



Title: Bioinformatic practical applications in biotechnology, medicine, environmental and agricultural sciences

Authors: RAGGI, Luciana, JIMENEZ-JACINTO, Verónica and ESCOBAR-ZEPEDA, Alejandra

Editorial label RINOE: 607-8695

VCIERMMI Control Number: 2023-02

VCIERMMI Classification (2023): 261023-0002

Pages: 68

RNA: 03-2010-032610115700-14

MARVID - Mexico

Park Pedregal Business. 3580-
Adolfo Ruiz Cortines Boulevard –
CP.01900. San Jerónimo Aculco-
Álvaro Obregón, Mexico City
Skype: MARVID-México S.C.
Phone: +52 | 55 6159 2296
E-mail: contact@marvid.org
Facebook: MARVID-México S. C.
Twitter: @Marvid_México

www.marvid.org

Holdings

Mexico	Colombia	Guatemala
Bolivia	Cameroon	Democratic
Spain	El Salvador	Republic
Ecuador	Taiwan	of Congo
Peru	Paraguay	Nicaragua

Introduction

Methodology

Results

Conclusions

References

Introduction

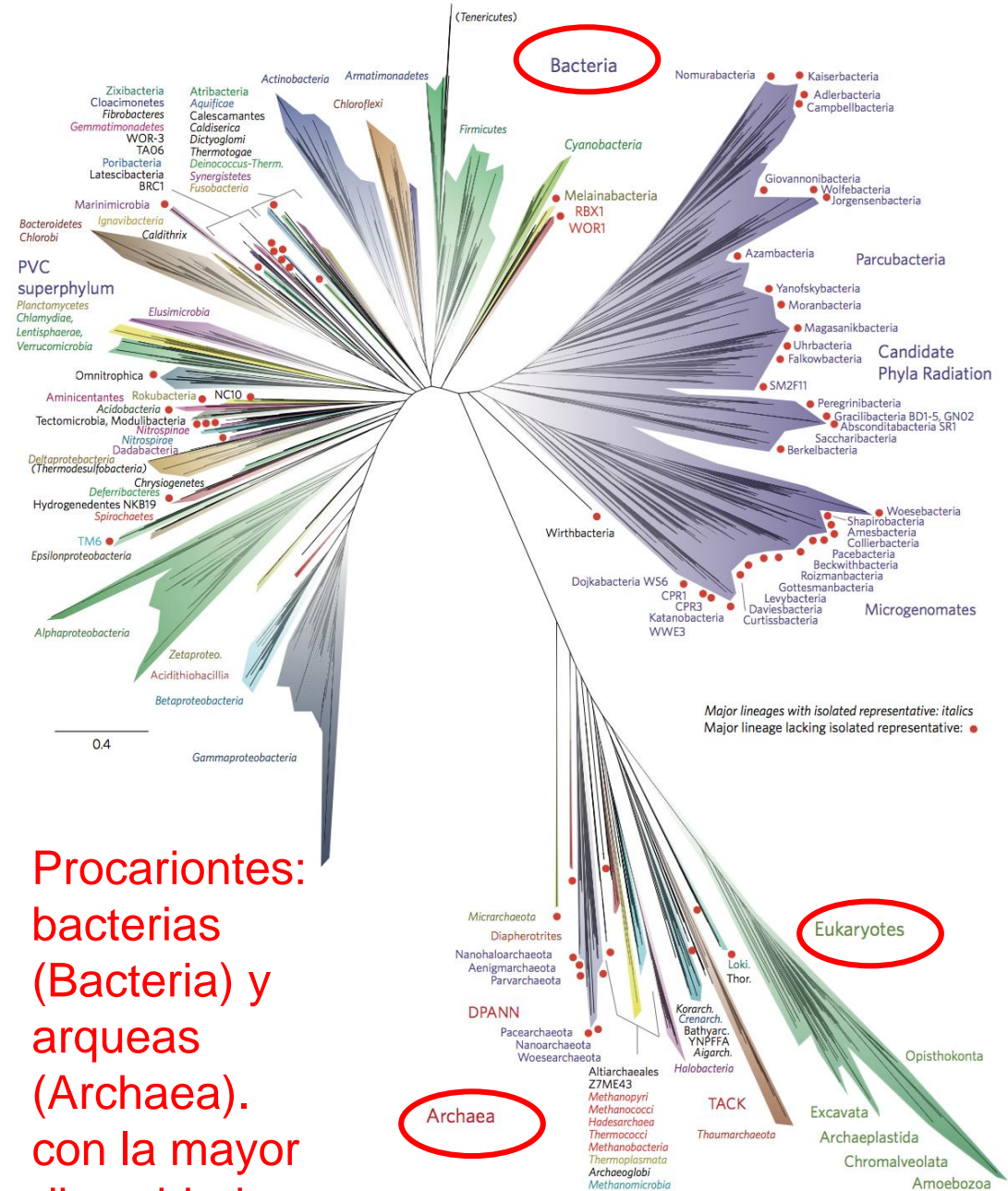
Tree of life - TOL

Árbol de la vida con 3 Dominios

(2 procarióticos y 1 eucariótico).

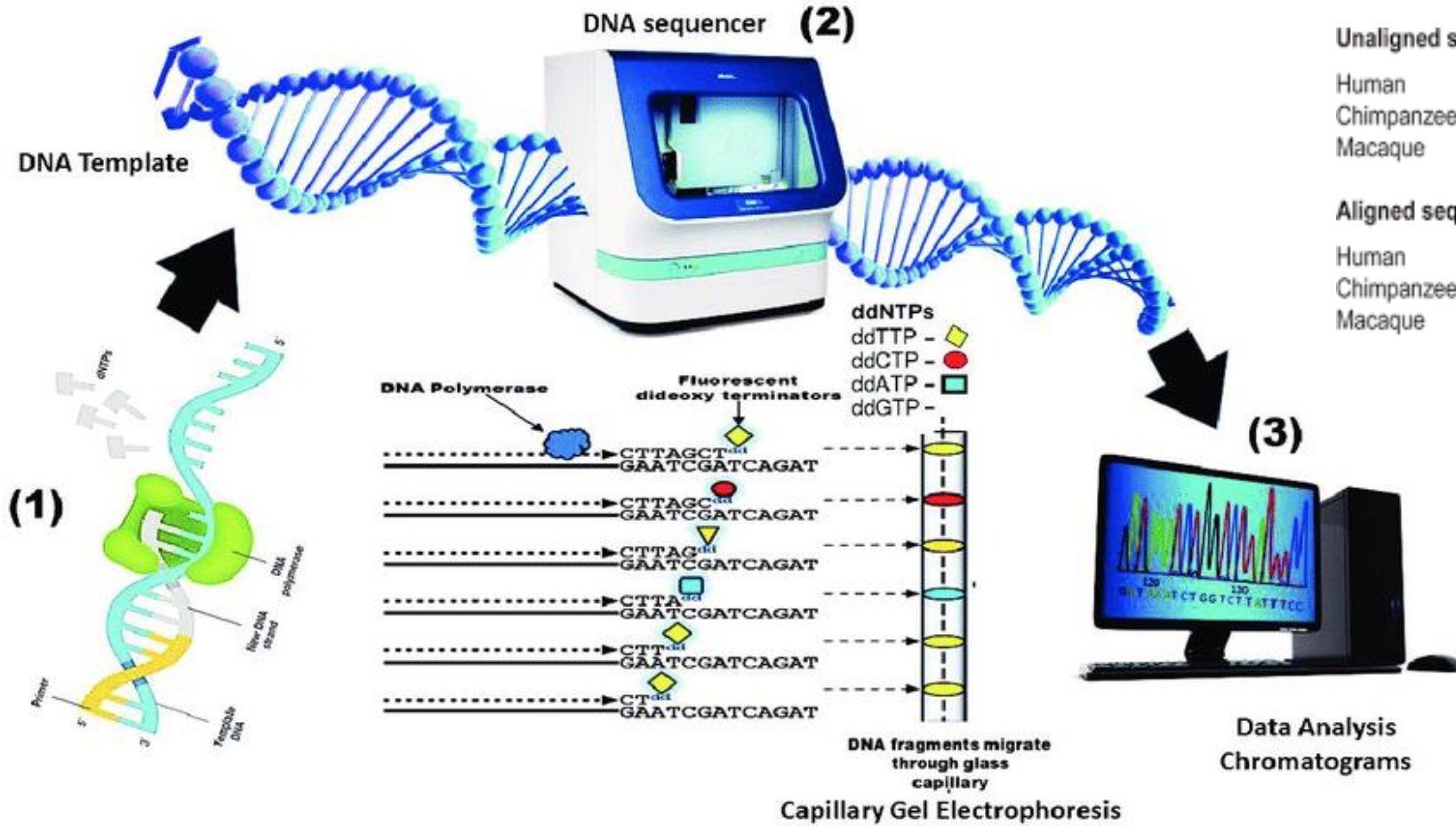
Todos los organismos vivos tienen características semejantes, en particular su información genética guardada en el DNA.

