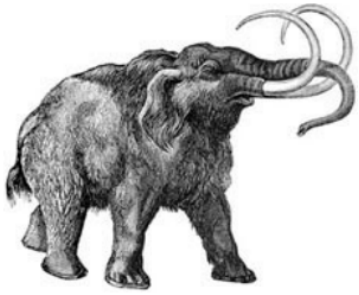
GENOMICA Y PROTEOMICA COMPUTACIONAL

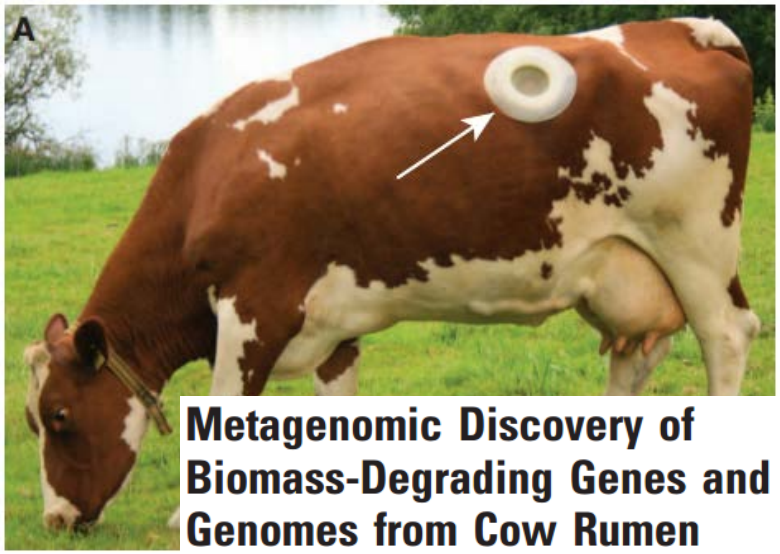
Práctica 11: Metagenómica (USEARCH)

Dr. Frank Solano

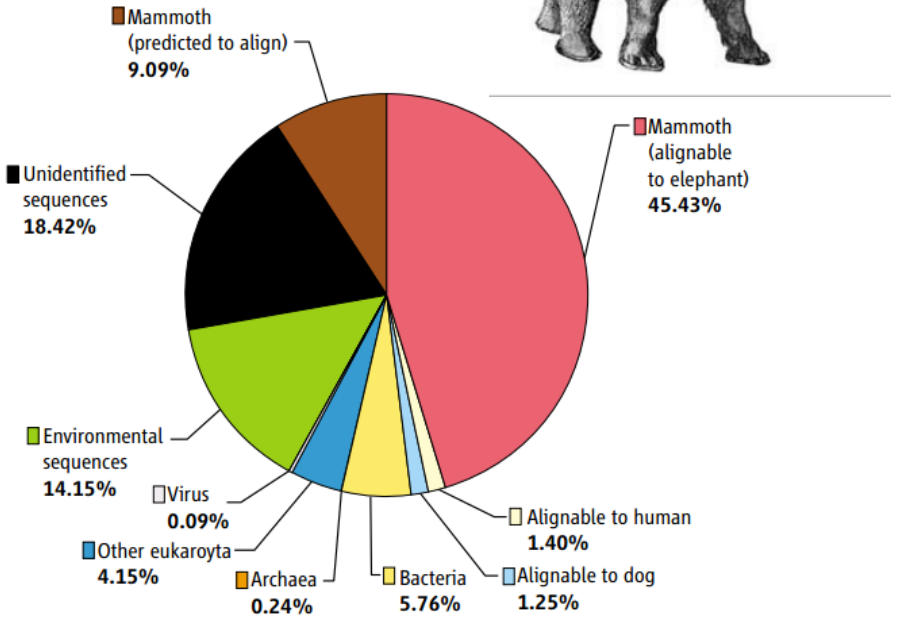
Metagenómica

- El estudio del ADN de los microrganismos no cultivados
- Genoma: la información genética de un solo organismo
- Metagenoma: la información genética de un conjunto de organismos
  - Suelos, aguas, costa, aire, huesos antiguos, muestras asociadas a un huésped, etc.





Genes para enzimas que degradan polisacáridos (biocombustibles)

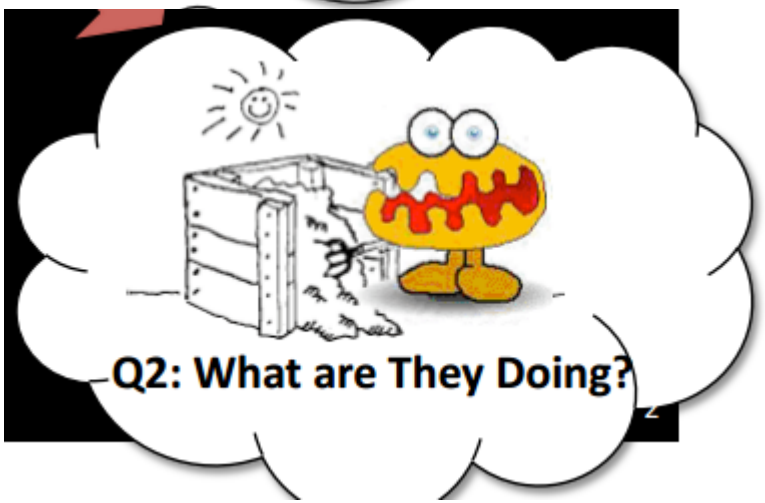


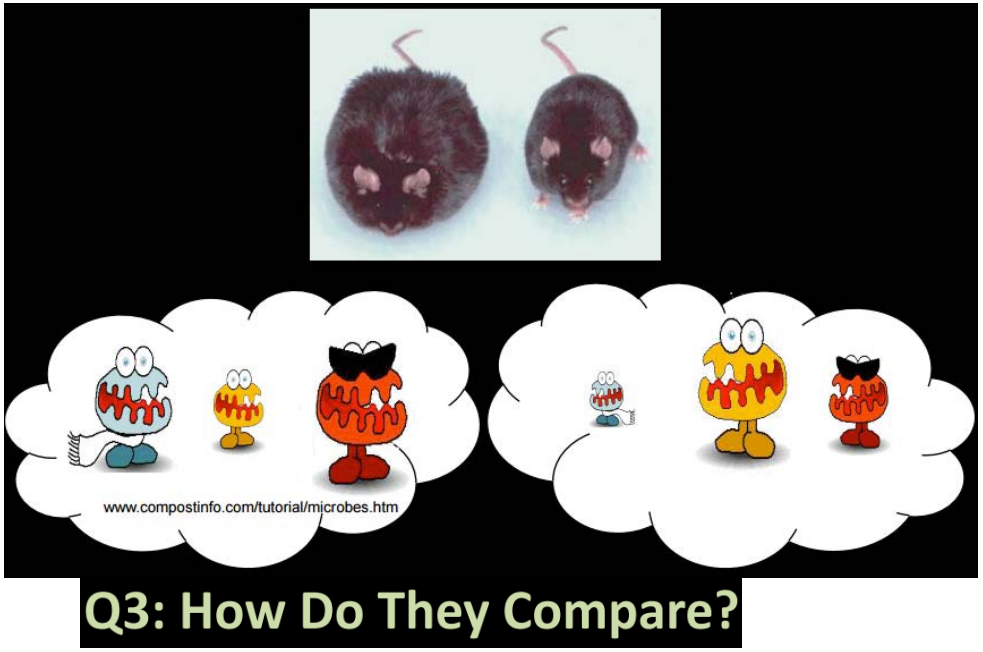
Metagenómica

- Objetivo: entender la composición y operación de un consorcio complejo de microrganismos en una muestra ambiental por medio de la secuenciación y análisis de su ADN
- Metatranscriptómica y metaproteómica:
  - El blanco es el ARN y las proteínas obtenidas de esas muestras.
- Avances en la NGS aumentan el número y el alcance de los proyectos de secuenciación ambiental

Metagenómica: preguntas bioinformáticas

- Quienes están ahí?
  - Contenido taxonómico
- Que están haciendo?
  - Funciones de genes
- Como se comparan?
  - Muestras diferentes





Metagenómica: preguntas bioinformáticas

- Quienes están ahí?
  - Contenido taxonómico
- Que están haciendo?
  - Funciones de genes
- Como se comparan?
  - Muestras diferentes

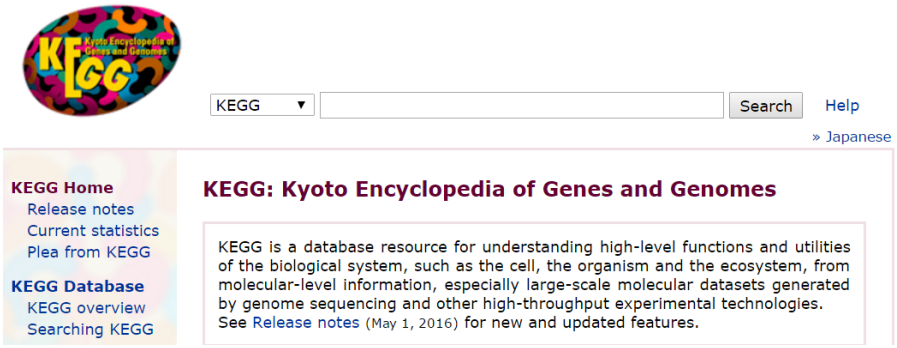
GREENGENES
The 16S rRNA Gene Database and Tools

Taxonomy
NCBI

rdp

silva
high quality ribosomal RNA databases

The Ribosomal Database Project



USEARCH
Ultra-fast sequence analysis

Metagenómica: preguntas bioinformáticas

- Quienes están ahí?
- Que están haciendo?
- Como se comparan?

