a broad picture of the community

Assembly-based



- Improved taxonomic classification
- Association between taxonomy and function
- Novel genes and novel taxa

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

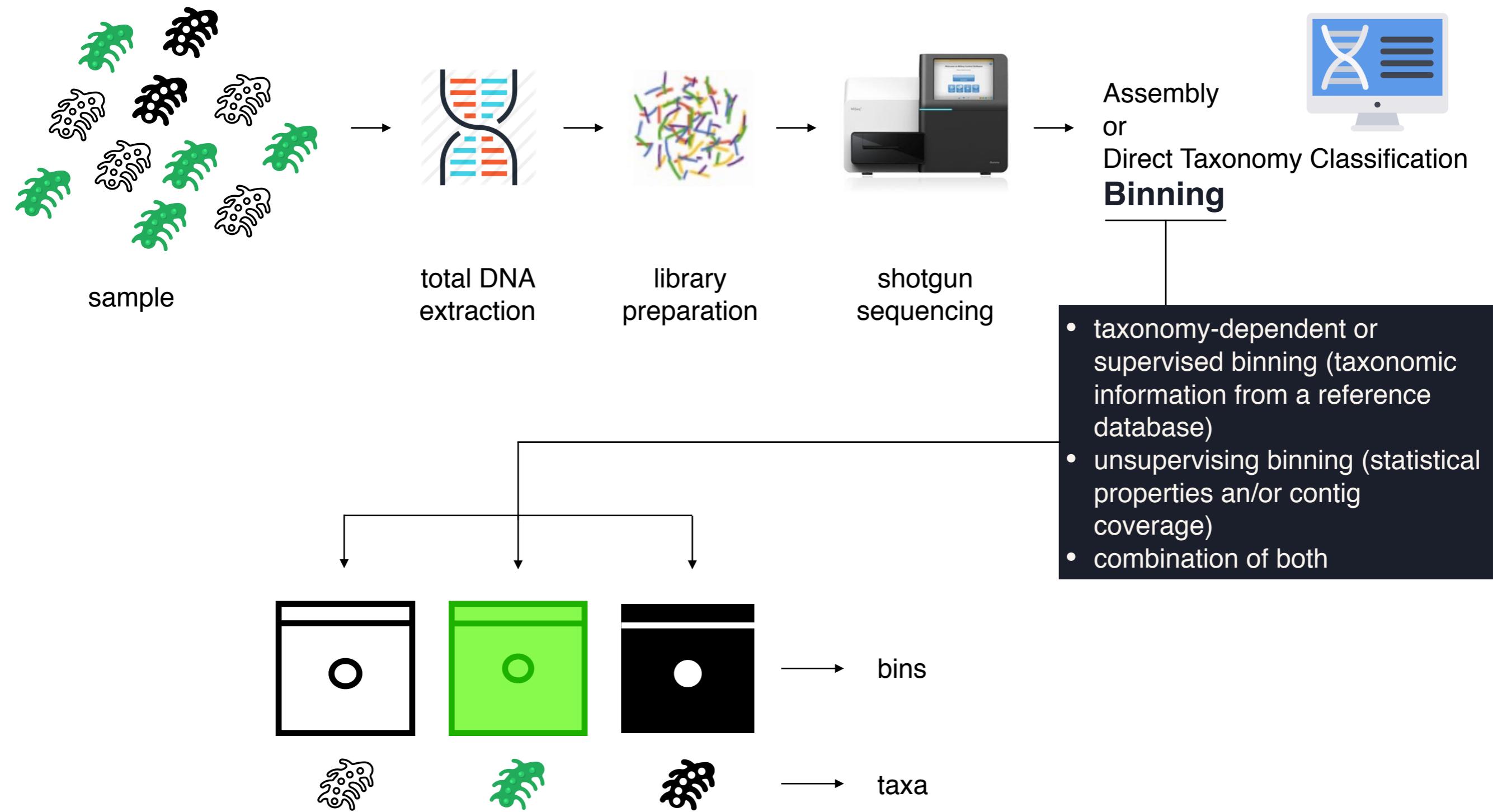


"Binning"