A partir de esta información se puede generar un árbol que los/nos une a todos los organismos vivos.



Procariontes:
bacterias
(Bacteria) y
arqueas
(Archaea).
con la mayor
diversidad.

Secuenciando DNA - base de la biología molecular



Unaligned sequences

Human **A****C****A****T****T****A****T****G****G****A****C****A****G****G****T****A****A****G****T****A****A****A****A****A****C****A****T****A****T****T**
 Chimpanzee **A****C****A****T****T****A****T****G****G****A****C****A****G****G****T****A****A****G****T****A****A****A****A****A****C****A****T****A****T****T**
 Macaque **A****T****A****T****A****C****A****T****T****A****C****G****G****A****C****A****G****T****A****A****G****T****A****A****A****A****C****A****T**

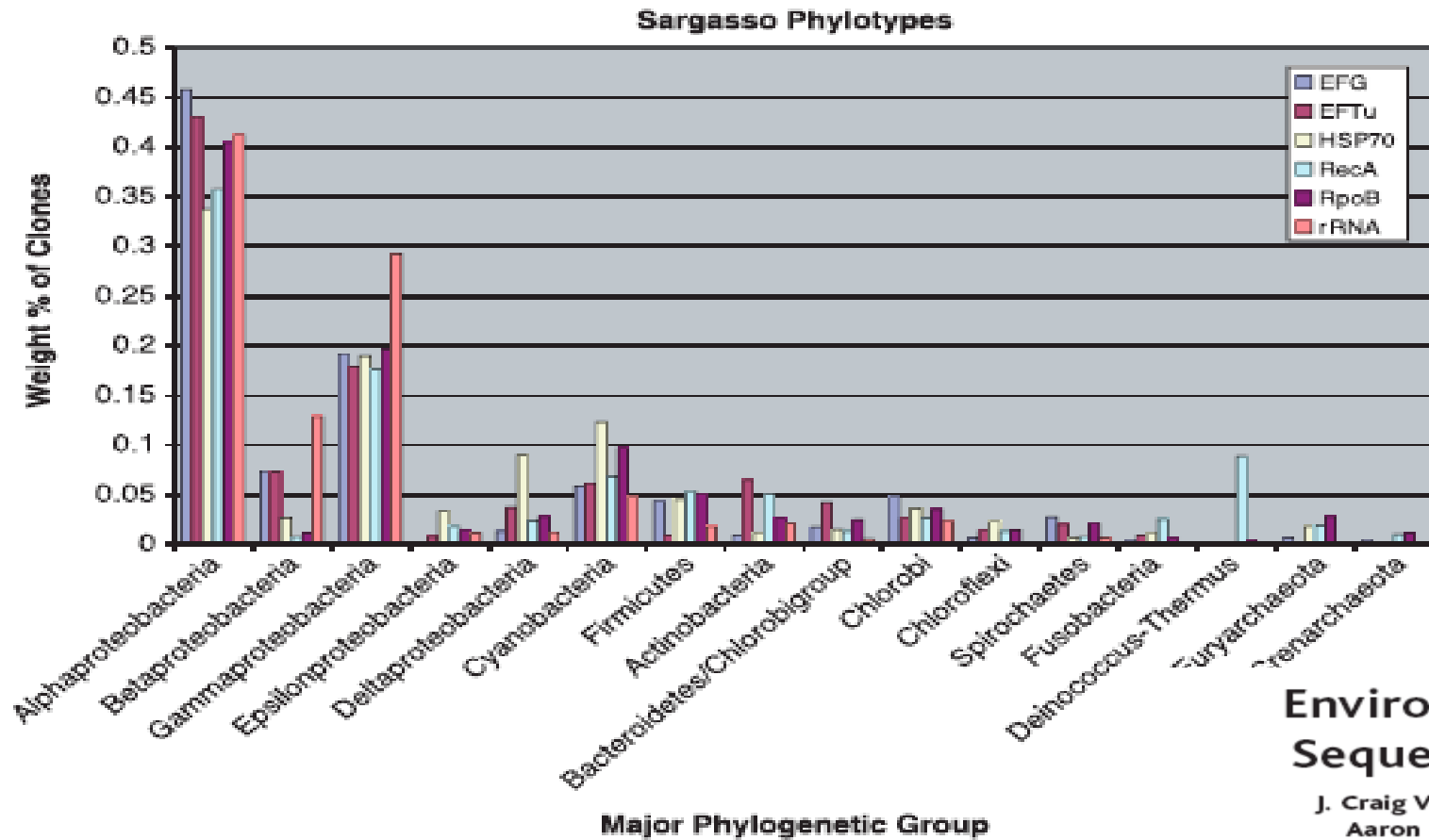
Aligned sequences

Human **A****C****A** **T****T****A****T****G****G****A****C****A****G****G****T****A****A****G****T****A****A****A****A****A****C****A****T****A****T****T**
 Chimpanzee **A****C****A** **T****T****A****T****G****G****A****C****A****G****G****T****A****A****G****T****A****A****A****A****A****C****A****T****A****T****T**
 Macaque **A****T****A****T****A****C****A** **T****T****A****C****G****G****A****C****A****G****T****A****A****G****T****A****A****A****A****C****A****T**

Frog
 Chicken
 Human
 Rabbit
 Mouse
 Opossum

G	C	T	G	A	C	T	T	C	T	G	A	G	G	T	T	
G	C	G	T	A	A	C	T	T	C	A	C	A	T	G	A	T
G	C	G	T	C	A	C	T	T	G	A	G	A	C	G	C	T
G	C	G	T	C	A	C	T	T	G	A	G	A	C	G	C	T
G	C	G	T	C	A	C	T	T	G	A	C	A	G	G	C	T
G	C	G	T	C	A	C	T	T	G	A	G	A	C	G	C	T

Uso de Oligos



Environmental Genome Shotgun Sequencing of the Sargasso Sea

J. Craig Venter,^{1*} Karin Remington,¹ John F. Heidelberg,³
 Aaron L. Halpern,² Doug Rusch,² Jonathan A. Eisen,³
 Dongying Wu,³ Ian Paulsen,³ Karen E. Nelson,³ William Nelson,³
 Derrick E. Fouts,³ Samuel Levy,² Anthony H. Knap,⁶
 Michael W. Lomas,⁶ Ken Neelson,⁵ Owen White,³
 Jeremy Peterson,³ Jeff Hoffman,¹ Rachel Parsons,⁶
 Holly Baden-Tillson,¹ Cynthia Pfannkoch,¹ Yu-Hui Rogers,⁴
 Hamilton O. Smith¹