Cold Spring Harbor Laboratory
bioRχiv
THE PREPRINT SERVER FOR BIOLOGY

New Results

SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences

Robert C. Edgar
doi: https://doi.org/10.1101/074161

Cold Spring Harbor Laboratory
bioRχiv
THE PREPRINT SERVER FOR BIOLOGY

New Results

UCHIME2: improved chimera prediction for amplicon sequencing

Robert C. Edgar
doi: https://doi.org/10.1101/074252

Search and clustering orders of magnitude faster than BLAST FREE

Robert C. Edgar Author Notes

Bioinformatics, Volume 26, Issue 19, 1 October 2010, Pages 2460–2461,
https://doi.org/10.1093/bioinformatics/btq461

Brief Communication | Published: 18 August 2013

UPARSE: highly accurate OTU sequences from microbial amplicon reads

Robert C Edgar

Nature Methods **10**, 996–998 (2013) | Download Citation

Error filtering, pair assembly and error correction for next-generation sequencing reads FREE

Robert C. Edgar, Henrik Flyvbjerg Author Notes

Bioinformatics, Volume 31, Issue 21, 1 November 2015, Pages 3476–3482,
https://doi.org/10.1093/bioinformatics/btv401

Cold Spring Harbor Laboratory
bioRχiv
THE PREPRINT SERVER FOR BIOLOGY

New Results Commen

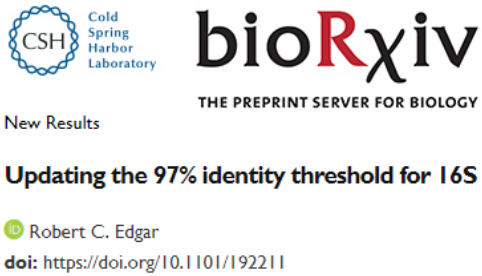
UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing

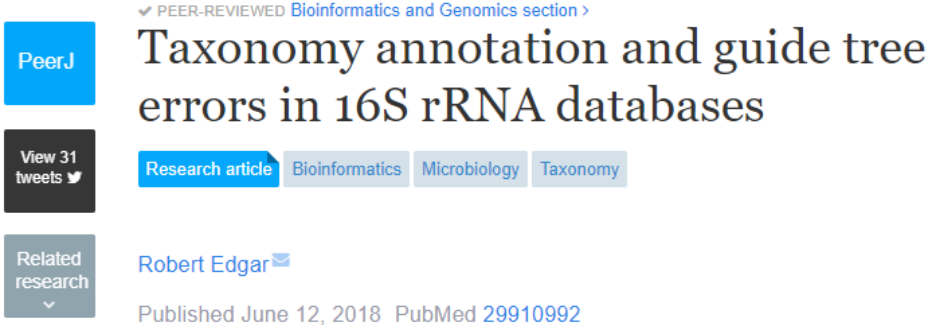
Robert C. Edgar
doi: https://doi.org/10.1101/081257

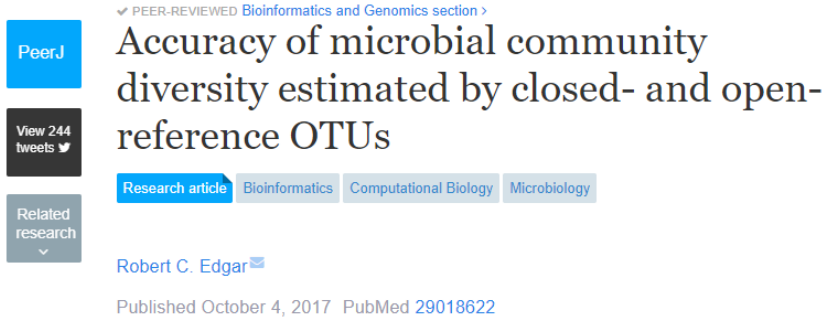
USEARCH
Ultra-fast sequence analysis

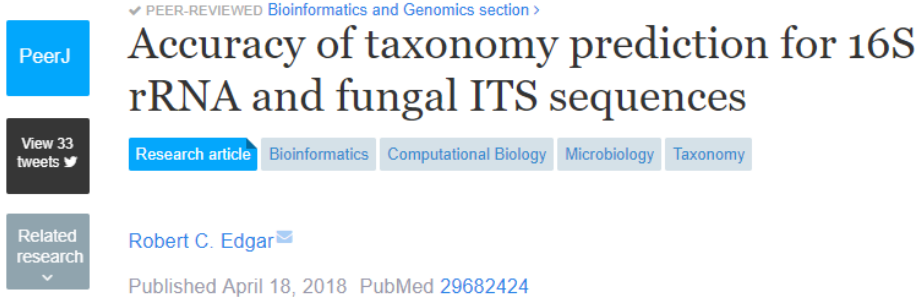
Metagenómica: preguntas bioinformáticas

- Quienes están ahí?
- Que están haciendo?
- Como se comparan?









Metagenómica: preguntas bioinformáticas

64bit

PeerJ

VSEARCH: a versatile open source tool for metagenomics

Torbjørn Rognes[1,2], Tomáš Flouri[3,4], Ben Nichols[5], Christopher Quince[5,6] and Frédéric Mahé[7,8]

[1] Department of Informatics, University of Oslo, Oslo, Norway
[2] Department of Microbiology, Oslo University Hospital, Oslo, Norway
[3] Heidelberg Institute for Theoretical Studies, Heidelberg, Germany
[4] Institute for Theoretical Informatics, Karlsruhe Institute of Technology, Karlsruhe, Germany
[5] School of Engineering, University of Glasgow, Glasgow, United Kingdom
[6] Warwick Medical School, University of Warwick, Coventry, United Kingdom
[7] Department of Ecology, University of Kaiserslautern, Kaiserslautern, Germany
[8] UMR LSTM, CIRAD, Montpellier, France

ABSTRACT

Background. VSEARCH is an open source and free of charge multithreaded 64-bit tool for processing and preparing metagenomics, genomics and population genomics nucleotide sequence data. It is designed as an alternative to the widely used USEARCH tool (*Edgar, 2010*) for which the source code is not publicly available, algorithm details are only rudimentarily described, and only a memory-confined 32-bit version is freely available for academic use.

Metagenómica: preguntas bioinformáticas

Bioinformatics, 2018, 1–3
doi: 10.1093/bioinformatics/bty071
Advance Access Publication Date: 14 February 2018
Applications Note

OXFORD

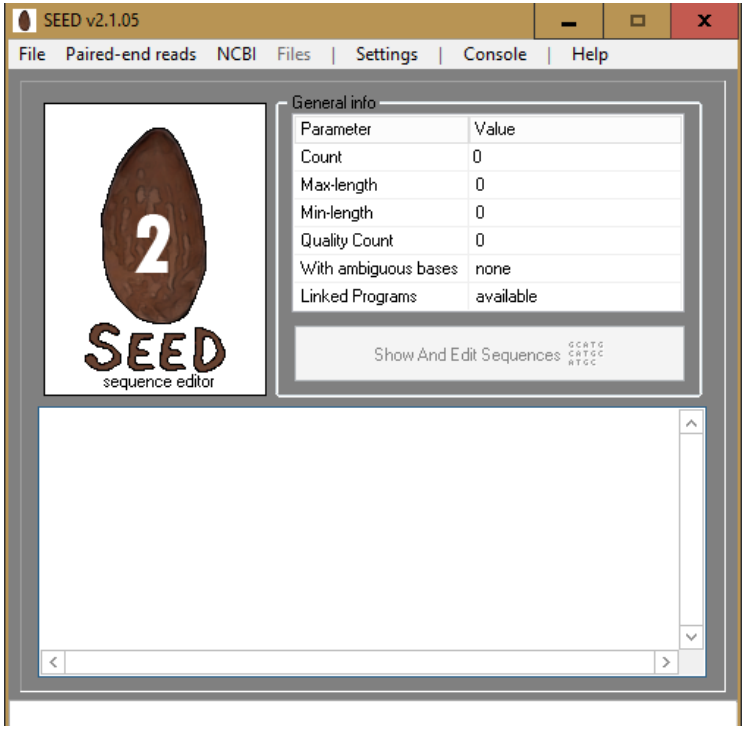
Sequence analysis

SEED 2: a user-friendly platform for amplicon high-throughput sequencing data analyses

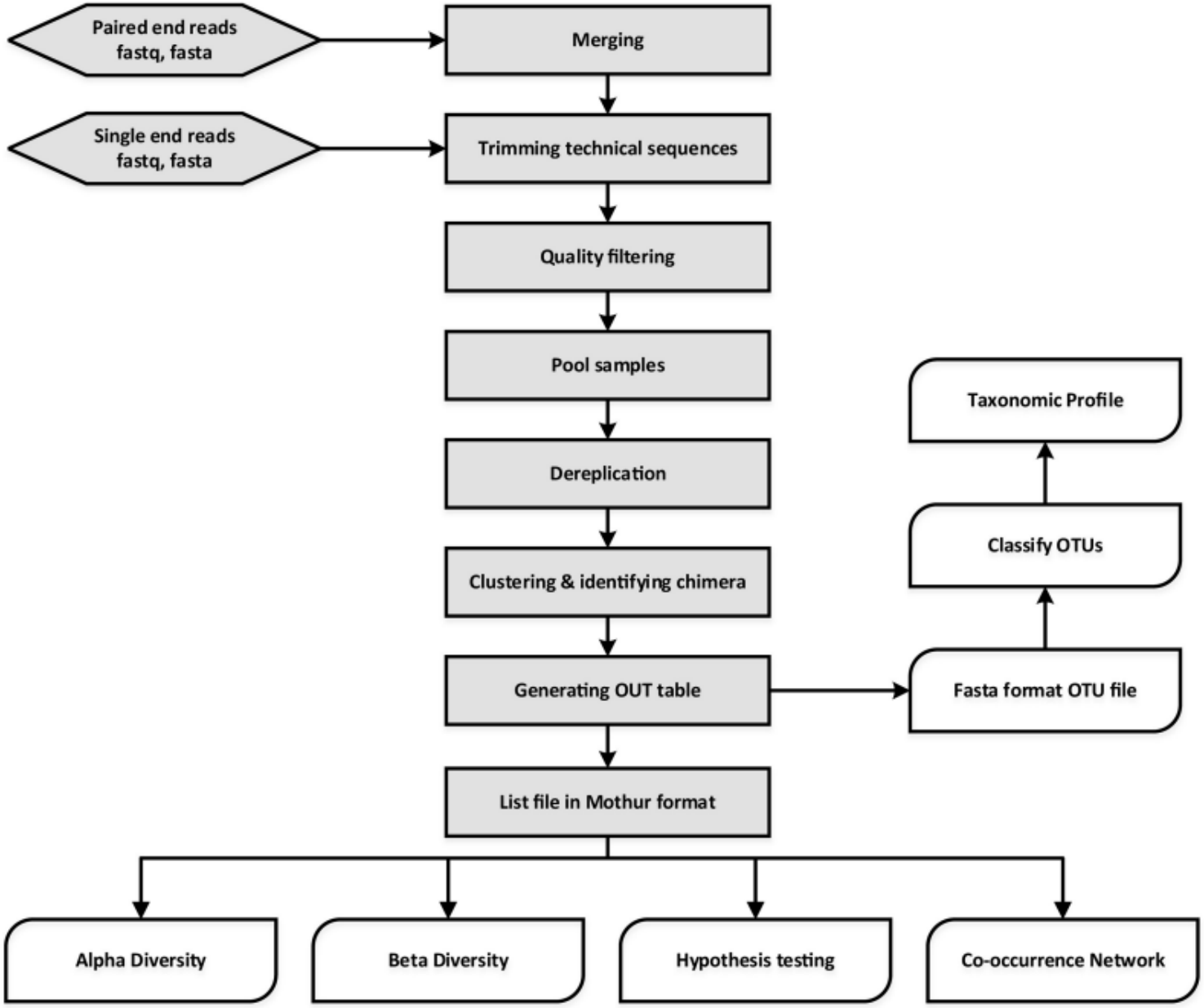
Tomáš Větrovský, Petr Baldrian and Daniel Morais*