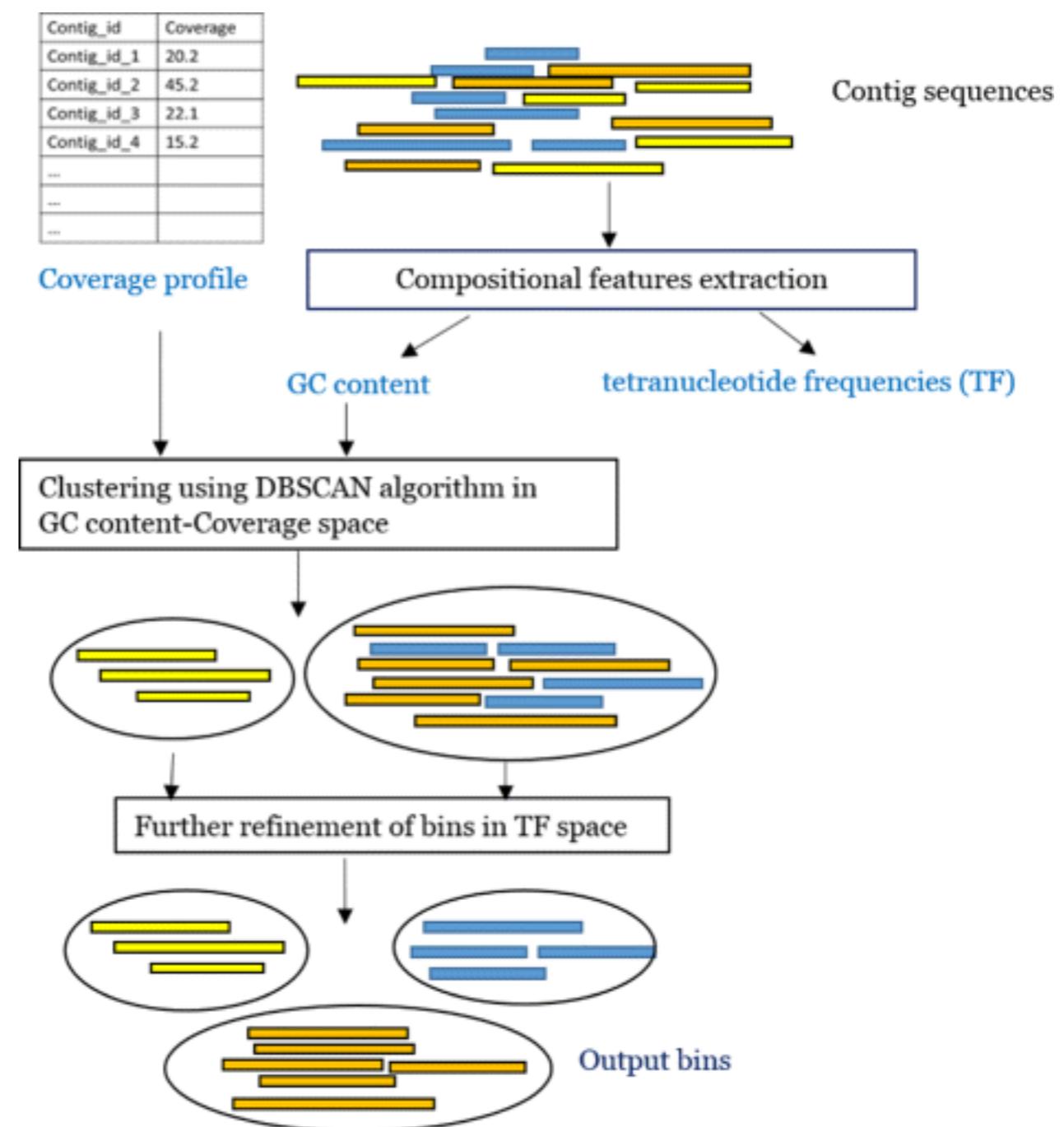
Complex sample

Binning



Binning

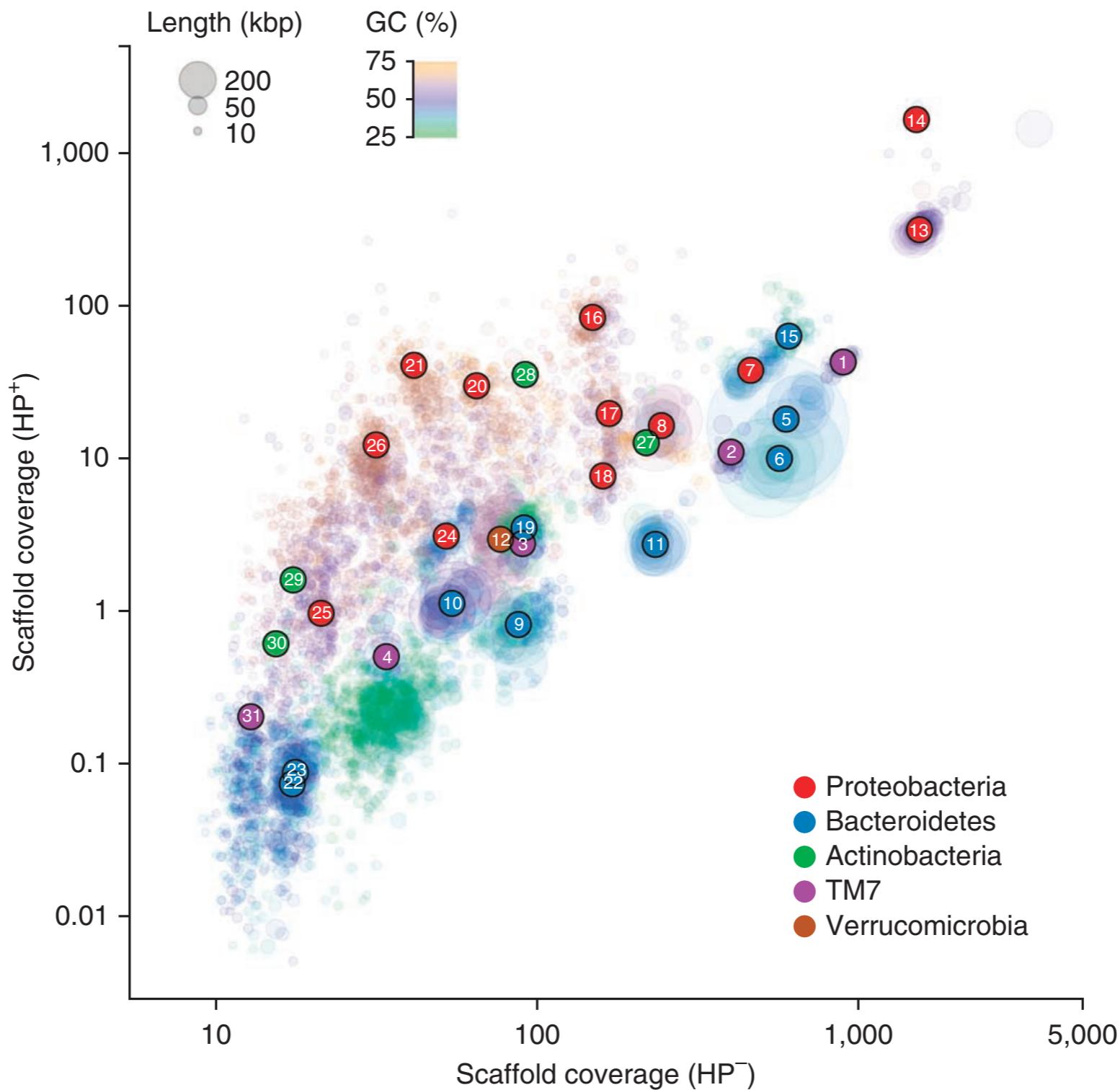
- Agrupar secuencias de acuerdo a características comunes.
- % GC, frecuencia de di-, tri-, o tetra-nucleótidos.
- Cobertura diferencial.