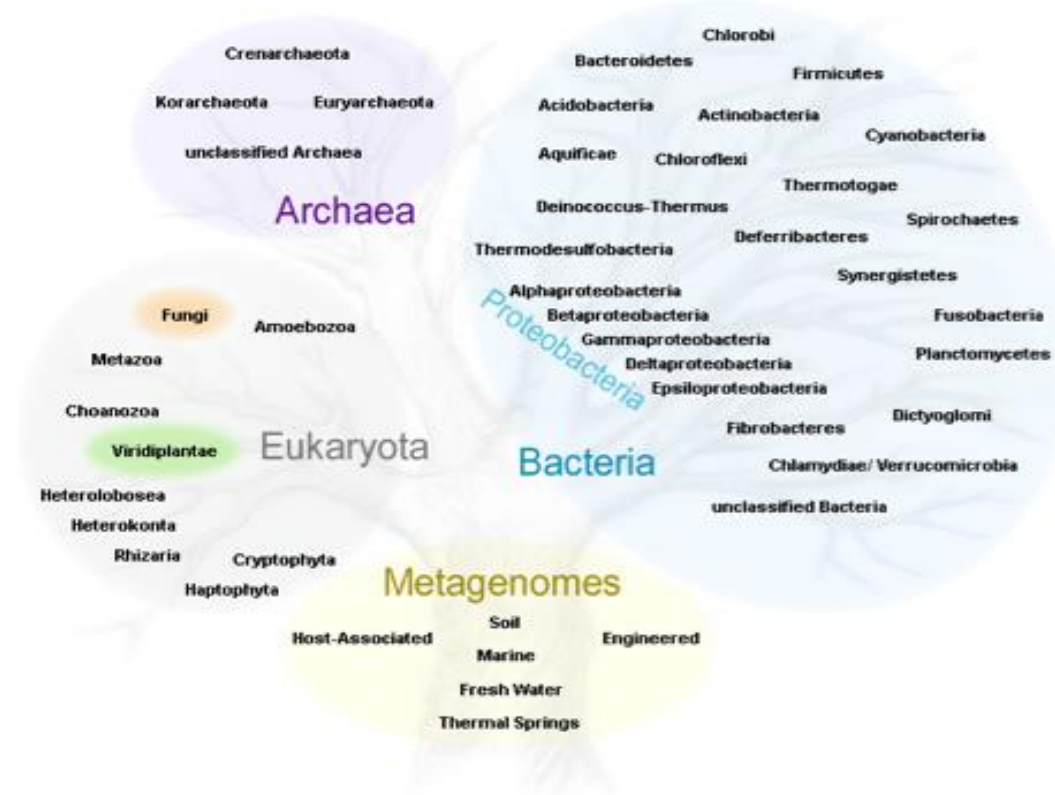
http://genome.jgi.doe.gov/

JGI Home Project List Login

by keyword Search Genomes

all JGI Organisms and Abiotrophia defectiva ATCC 49176 and Home

GO!



"Tree of Life" drawing by Leila Hornick, copyright 2005

TREE OF LIFE: To use the tree navigation: click a branch name and select the system for the organism of interest.



[Integrated Microbial Genomes \(IMG\)](#) and [Metagenomes \(IMG/M\)](#) - resources for comparative analysis and annotation of all publicly available genomes.

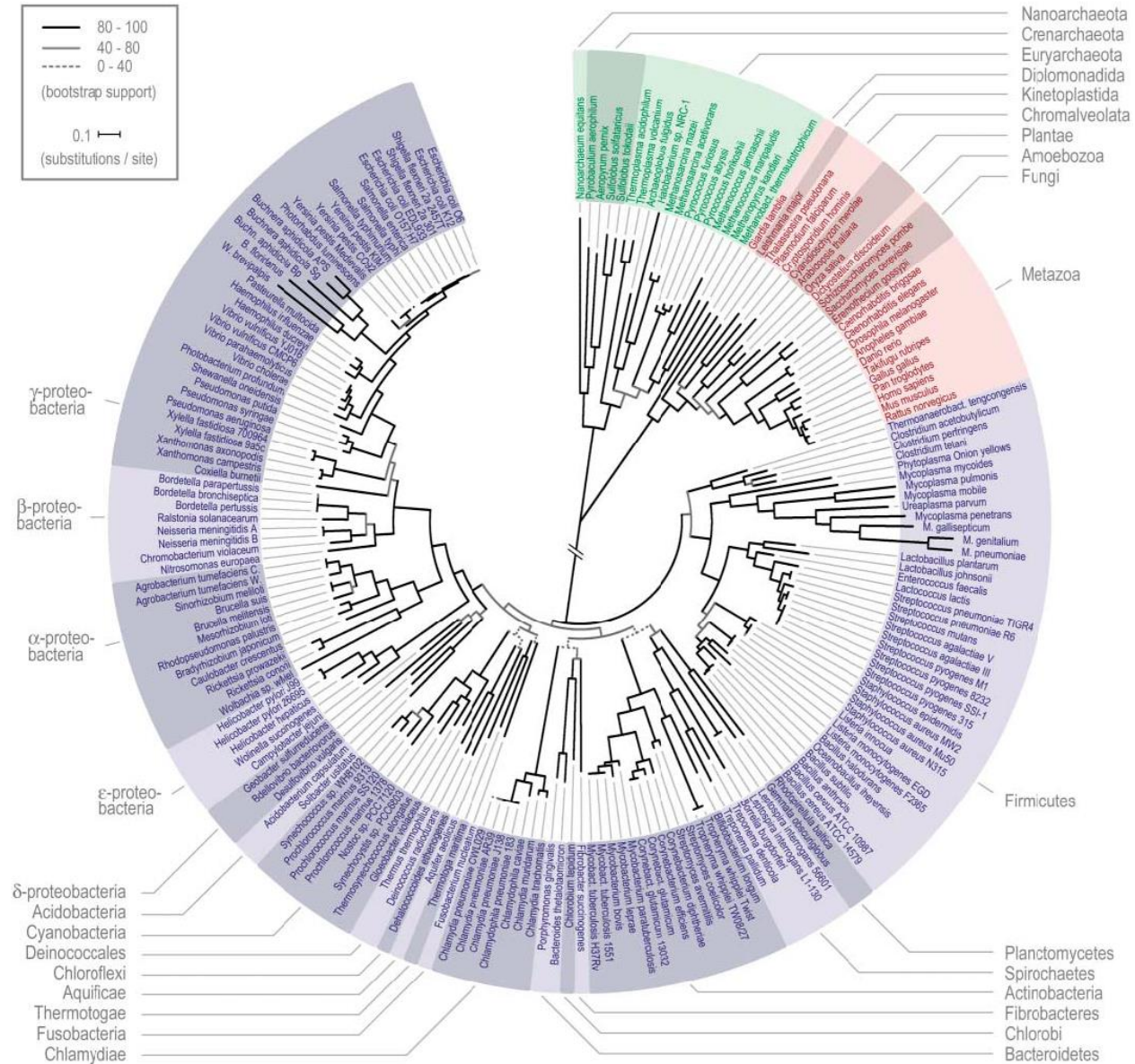


[Mycocosm](#) - the Fungal Genomics Resource. Provides access to the annotated fungal genomes and tools for their analysis.



[Phytozome](#) - a comparative hub for green plant genomes and gene family data and analysis.