Institute of Microbiology of the CAS, 14220 Prague 4, Czech Republic

64bit



Metagenómica: flujo de trabajo

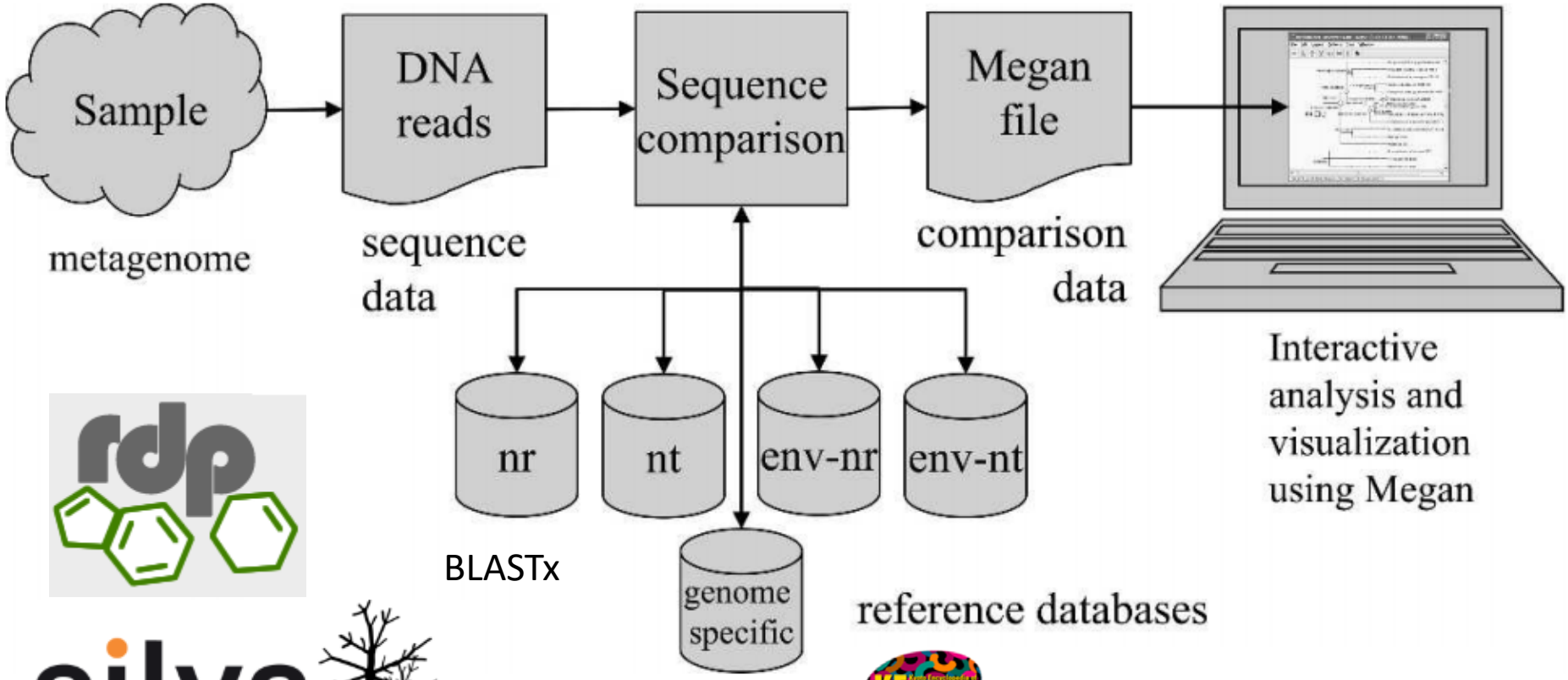


Metagenómica

Taxonomy

The SEED
Home of the SEED.

COGs
Phylogenetic classification of proteins encoded in complete genomes



rdp

silva
high quality ribosomal RNA databases

GREENGENES
The 16S rRNA Gene Database and Tools

KEGG Search Help
» Japanese

KEGG Home
Release notes
Current statistics
Plea from KEGG
KEGG Database
KEGG overview
Searching KEGG

KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies.
See Release notes (May 1, 2016) for new and updated features.

Metagenómica

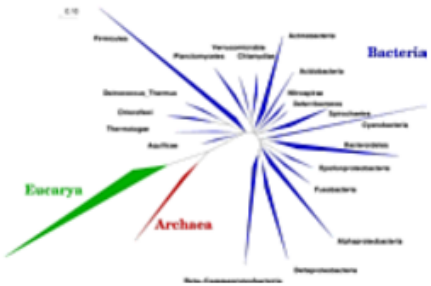
- Bases de datos

silva
high quality ribosomal RNA databases

Search Aligner Download Documentation Projects FISH & Probes Contact

SSU Ref NR (non-redundant)

Why a non-redundant version of the SILVA SSU Ref database?

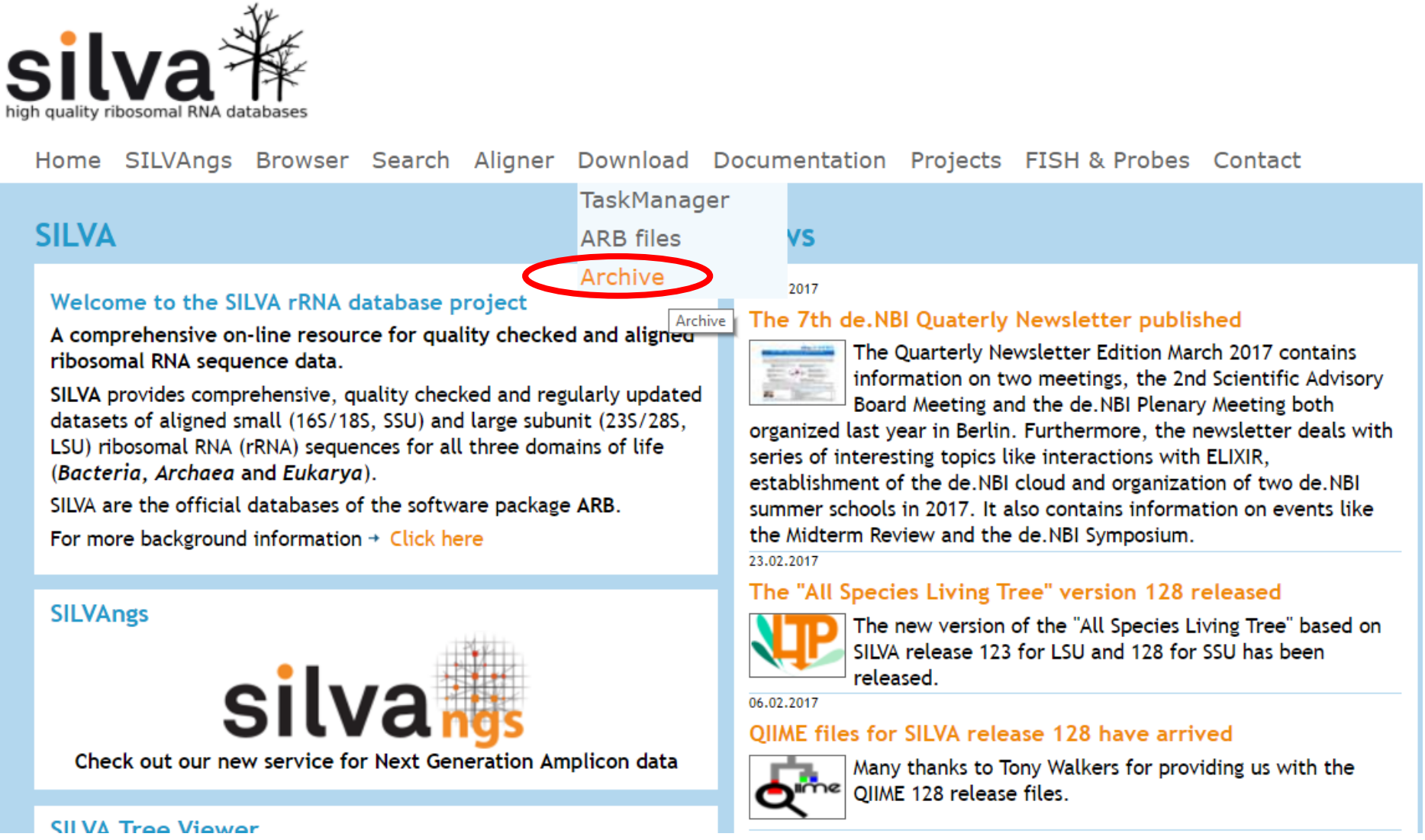


For users interested in representative (rRNA) sequence collections, the rapid growth of the data sets has led to immense hardware requirements paired with a significantly increased amount of time to analyse the data. In case of large databases such as the current rRNA data sets, ARB especially requires large amounts of main memory (RAM) to be able to load the database (see also box below).
Basically, there are two options to face this problem: (1) a hardware upgrade to provide the amount of RAM required or (2) a reduction of the number of sequences in the ARB database to bring the RAM requirements down to the current hardware specifications.
For multiple reasons, the second option should be prefered as long as the resulting data set still is "representative" - a very important parameter in environmental microbiology. Therefore, the SILVA project has addressed this task, resulting in a **"non-redundant" (NR) SSU Ref dataset** build by a dereplication of the full SSU Ref using a 99% identity criterion.
As of SILVA release 119 the SSU Ref NR is the only SSU dataset with a manual curated guide tree. SSU Ref is still provided as an ARB dataset but without the guide tree.