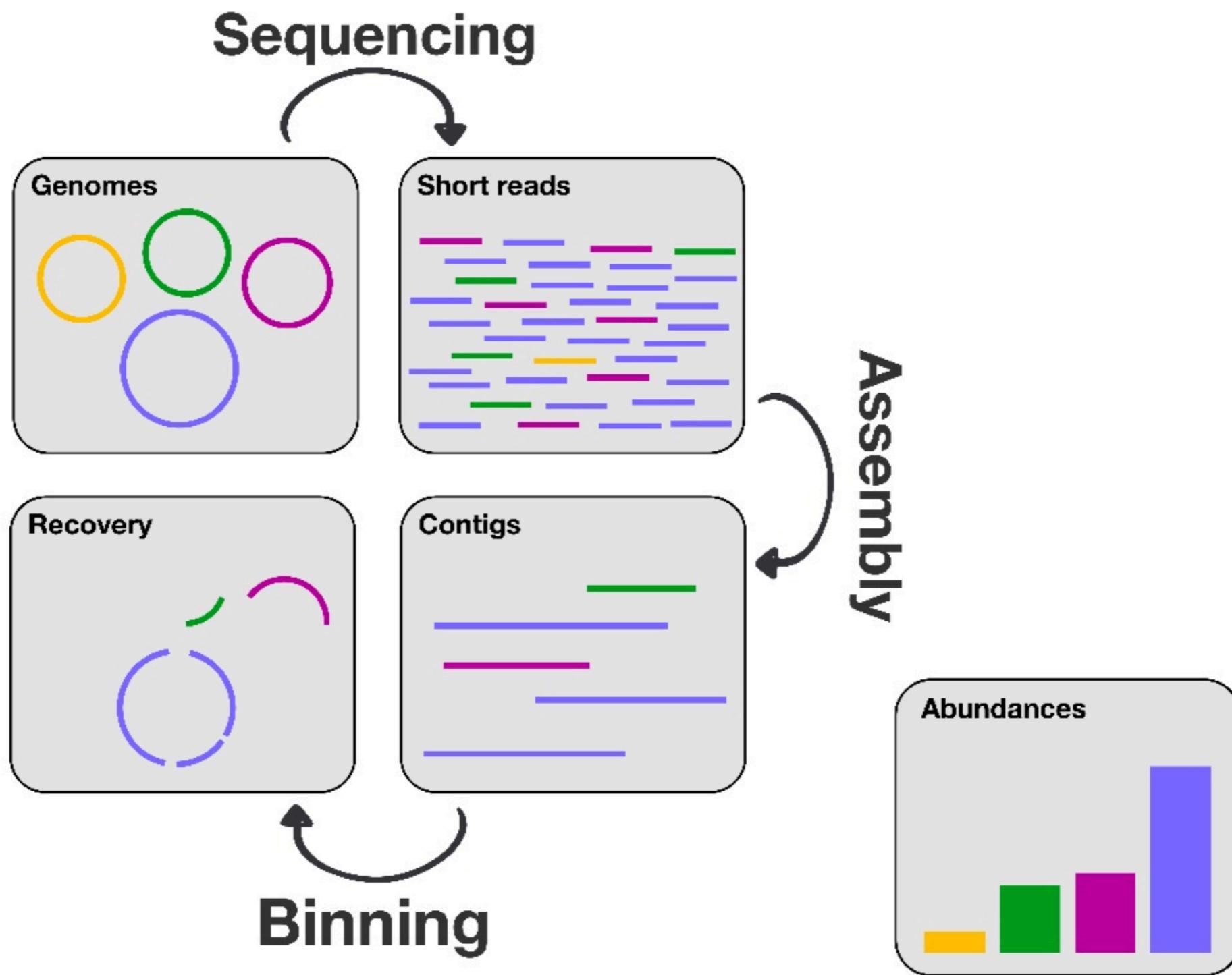
Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes

Mads Albertsen¹, Philip Hugenholtz^{2,3}, Adam Skarszewski², Kåre L Nielsen¹, Gene W Tyson^{2,4} & Per H Nielsen¹