Árbol de genomas completos

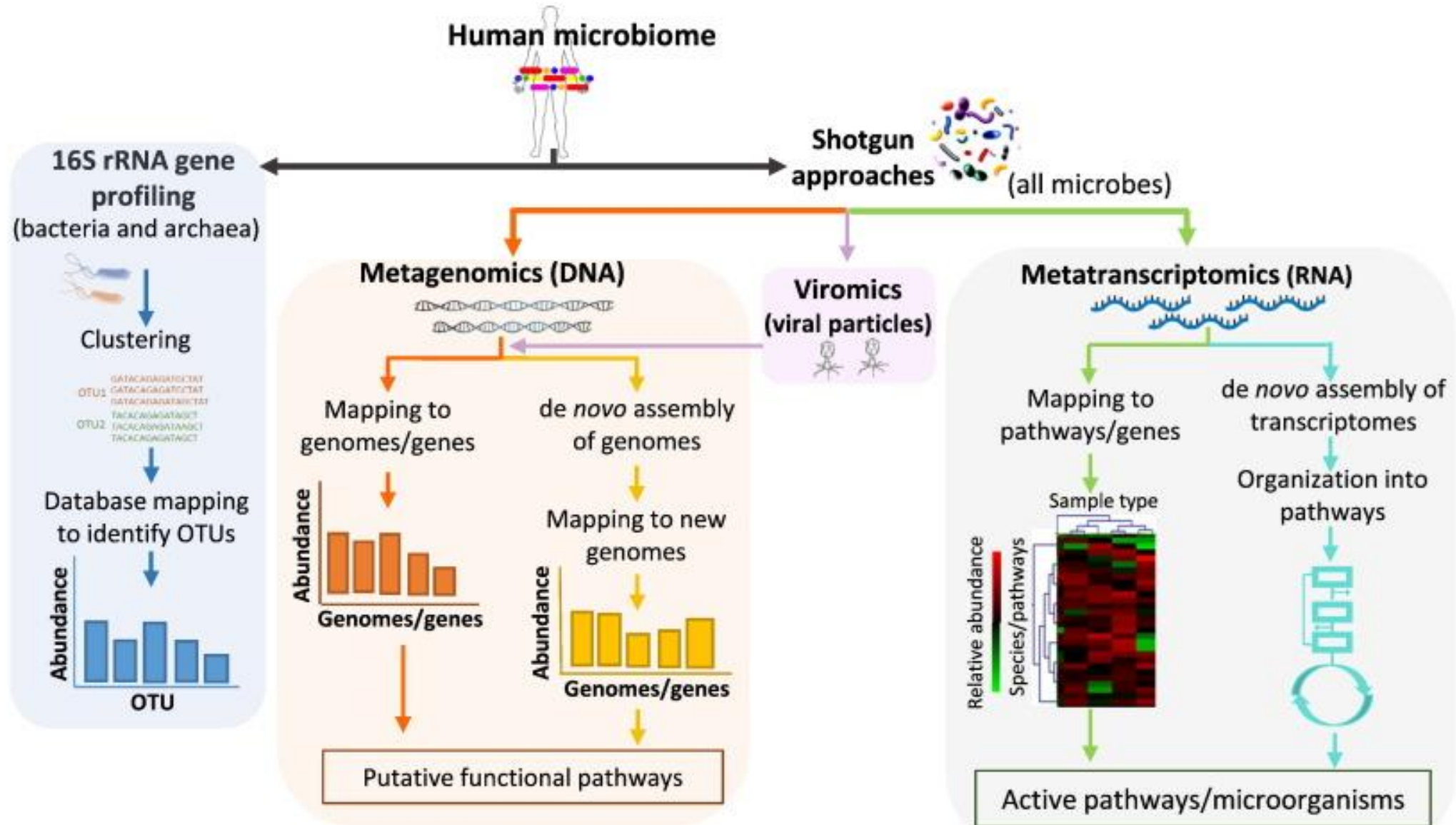


Arqueas

Eucariontes

Bacterias

Microbiómica

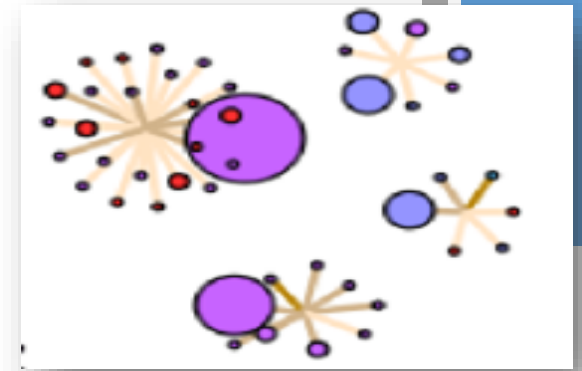


Methodology

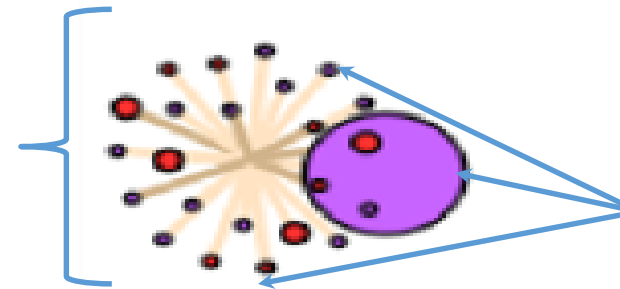


Immune diversity

- Identificación de los componentes del repertorio de IGHV
 - Grupos clonales y linajes VH
 - Frecuencia absoluta y relativa

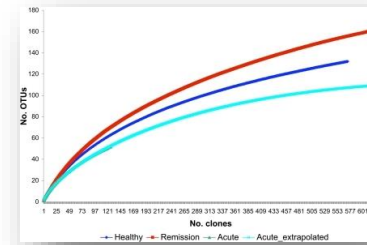


Grupo Clonal



Idiotipos

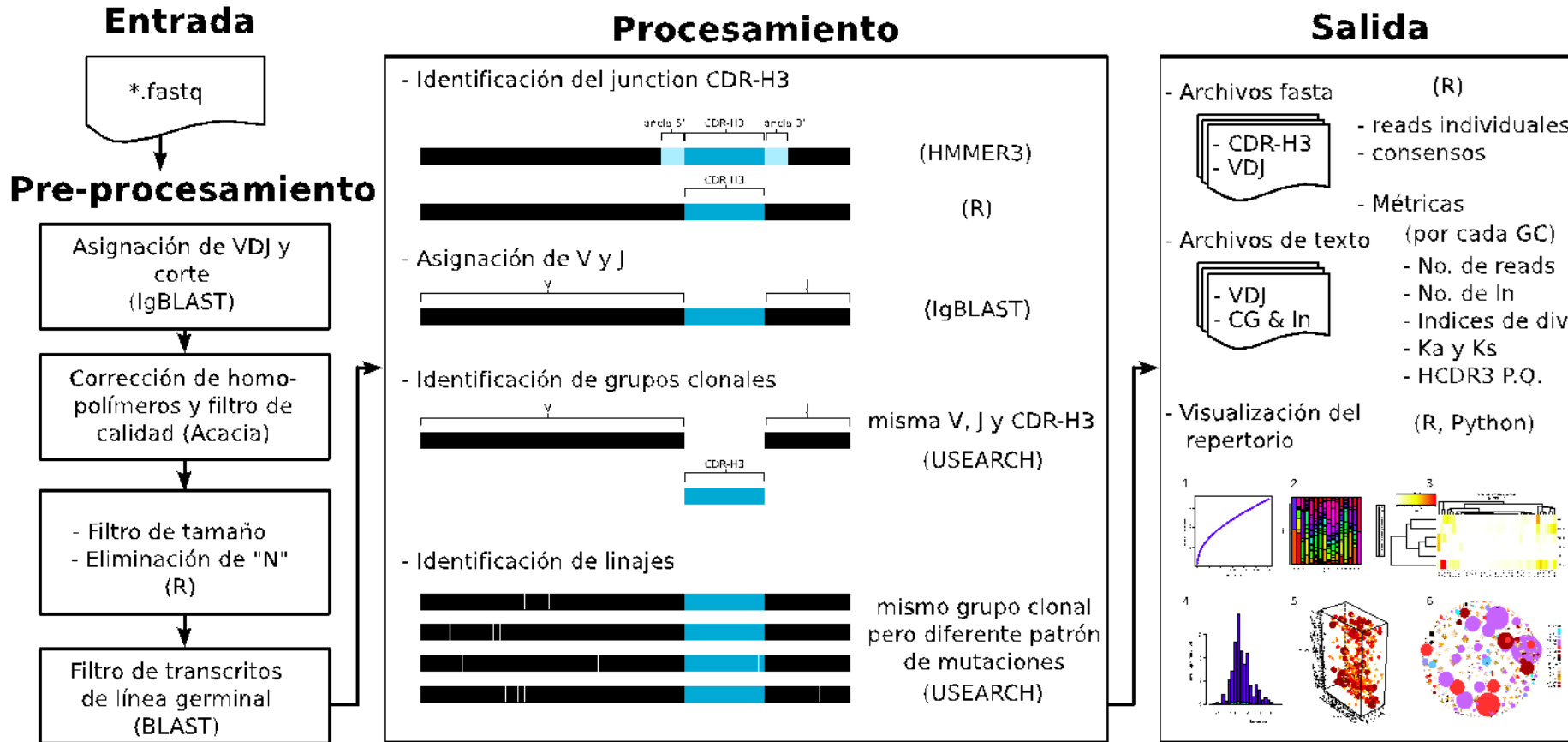
- Descripción del repertorio
 - Índices de diversidad
 - Gráficas de rarefacción
 - Medición de HSM



$$H' = - \sum_{i=1}^R p_i \ln p_i$$

- Comparación del repertorio
- Minería de repertorio de anticuerpos para su producción *in silico* de mAbs

Diagrama de Flujo ImmuneDiversity



Requerimientos para su instalación

GNU/Linux or MacOS :