Metagenómica

- Bases de datos

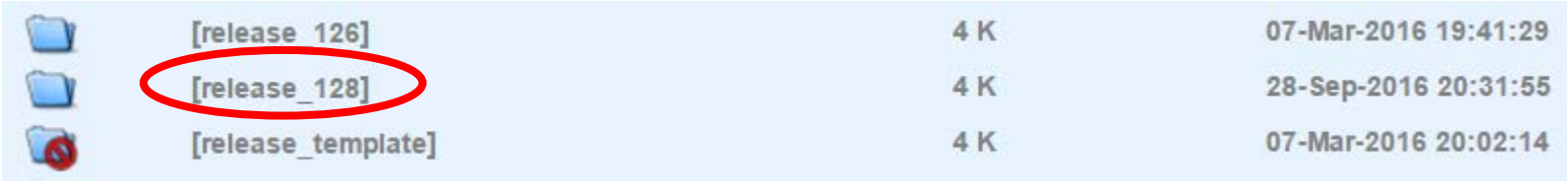
https://www.arb-silva.de/



Metagenómica

https://www.arb-silva.de/

- Bases de datos



	[release_126]	4 K	07-Mar-2016 19:41:29
	[release_128]	4 K	28-Sep-2016 20:31:55
	[release_template]	4 K	07-Mar-2016 20:02:14

File Repository

Listing fileadmin/silva_databases/release_128/

	Filename	Filesize	Last modification
	[..]	4 K	13-Mar-2017 09:45:58
	[ARB_files]	4 K	30-Sep-2016 17:40:30
	[Exports]	4 K	28-Sep-2016 20:32:41
	[Fields_description]	4 K	21-Sep-2016 16:01:20
	[user]	4 K	29-Feb-2016 22:58:20

Metagenómica

- Bases de datos

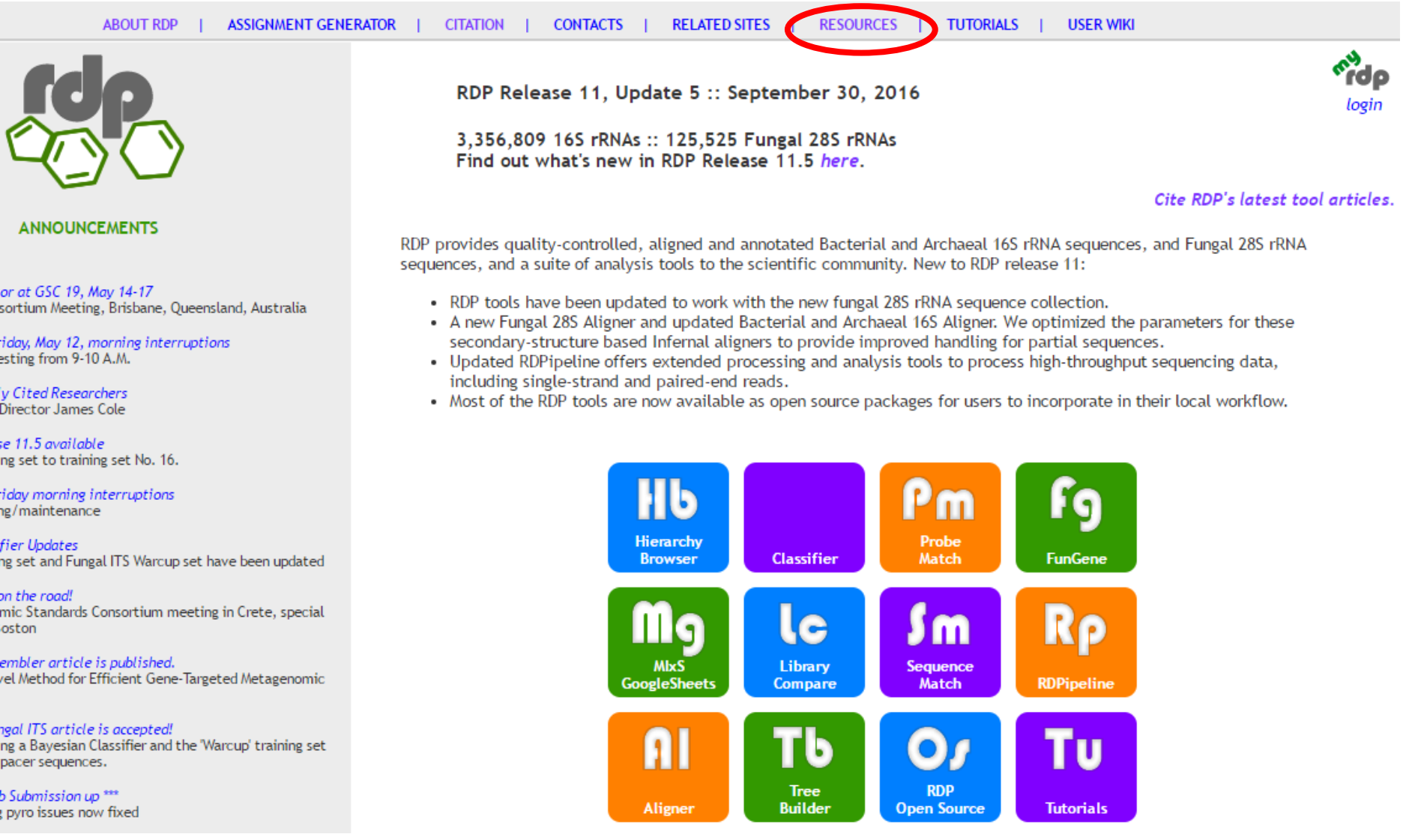
https://www.arb-silva.de/

SILVA_128_SSUParc_tax_silva.fasta.gz	796 M	28-Sep-2016 20:18:38
SILVA_128_SSUParc_tax_silva.fasta.gz.md5	71 B	28-Sep-2016 20:18:38
SILVA_128_SSUParc_tax_silva_trunc.fasta.gz	792 M	28-Sep-2016 20:18:48
SILVA_128_SSUParc_tax_silva_trunc.fasta.gz.md5	77 B	28-Sep-2016 20:18:48
SILVA_128_SSURef_Nr99_tax_silva.fasta.gz	150 M	28-Sep-2016 20:19:04
SILVA_128_SSURef_Nr99_tax_silva.fasta.gz.md5	75 B	28-Sep-2016 20:19:04
SILVA_128_SSURef_Nr99_tax_silva_full_align_trunc.fasta.gz	2 G	28-Sep-2016 20:19:20
SILVA_128_SSURef_Nr99_tax_silva_full_align_trunc.fasta.gz.md5	92 B	28-Sep-2016 20:19:20
SILVA_128_SSURef_Nr99_tax_silva_trunc.fasta.gz	150 M	28-Sep-2016 20:19:23
SILVA_128_SSURef_Nr99_tax_silva_trunc.fasta.gz.md5	81 B	28-Sep-2016 20:19:23
SILVA_128_SSURef_tax_silva.fasta.gz	360 M	28-Sep-2016 20:19:28
SILVA_128_SSURef_tax_silva.fasta.gz.md5	70 B	28-Sep-2016 20:19:28
SILVA_128_SSURef_tax_silva_full_align_trunc.fasta.gz	4 G	28-Sep-2016 20:20:27
SILVA_128_SSURef_tax_silva_full_align_trunc.fasta.gz.md5	87 B	28-Sep-2016 20:20:27
SILVA_128_SSURef_tax_silva_trunc.fasta.gz	358 M	28-Sep-2016 20:20:32
SILVA_128_SSURef_tax_silva_trunc.fasta.gz.md5	76 B	28-Sep-2016 20:20:32

Metagenómica

- Bases de datos

https://rdp.cme.msu.edu/index.jsp



Metagenómica

- Bases de datos

https://rdp.cme.msu.edu/index.jsp

Release Alignment:

These files reflect the contents of the current RDP Release data. Please see the release notes for more information.

Gene	Aligned FASTA	Unaligned	Coverage Chart
Bacteria 16S	Aligned	Fasta, Genbank	Excel
Archaea 16S	Aligned	Fasta, Genbank	Excel
Fungal 28S	Aligned	Fasta, Genbank	Excel

Details of sequence quality check by RDP is available.

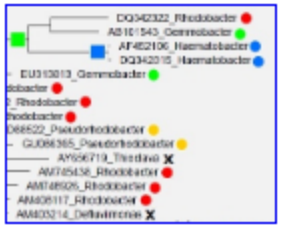
Metagenómica

- Bases de datos: cual usar?

Use a small database with authoritative classifications

I recommend using a authoritatively classified sequences, e.g. for 16S the most recent RDP training set or LTP release.

Taxonomy annotations in large databases are unreliable predictions



The taxonomy annotations in the large 16S databases (SILVA, Greengenes, or the full RDP database) are mostly computational predictions from 16S sequences. Roughly one in five of these predictions are wrong, probably because the guide trees have pervasive branching order errors. Therefore, using annotations from large databases adds a substantial error rate in the reference dataset on top of the intrinsic error rate of a prediction algorithm such as SINTAX or the Naive Bayesian Classifier. With these considerations in mind, I believe it is best to use a database of type strain and isolate sequences rather than Greengenes, SILVA or RDP.

https://www.drive5.com/usearch/manual/faq_tax_db.html

Metagenómica

- Datos

SCIENTIFIC REPORTS

www.nature.com/scientificreports

OPEN

Variation in koala microbiomes within and between individuals: effect of body region and captivity status

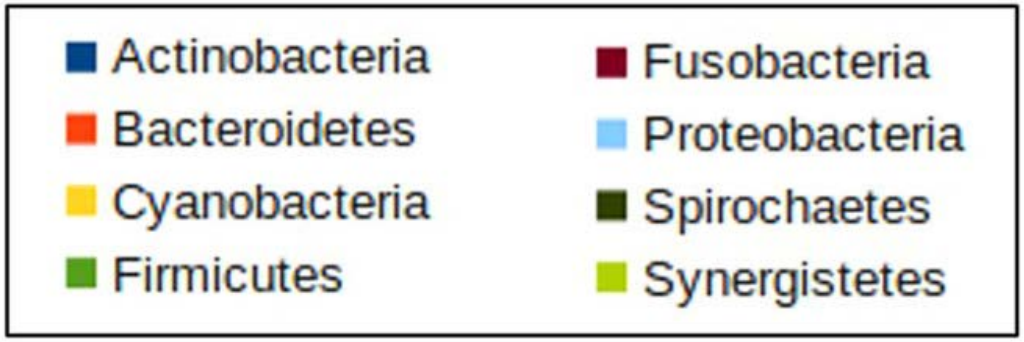
SCIENTIFIC REPORTS | 5:10189 | DOI: 10.1038/srep10189