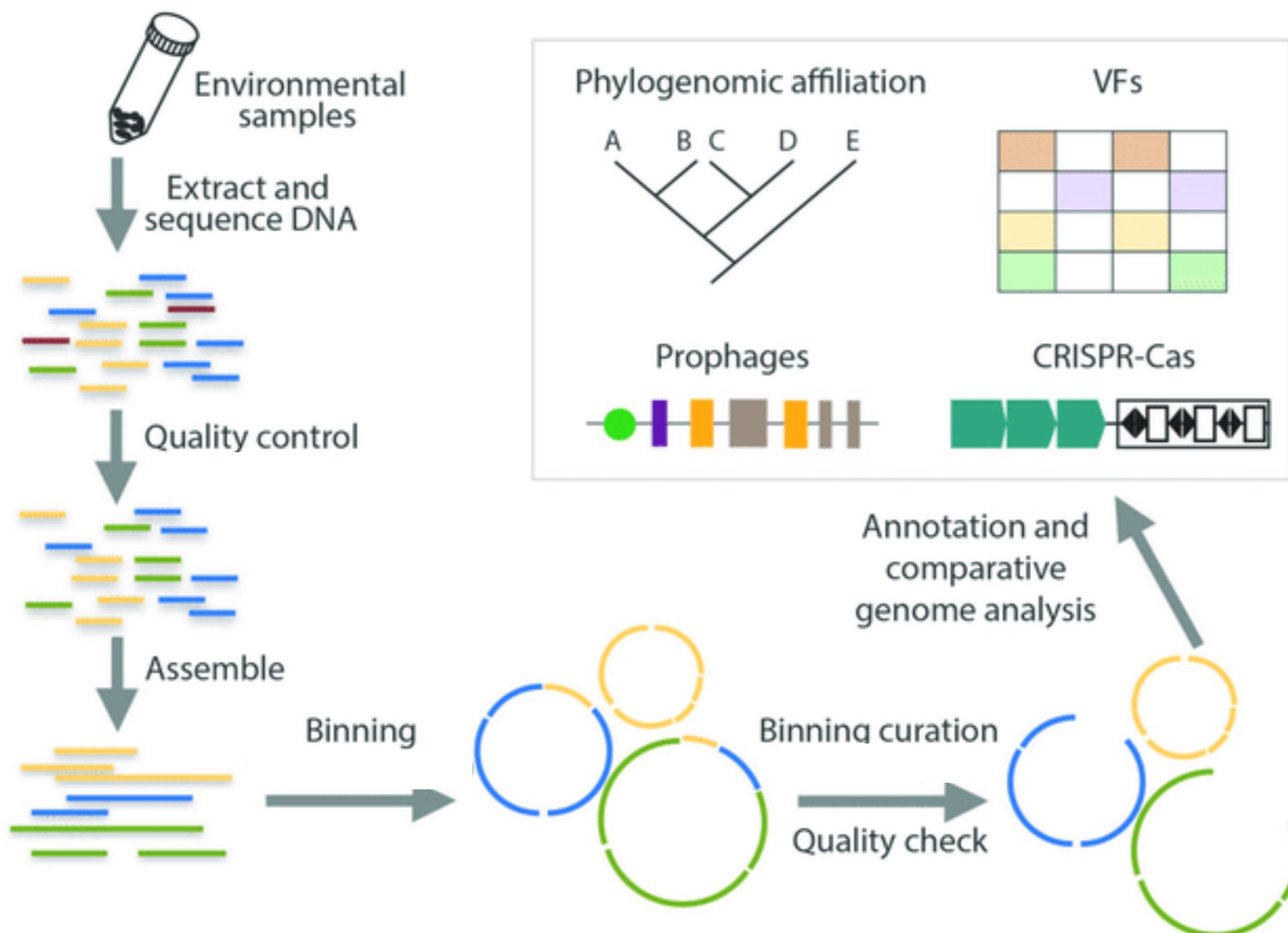
NATURE BIOTECHNOLOGY VOLUME 31 NUMBER 6 JUNE 2013



Genome-resolved metagenomics through binning



Genome-resolved metagenomics through binning



Metagenómica: microscopio moderno