- R \geq 3.0
- Python \geq 2.7 or Python 3
- Java \geq 1.6
- HMMER 3.0



Dependencias



- seqinr
- vegan
- ShortRead

- gplots
- Stringr
- Peptides



- iGraph 0.6.5
- Psutil
- pandas

Otros

- Blast 2.2.22 (incluido)
- acacia 1.52.b0 (incluido)
- igBlast (incluido)
- usearch \geq 6.1.544



Download

<https://bitbucket.org/ImmunediveRsity/immunediversity>



The screenshot shows the Bitbucket interface for the repository ImmunediveRsity/immunediversity. The top navigation bar includes the Bitbucket logo, 'Teams', 'Projects', 'Repositories', and 'Snippets' menus, along with a search bar and user profile. The left sidebar contains 'ACTIONS' (Clone, Create branch, Create pull request, Compare, Fork) and 'NAVIGATION' (Overview, Source, Commits, Branches, Pull requests, Wiki, Downloads). The main content area displays the repository overview with a table of statistics: Last updated (2015-07-23), Language (R), Access level (Write), 1 Branch, 0 Tags, 0 Forks, and 2 Watchers. Below this is the repository title 'ImmunediveRsity' and a description: 'Evaluation of the antibody repertoire by analyzing HTS data.' A 'Description:' section follows, explaining the project's goal. The 'Version' section shows '1.0.8'. On the right, the 'Recent activity' section lists four commits by Andres Aguilar, including a push to the main branch and two README updates.

Bitbucket Teams Projects Repositories Snippets Find a repository... ?

ImmunediveRsity / ImmunediveRsity

Overview

Last updated	2015-07-23	1	0
Language	R	Branch	Tags
Access level	Write (revoke)	0	2
		Forks	Watchers

Edit README

ImmunediveRsity

Evaluation of the antibody repertoire by analyzing HTS data.

Description:

Manipulation and processing of HTS reads to identify VDJ usage and clonal origin to gain insight of the antibody repertoire of a given organism.

Version

1.0.8

Recent activity

- 1 commit
Pushed to ImmunediveRsity/immunediversity
| 0927a69 Modificado el script find_CDR3
Andres Aguilar · 2015-07-23
- 1 commit
Pushed to ImmunediveRsity/immunediversity
| 8cf0210 README update
Andres Aguilar · 2015-06-02
- 2 commits
Pushed to ImmunediveRsity/immunediversity
| ad180d4 README update
| def8797 README update
Andres Aguilar · 2015-06-02
- 1 commit
Pushed to ImmunediveRsity/immunediversity
| 5a481a4 Modificada la version y actualizad...



Instalación

-Tiene su propio instalador “install.sh”

-Se necesitan permisos de SUDO

```
*****
                        ImmunediveRsity
Evaluation of the antibody repertoire by analyzing HTS data
*****

Verify prev installation...
ImmunediveRsity [OK]

Verifying dependencies...

Python [OK]
iGraph is installed [OK]
iGraph(python) [OK]
psutil is installed [OK]
psutil(python) [OK]
> source('installer/dependency_checker.R') [OK]
R libraries [OK]
HMMR [OK]
INTR0 TO CONTINUE...

root password (sudo)
Starting installation process

Creating ImmunediveRsity directory [OK]
Moving uninstall.sh [OK]
Creating /opt/ImmunediveRsity/uninstaller [OK]
Moving uninstaller scripts [OK]
Moving COPYING to /opt/ImmunediveRsity [OK]
Moving README.md to /opt/ImmunediveRsity [OK]
Creating ImmunediveRsity/bin directory [OK]
Moving bin/ImmunediveRsity [OK]
Moving bin/ImmunediveRsity_notifier [OK]
Moving bin/ImmunediveRsity_subset [OK]
Moving /bin/acacia [OK]
Moving bin/cutVDJ.pl [OK]
Moving bin/parseKaks [OK]
Moving bin/network_graph [OK]
Moving bin/complete_table [OK]
Moving bin/div_table [OK]
Moving bin/usearch [OK]
Moving bin/parseKaks_subset [OK]
Creating blast-2.2.22 directory [OK]
Creating bin/blast-2.2.22/bin directory [OK]
```





¿qué datos necesitamos para correrlo?

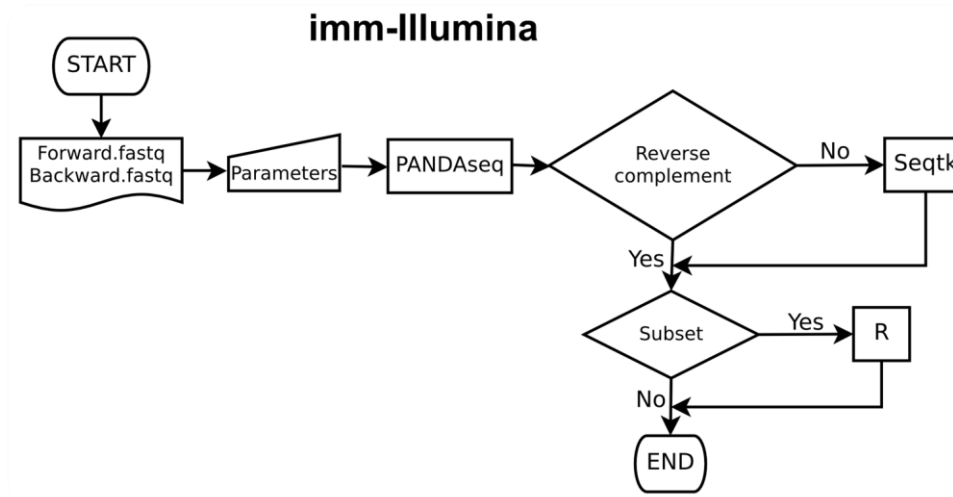
Secuencias de Inmunoglobulinas

ROCHE 454

- Carpeta con los archivos Fastq que se van a procesar
- En reverso complementario
- Longitud > 300pb

Illumina

- Correr script para illumina
- Carpeta con los archivos Fastq que se van a procesar
- En reverso complementario
- Longitud > 300pb