Received: 02 December 2014
Accepted: 25 March 2015
Published: 11 May 2015

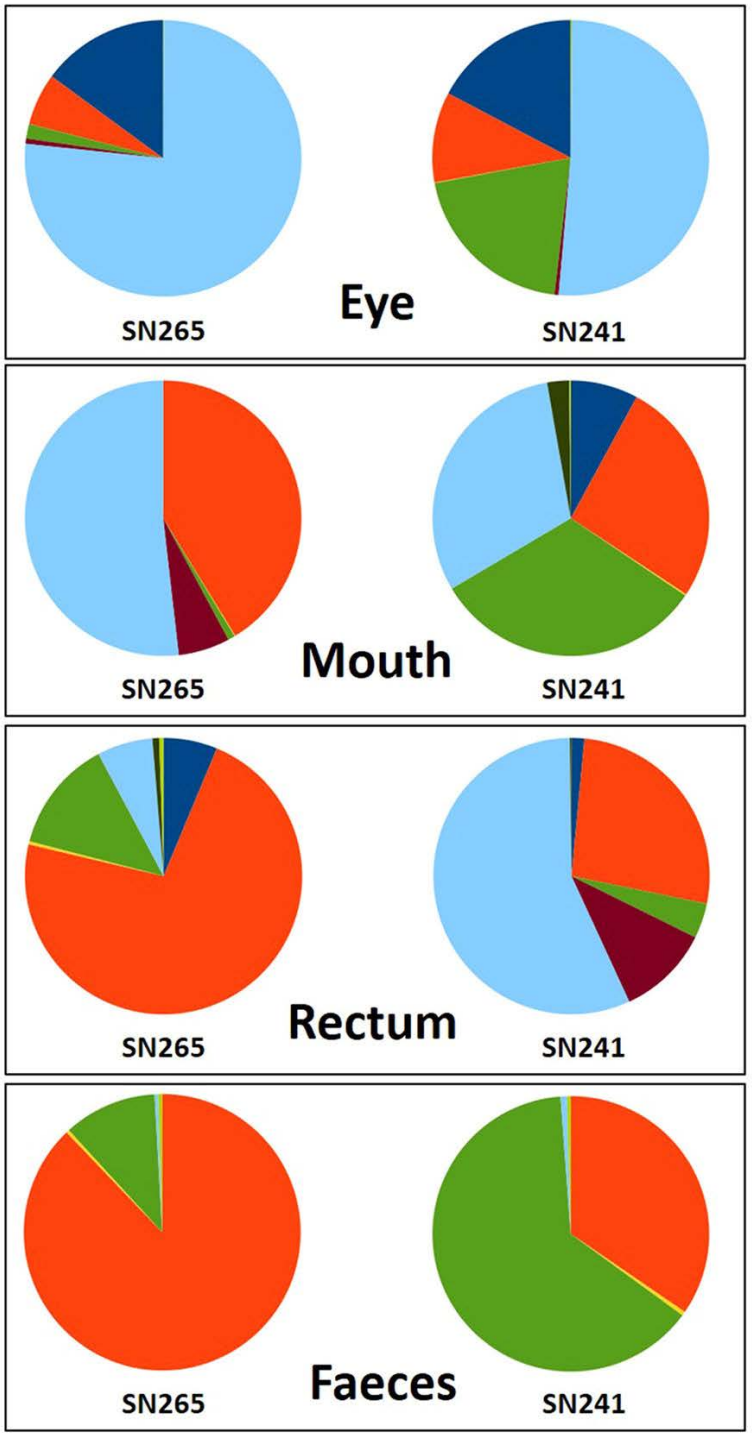
Niccoló Alfano[1], Alexandre Courtiol[1], Hanna Vielgrader[2], Peter Timms[3], Alfred L. Roca[4] & Alex D. Greenwood[1]

Metagenómica

- Datos



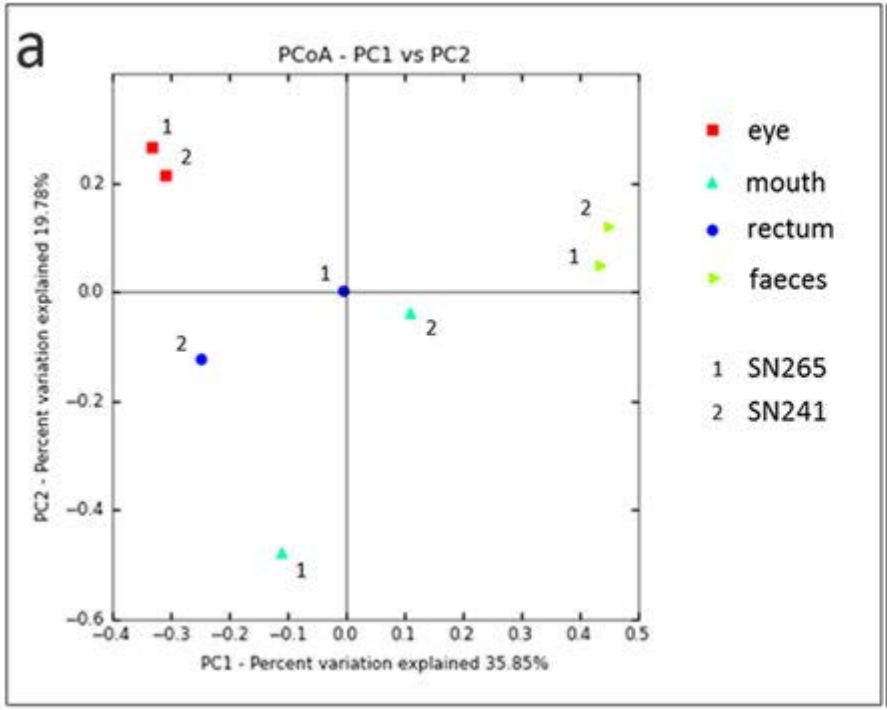
Dos koalas capticos del zoologico Tiergarten Schönbrunn en Vienna (Austria): un macho de 14 años (SN241) y una hembra de 12 años (SN265)