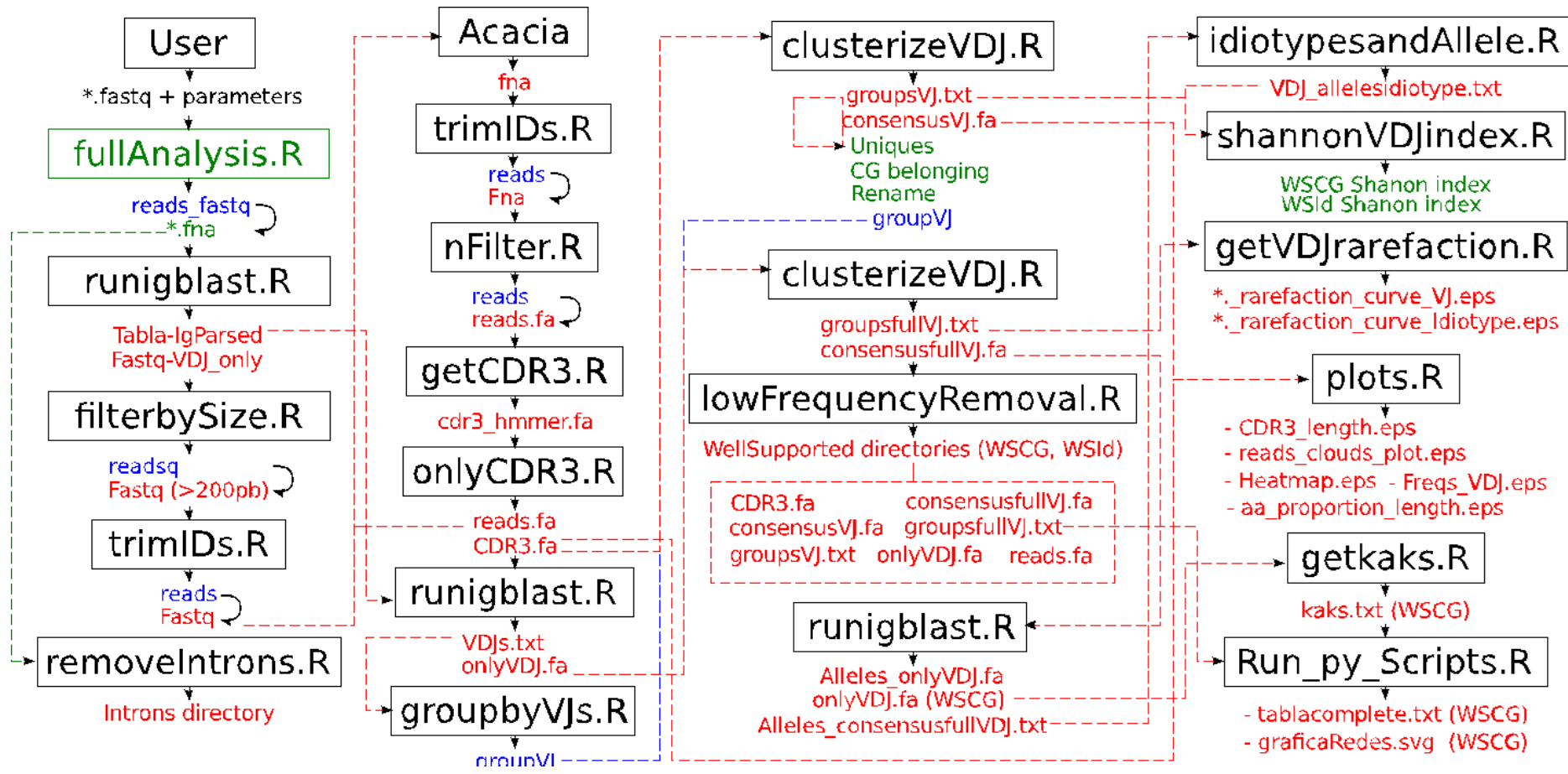
¿Cómo correrlo?

```
~ % ImmunediveRsity -h
usage: ImmunediveRsity [-h] -i INPUT -o OUTPUT [-c CORES] [-l READLENGTH]
                        [-s SPECIES] [-d ID] [-q QFRACT] [-t TFRACT] [-a] [-r]
                        [-cc CGFREQ_CUT] [-ic IFREQ_CUT] [-n NOTIFY] [-m] [-p]
                        [-mq MEAN_QUALITY] [-v]
```

```
test % ImmunediveRsity -i /Users/TinaGodoy/Desktop/test/fastq -o /Users/TinaGodoy/Desktop/test/results -c 3 -p -r
Searching in /Users/TinaGodoy/Desktop/test/fastq
>Library den1Hre_i10 found!

Creating dirs in /Users/TinaGodoy/Desktop/test/results
>Dir den1Hre_i10 created!
  config.out created!
Runner.sh done!!!
test % █
```

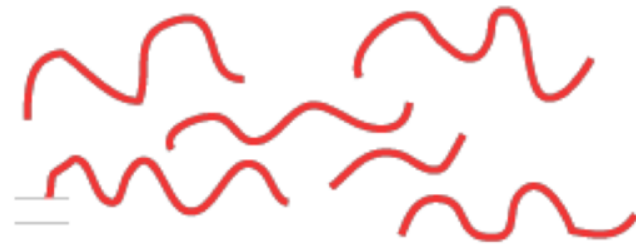
¿Qué hace mientras corre?



Transcriptomics

The study of the complete set of RNAs (**transcriptome**) encoded by the genome of a specific cell or organism at a specific time or under a specific set of conditions

Isolate total RNA



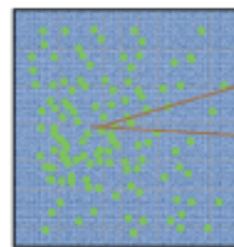
Convert to double stranded DNA



Hybridize to tiling array

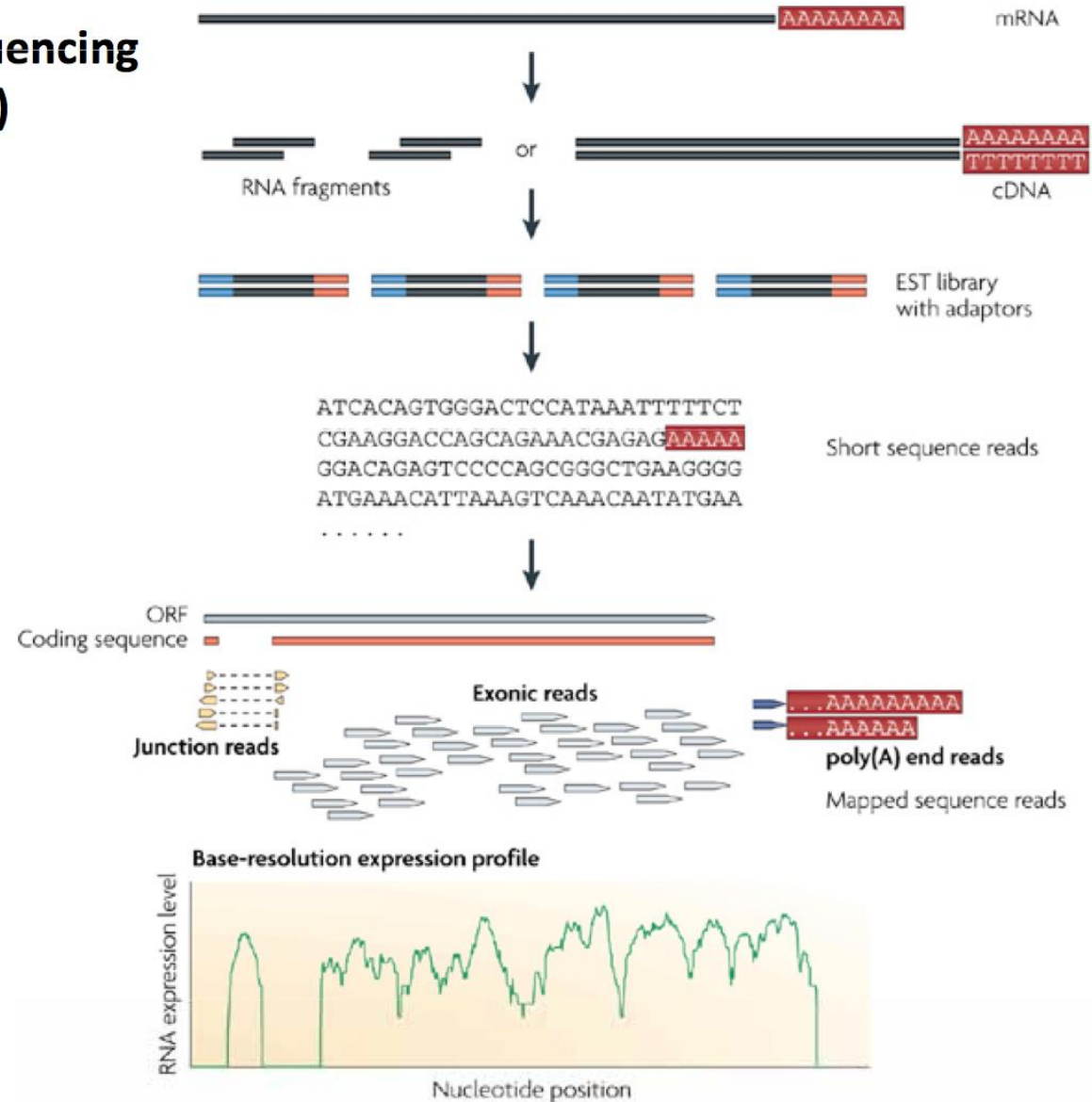


Analyze gene expression



Experimental Approaches Used to Characterize the Human Transcriptome

RNA Sequencing (RNA Seq)

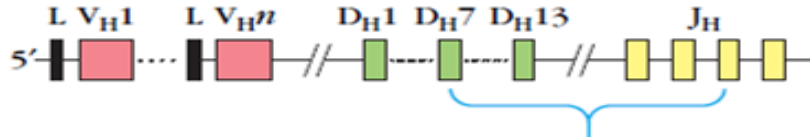


Results

Linfocito B

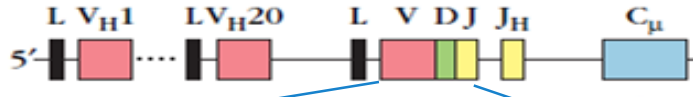
Línea germinal

DNA de
cadena H
de la
línea germinal



Recombinación somática

DNA de
cadena H
reordenado



CDR3



In silico

Base de datos

- 40 segmentos V [239 alelos]
- 25 segmentos D [30 alelos]
- 6 segmentos J [13 alelos]

Asignación de segmentos

6,000 rearrreglos VDJ
>93,000 rearrreglos con alelos
Diversidad en el CDR3: $\sim 2 \times 10^7$
Diversidad por individuo: $\sim 1.8 \times 10^{11}$
Diversidad poblacional: $\sim 2.8 \times 10^{12}$

The Immune System, 3ed, 2009



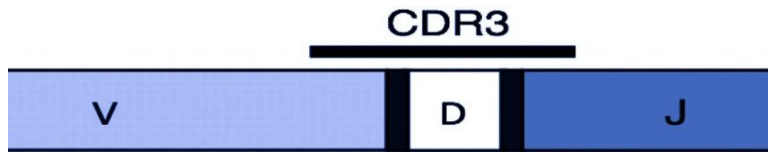
V	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHV5-51*01	91.53	295	24	1	1	294	1	295	6e-107	377
V	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHV5-51*03	91.16	294	25	1	1	293	1	294	1e-105	372
V	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHV5-51*02	90.85	295	26	1	1	294	1	295	4e-105	371
D	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHD1-26*01	100.00	7	0	0	301	307	4	10	8.4	14.4
D	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHD2-15*01	100.00	7	0	0	302	308	10	16	8.4	14.4
D	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHD3-22*01	100.00	7	0	0	302	308	16	22	8.4	14.4
J	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHJ3*02	92.00	50	4	0	314	363	1	50	1e-13	60.3
J	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHJ3*01	90.00	50	5	0	314	363	1	50	2e-10	50.3
J	reversed HA7IYCV01.IGHV5-51.IGHJ3.4.1	IGHJ4*03	96.00	25	1	0	330	354	15	39	2e-07	40.4

+

Identificación del CDRH3 (origen clonal)

Huella digital de los anticuerpos

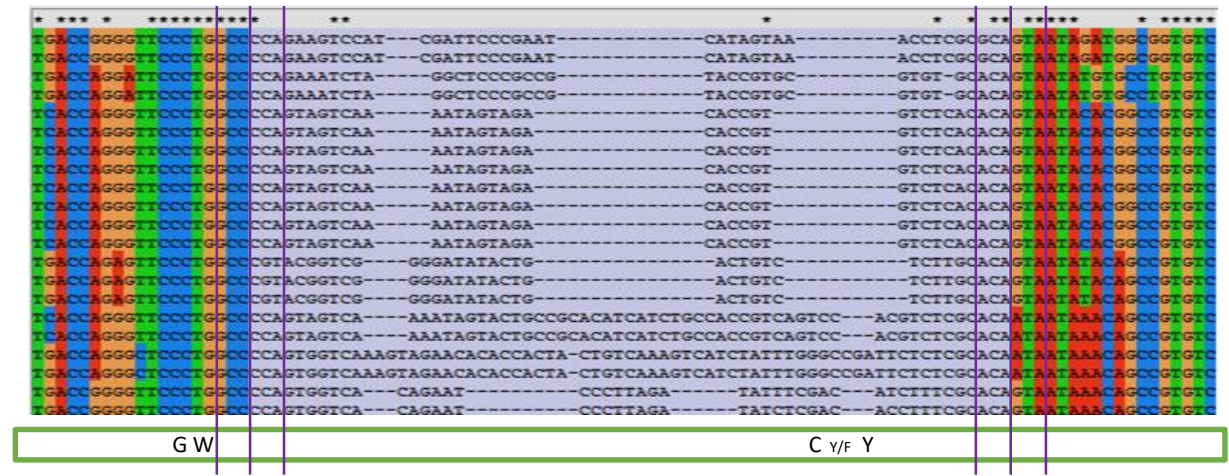
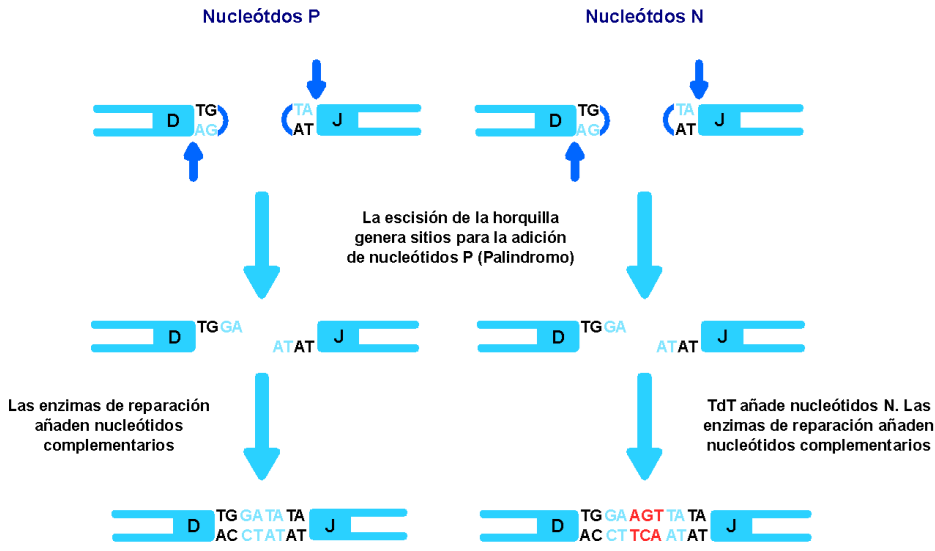
CDRH3



Identificación del CDRH3



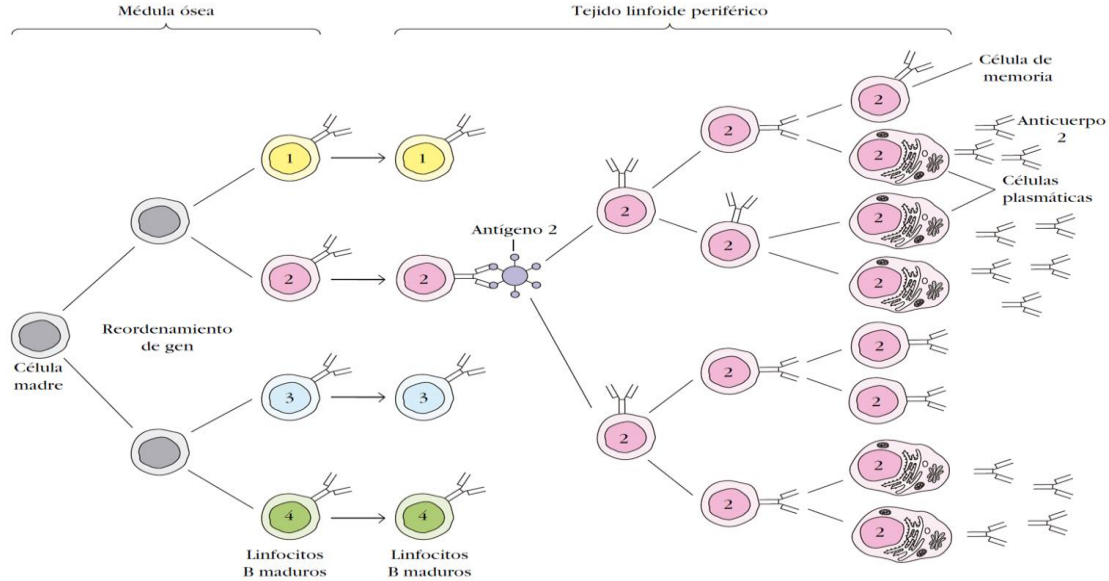
- Perfil con rearrreglos de la línea germinal
- 3 rondas de entrenamiento con secuencias verdaderas