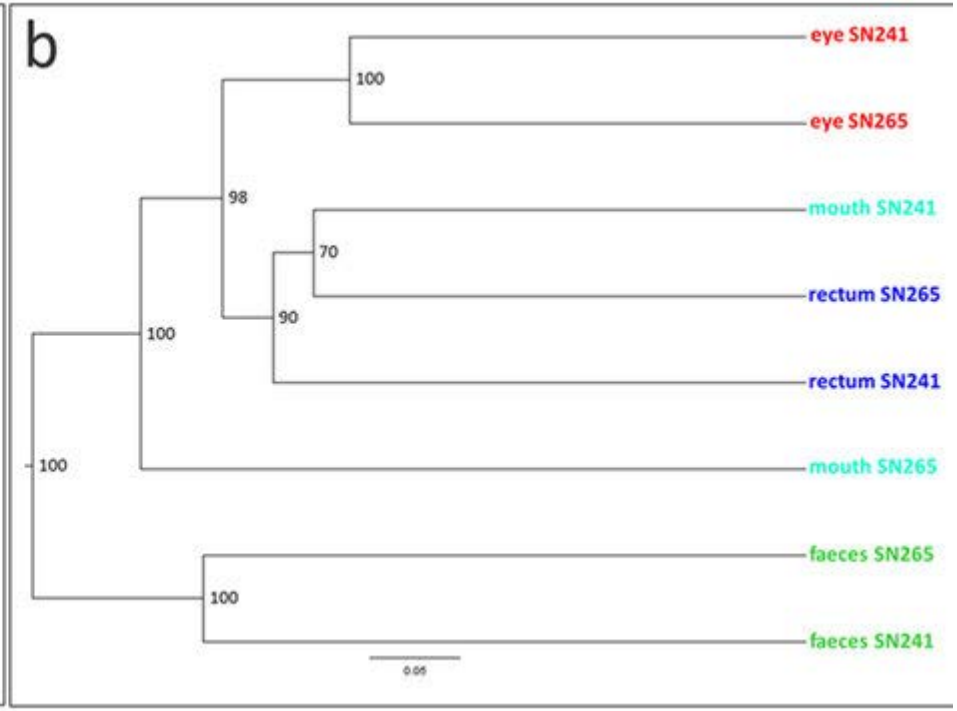
Metagenómica

- Datos

Análisis de componentes principales

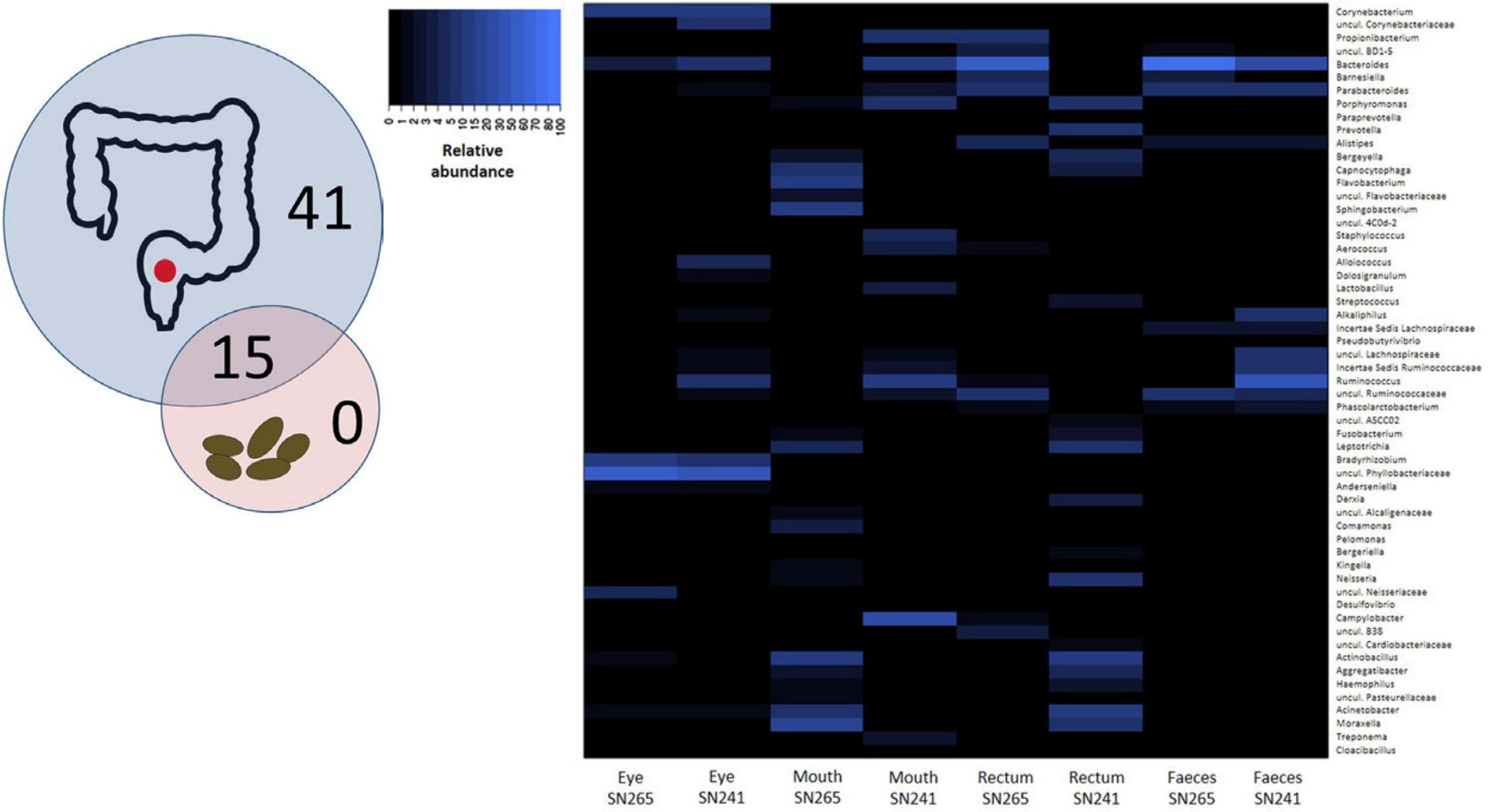


Medición de distancias filogenéticas



Metagenómica

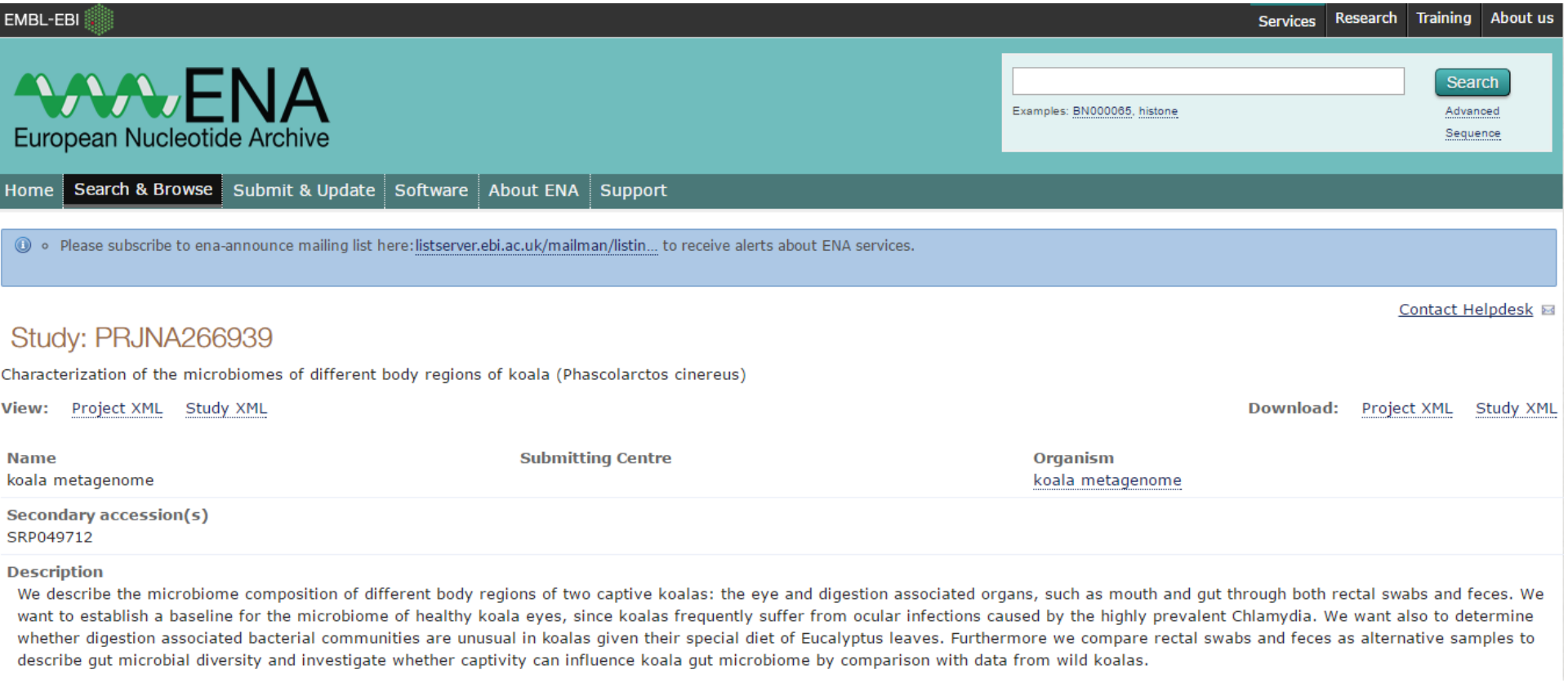
- Resultados



Metagenómica

https://www.ebi.ac.uk/ena/data/view/PRJNA266939

- Estudio PRJNA266939



Metagenómica

Showing results 1 - 8 of 8 results

Study accession	Sample accession	Secondary sample accession	Experiment accession	Run accession	Tax ID	Scientific name	Instrument model	Library layout	FASTQ files (FTP)
PRJNA266939	SAMN03176826	SRS742864	SRX757717	SRR1648867	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176863	SRS743141	SRX757994	SRR1648868	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176856	SRS743421	SRX758296	SRR1649356	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176857	SRS743422	SRX758297	SRR1649357	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176859	SRS743423	SRX758298	SRR1649358	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176860	SRS743424	SRX758299	SRR1649359	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176861	SRS743425	SRX758300	SRR1649360	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2
PRJNA266939	SAMN03176862	SRS743426	SRX758301	SRR1649361	1573835	koala metagenome	Illumina MiSeq	PAIRED	File 1 File 2

Ojo
Koala Pci_SN241 (SRR1648868)
Koala Pci_SN265 (SRR1648867)

Heces
Koala Pci_SN241 (SRR1649361)
Koala Pci_SN265 (SRR1649360)

SN_241: Macho
SN_265: Hembra

Metagenómica

- Metodología
  - Universal primers 27F (5´-AGAGTTTGATCCTGGCTCAG-3´) and 338R (5´-TGCTGCCTCCCGTAGGAGT-3´)[53] were used for PCR amplification of the V1–V2 hypervariable regions of the bacterial 16S rRNA gene.
  - The libraries were first amplified in a 50 µl volume reaction using 5 µl of DNA library, 0.5 µl Herculase II Fusion DNA Polymerase (Agilent Technologies Inc.), 10 µl Herculase II Reaction Buffer (5x), 0.5 µl dNTPs (25 mM), 1 µl Single Index Primer P5 (10 µM), 1 µl Illumina Index Primer P7 (10 µM) and sterile distilled water to volume. A unique P7 Index Primer was used for each library to allow for subsequent sample discrimination after the sequencing of pooled libraries.
  - The indexed DNA libraries were then pooled at equimolar concentrations for paired-end sequencing (2 × 250) on an Illumina MiSeq v2 platform at the Danish National High-Throughput DNA Sequencing Centre in Copenhagen, Denmark.

Metagenómica

- Pre-procesamiento

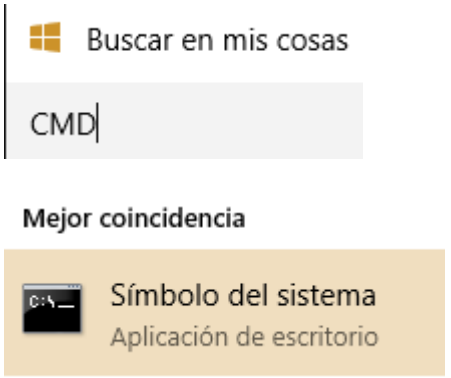
http://www.drive5.com/usearch/download.html

Introduction

In SSU metagenomics, next-generation reads are clustered into Operational Taxonomic Units (OTUs). This requires quality filtering, dereplication, discarding singletons (optional), and finally clustering into OTUs, typically at a 97% identity threshold.

Benchmark results

The OTU benchmark uses 454 Titanium and Illumina MiSeq reads of Even and Staggered mock communities used for protocol development in the Human Microbiome Project (HMP). USEARCH results were obtained with the same parameters for all samples. The number of reads per sample ranges from 10,000 (Titanium) to two million (MiSeq). The accuracy of UPARSE was compared to recommended procedures (Sept. 2012) for mothur, QIIME and AmpliconNoise.



UPARSE OTU analysis

USEARCH benchmarks

Metagenómica

- USEARCH

http://www.drive5.com/usearch/download.html

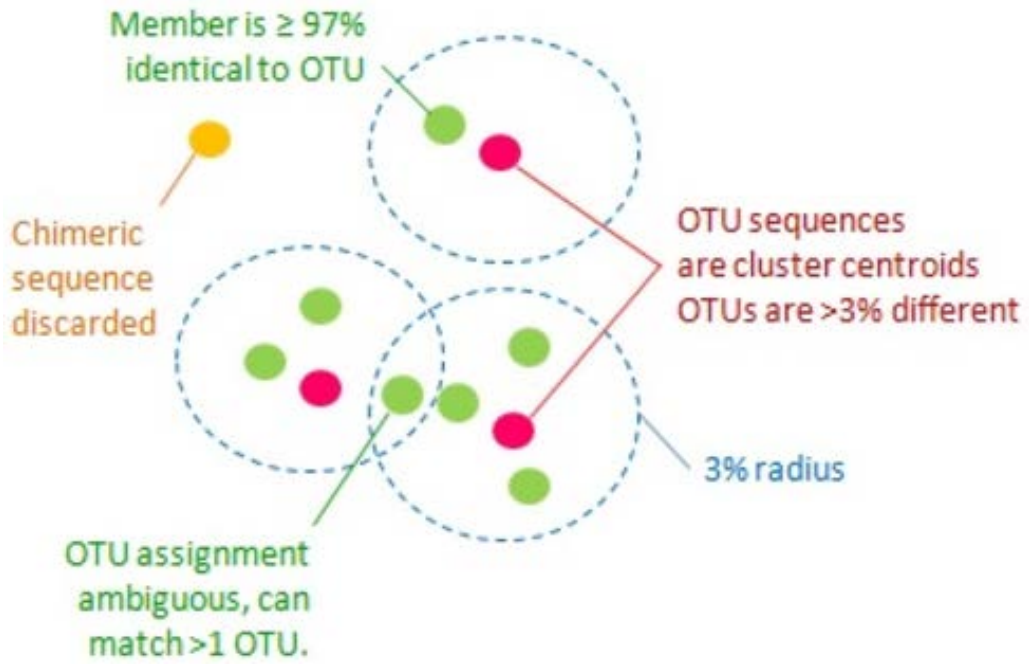
Read preparation
Assemble paired reads, quality filter, trim lengths, find unique sequences

OTU clustering / denoising
Select OTU sequences

Construct OTU table
Map reads to OTUs to get counts per sample

Quality control
Check OTU sequences and analyze control samples

Diversity and taxonomy analysis
Calculate alpha and beta diversity from OTU table
Predict taxonomy for OTU sequences



Clustering criteria

1. All pairs of OTU sequences should have < 97% pair-wise sequence identity.
2. Chimeric sequences should be discarded.
3. All non-chimeric input sequences should match at least one OTU with ≥ 97% identity.

Metagenómica

http://www.drive5.com/usearch/download.html

- USEARCH

Read preparation
Assemble paired reads, quality filter, trim lengths, find unique sequences

Read preparation

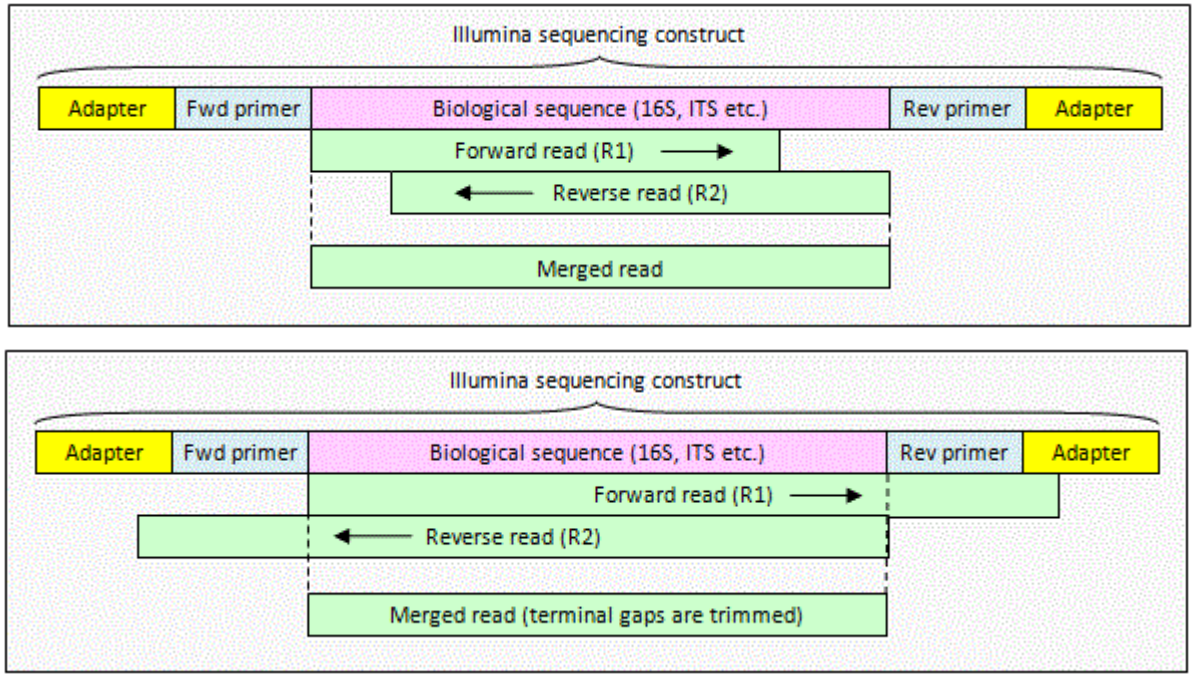
See also
OTU / denoising pipeline

The table below summarizes the read preparation steps that I recommend for an OTU / denoising pipeline. Follow the links for details.

	Step	Description
	Understand your reads	Investigate your reads, don't blindly follow a recipe
	Demultiplex	Assign reads to samples using index reads or barcodes
1	Merge pairs	Merge paired reads to get consensus sequences and Q scores
3	Strip primers	Primer-binding sequence should be removed before quality filtering
2	Orient	If you have reads on both strands, orient before trimming and finding uniques
3	Strip machine sequences	Remove machine-specific sequences e.g. TCAG for 454
4	Length trimming	Remove low-quality tails, make sure 3' ends align
5	Quality filtering	Making OTUs (but not the OTU table) needs high-quality reads
	Pool samples	Reads for all samples should be combined
6	Discard singletons	Remove low-abundance reads, which often have errors

Metagenómica

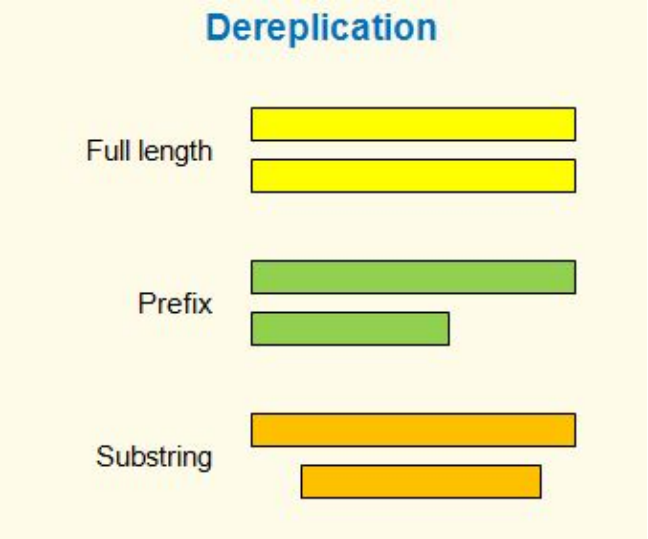
- USEARCH



The non-biological sequence will be deleted automatically because fastq_mergepairs detects staggered alignments and deletes terminal gaps before building a consensus sequence.

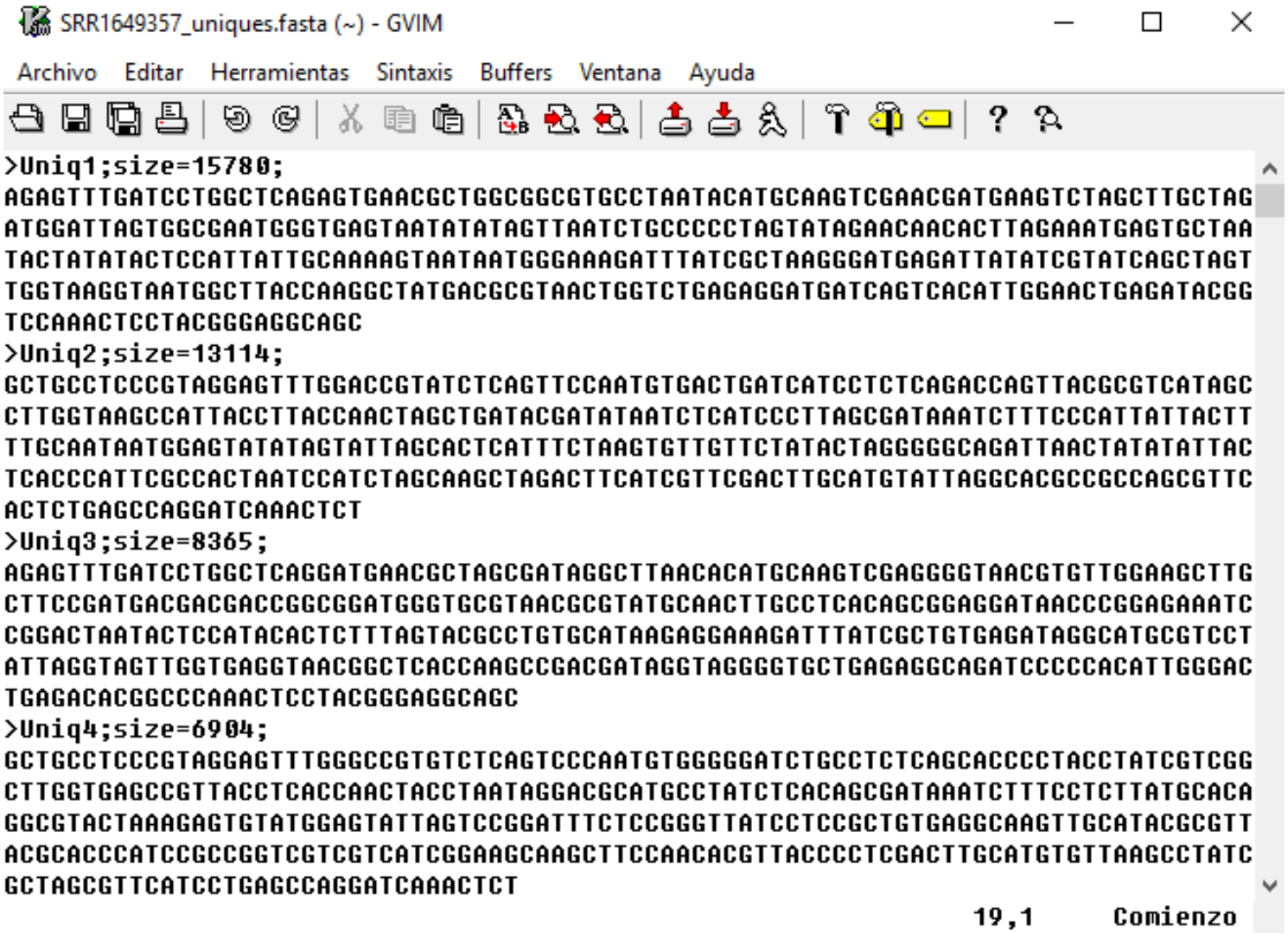
Metagenómica

- USEARCH Paso A: preparación de las lecturas
- 1. Unir lecturas R1 y R2: usearch -fastq_mergepairs
- 2. Orientar secuencias: usearch -orient
- 3. Eliminar secuencias de primers: usearch -fastx_truncate
- 4. Filtrar por longitud y calidad: usearch -fastq_filter
- 6. Dereplicar: usearch -fastx_uniques



Metagenómica

- USEARCH



Metagenómica

Singletons

Definition of a singleton

A singleton is a read with a sequence that is present exactly once, i.e. is unique among the reads.

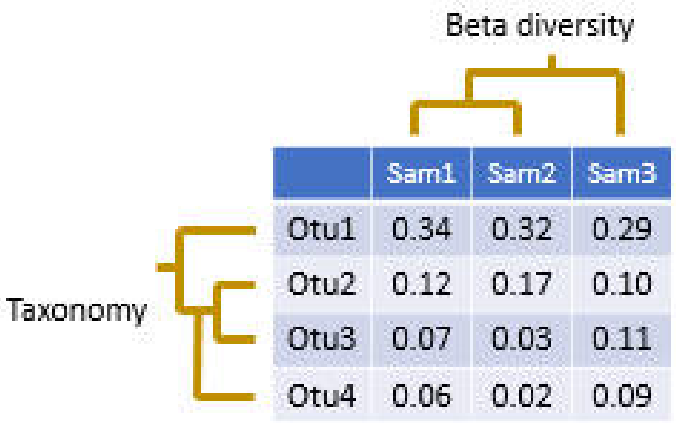
- USEARCH Paso B: Generar OTUs y tabla de OTUs
- 7. Genearar OTUs: usearch -cluster_otus
- 8. Crear tabla de OTUs: usearch -otutab
- 9. Normalizar la tabla de OTUs: usearch -otutab_norm
- 10. Eliminar OTUs poco representativos: usearch -otutab_trim

Metagenómica

- USEARCH

Clustering and chimera filtering Asignar lecturas a OTU (OTU table)

#OTU ID	F3D0	F3D141	F3D142	F3D143	F3D144	F3D145	F3D146	F3D147
OTU_6	749	535	313	372	607	849	493	2025
OTU_25	29	57	14	2	14	22	16	127
OTU_1	613	497	312	247	472	719	349	1720
OTU_8	426	378	255	237	382	627	330	1417
OTU_31	149	38	10	19	25	21	43	31
OTU_2	366	392	327	185	313	542	248	1367
OTU_7	196	370	92	107	48	155	74	105
OTU_10	46	169	87	109	171	209	120	864
OTU_80	26	6	0	1	4	8	18	11



Metagenómica

- USEARCH Paso C: Control de calidad de los OTUs
- 11. Alinear OTUs a base de datos: usearch -usearch_global
- 12. Verificar longitud de los OTUs: usearch -fastx_info
- 13. Verificar orientacion de los OTUs: usearch –orient
- 14. Filtrar OTUs de baja complejidad: usearch -filter_lowc
- 15. Filtrar secuencias PhiX en OTUs: usearch -search_phix

Metagenómica

- USEARCH
- Control de calidad de los OTUs obtenidos

Issue	Description
Alignments	Do the OTU sequences align well to a reference database for your gene?
Missing OTUs	Do all OTUs appear in the OTU table?
Coverage	How much of the data is explained by the OTUs?
Short contructs	Bad sequencing construct created by PCR
Strand duplicates	Sequences of both plus and minus strands
Offsets	Sequences start at different positions in the gene
Cross-talk	Reads assigned to the wrong sample.
Sequence error	Polymerase errors and bad base calls
Low complexity	Sequencer noise
PhiX	Unfiltered spike-in
Chimeras	Unfiltered PCR chimeras
Mistargeting	Primers amplify a different region
Contaminants	Self-explanatory
Primers	Primer-binding sequences should be stripped at the start of the pipeline
Tight OTUs	OTUs >97% identical

Metagenómica

- USEARCH Paso D: Análisis de diversidad
- 16. Alpha diversidad: usearch -alpha_div

- 17. Rarefaccion: usearch -alpha_div_rare

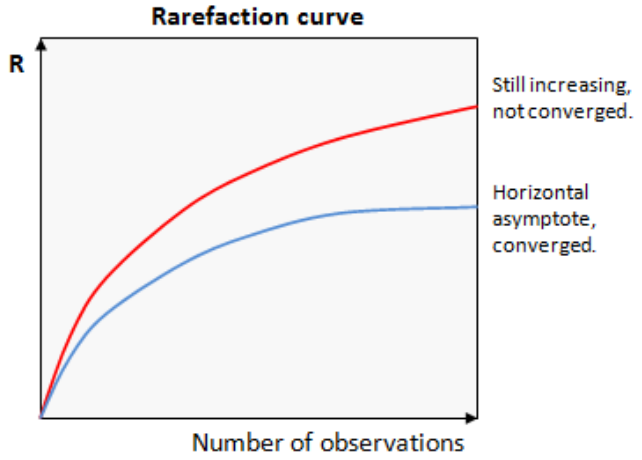
- 18. Beta diversidad: usearch -beta_div

Metagenómica

- USEARCH
- Estimating microbial diversity
  - Diversidad alfa: The diversity in a single sample (alpha diversity) is commonly measured using metrics such as the Shannon index and the Chao1 estimator
  - Diversidad beta: The variation between pairs of samples (beta diversity) is measured using metrics such as the Jaccard distance or Bray-Curtis dissimilarity.
  - Such metrics are calculated from OTU frequencies.

  - Other metrics, e.g. unweighted UniFrac (called unifrac_binary in usearch) use presence / absence only, effectively considering a count to be one if it is any non-zero value.

  - Predecir taxonomía de los OTUs (SINTAX



Metagenómica

- USEARCH Paso E: Asignación taxonómica
- 19. Asignar taxonomía a los OTUs: usearch –sintax

- 20. Resumen de reporte de taxonomía a filo y genero: usearch -sintax_summary

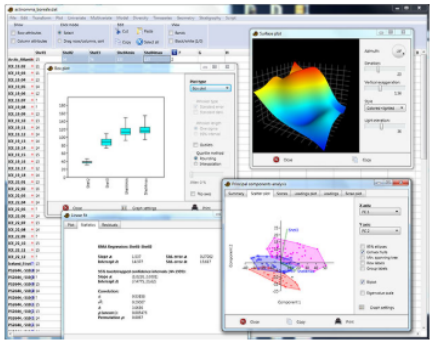
- 21. Análisis estadístico con PAST3: generar PCoA o NMDS

Metagenómica

https://folk.uio.no/ohammer/past/

- Análisis estadístico con PAST3

Past 3.x - the Past of the Future



Past is free software for scientific data analysis, with functions for data manipulation, plotting, univariate and multivariate statistics, ecological analysis, time series and spatial analysis, morphometrics and stratigraphy.

Past went through a complete redesign in 2013.

Past works under Windows 7, Windows 10, and Mac OSX.

Current version (April 2019): **3.24**

Downloads

Windows, zipped exe, version 3.24 (5.5 MB)
Mac, version 3.24 (24 MB), only OSX 10.8 (Mountain Lion) and later

PDF manual (8 MB)

The old Past (version 2.17) with manual is still available, here.

Resources

Frequently Asked Questions!

Terms, reference, etc.

Version history

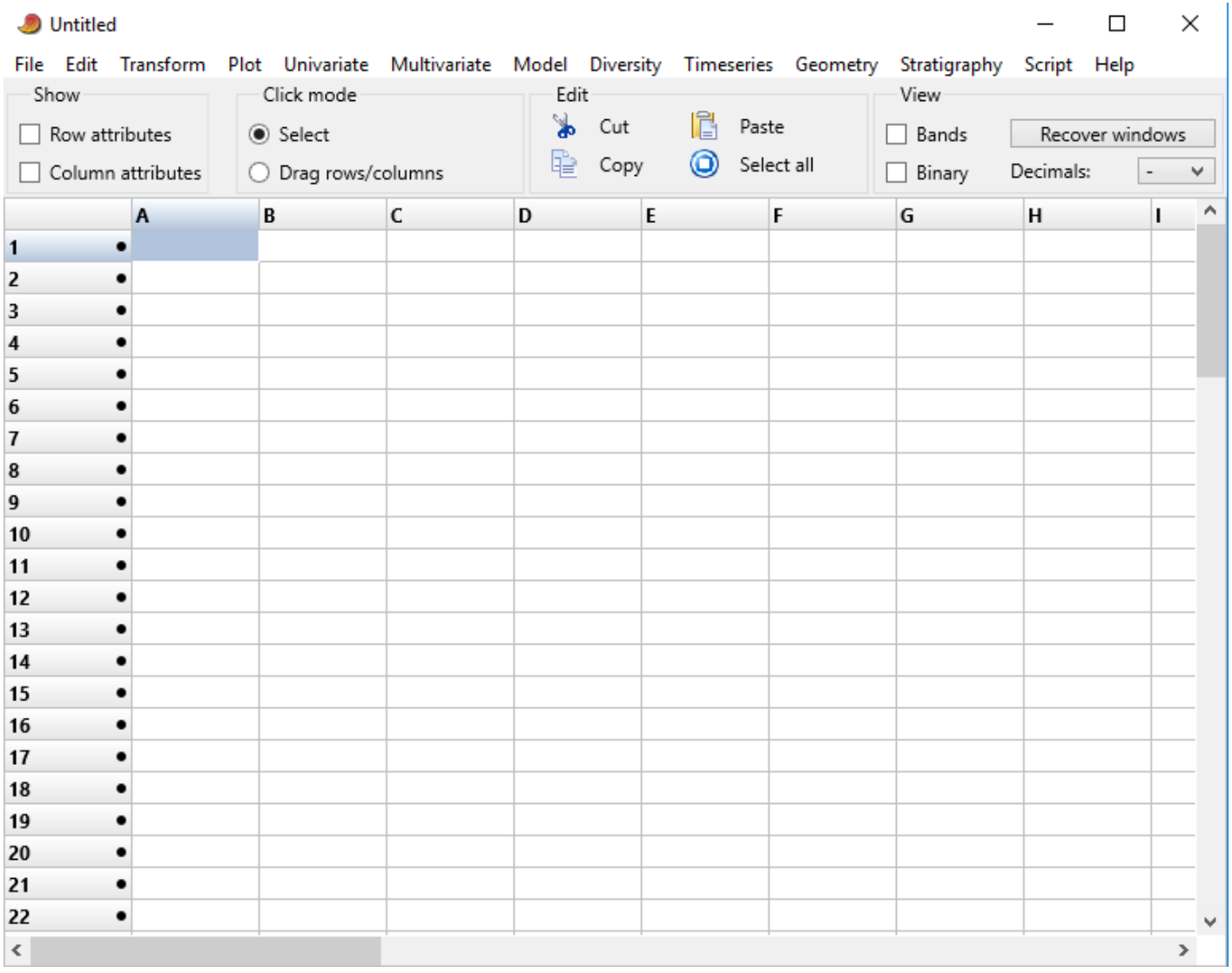
PAST mailing list

Øyvind Hammer, Natural History Museum, University of Oslo [ohammer(at)nhm.uio.no]

Metagenómica

https://folk.uio.no/ohammer/past/

- Análisis estadístico con PAST3



Metagenómica

- Tarea

Caporaso *et al. Genome Biology* 2011, **12**:R50
http://genomebiology.com/2011/12/5/R50

RESEARCH Open Access

Moving pictures of the human microbiome

J Gregory Caporaso[1], Christian L Lauber[2], Elizabeth K Costello[3], Donna Berg-Lyons[2], Antonio Gonzalez[4], Jesse Stombaugh[1], Dan Knights[4], Pawel Gajer[5], Jacques Ravel[5], Noah Fierer[2,6], Jeffrey I Gordon[7] and Rob Knight[1,8*]

Abstract

Background: Understanding the normal temporal variation in the human microbiome is critical to developing treatments for putative microbiome-related afflictions such as obesity, Crohn's disease, inflammatory bowel disease and malnutrition. Sequencing and computational technologies, however, have been a limiting factor in performing dense time series analysis of the human microbiome. Here, we present the largest human microbiota time series analysis to date, covering two individuals at four body sites over 396 timepoints.

Results: We find that despite stable differences between body sites and individuals, there is pronounced variability in an individual's microbiota across months, weeks and even days. Additionally, only a small fraction of the total taxa found within a single body site appear to be present across all time points, suggesting that no core temporal microbiome exists at high abundance (although some microbes may be present but drop below the detection threshold). Many more taxa appear to be persistent but non-permanent community members.