Identificación de los Grupos Clonales

La expansión clonal



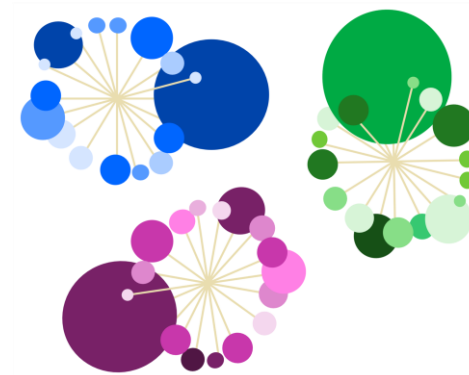
Agrupación por Origen Clonal

Mismo segmento
VH

Mismo segmento
JH

Misma longitud
en el CDRH3

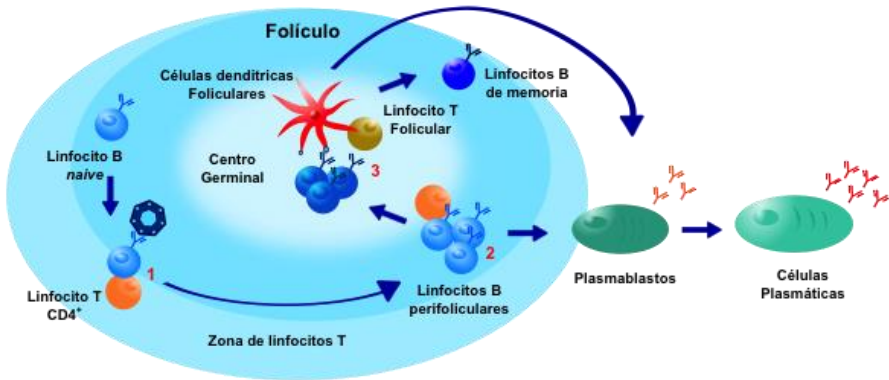
Identidad 97% en
el CDRH3



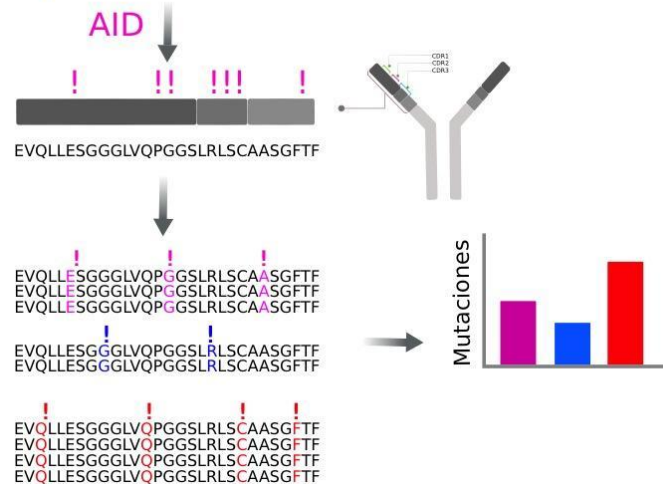


Identificación de linajes VH

La maduración de afinidad



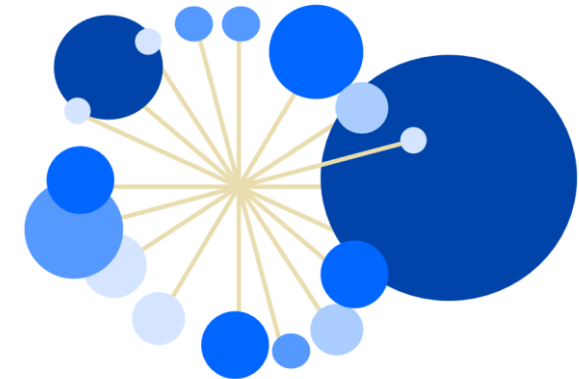
Hipermutación Somática



Identificación de linajes VH

Mismo Grupo Clonal

Identidad 98.5%





¿De donde sacaron los cortes?

- Se hizo la estandarización con tres sets de datos:
 - **Ratón transgénico MD4.** IGHVV4-3*01- IGHD-4-1*01-IGHJ2*01
 - **IGHV1-3.** Set humano curado a mano
 - **Stanford22.** Secuencias no relacionadas clonalmente.

Table 1. Overview of the reference sequencing sets

Set	Sequenced reads	After filters ¹	Observed clonotypes	Expected clonotypes	Well supported clonotypes ²	Observed lineages (without singletons)	Expected lineages	Well supported lineages ³
MD4	5,359	99.6%	10	1	1	21(7)	1	1
IGHV1-3	1,044	95.2%	1	1	1	469 (52)	10	7
Stanford22 ⁴	13,141	100%	11,779	13,141	NA	12,421	13,141	NA

¹Percent of reads that pass the pre-processing filters.

²Number of clonotypes whose corresponding lineages are composed ≥ 6 reads.

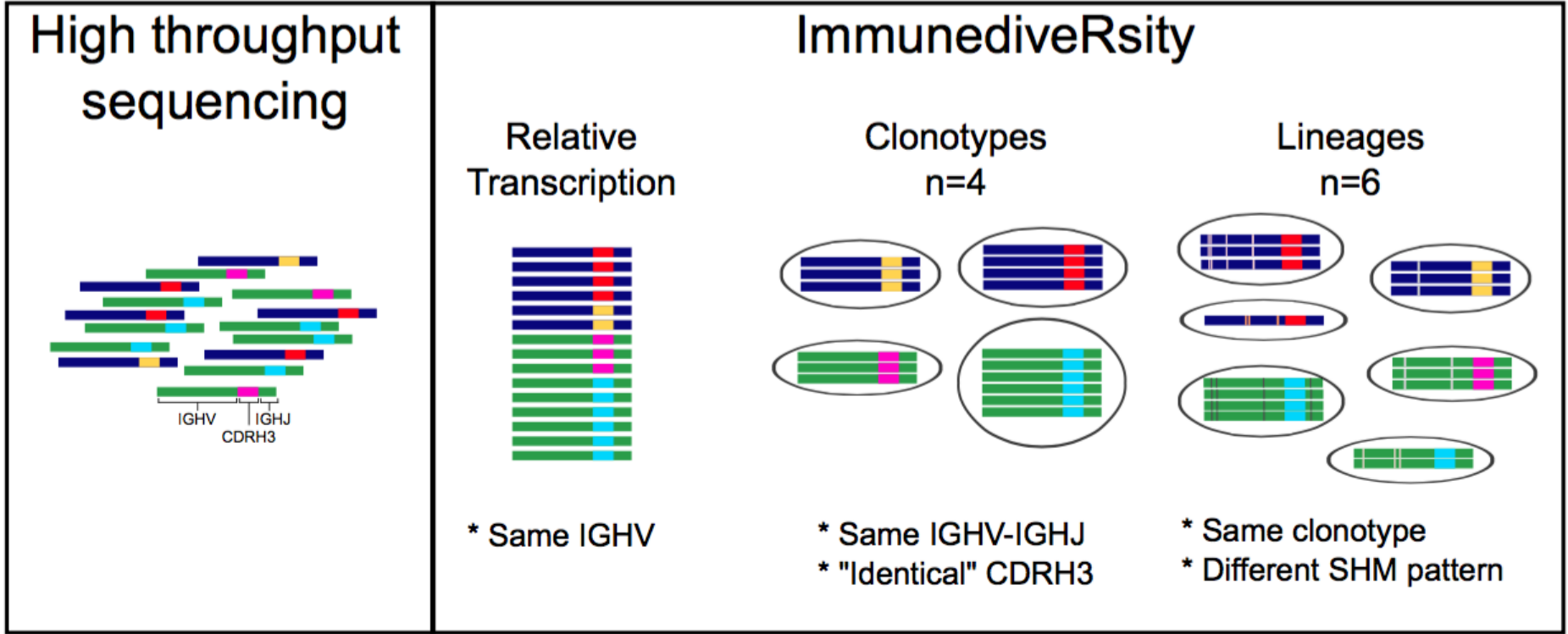
³Lineages composed of ≥ 6 reads.

⁴The publicly available Stanford22 set was published as a set of non-clonally related immunoglobulin sequences; ²⁸ we removed one read with a duplicated identifier and 11 with duplicated sequences.

NA, not applicable.



¿Qué información obtenemos?





Output

Archivos de texto -> Tablas

indices_table_unicos.txt

1	CG_ID	No_of_id	No_of_reads	reads_norm	Gini	Shannon	Shannon_norm	Shannon_pon	Shannon_Var	sm	nsm	Id_ID	VDJ	VDJ_AA	CDR3	CDR3_AA	CDR3_AA_length	aindex	hindex	bomanindex	charge	GRAVY
2	den1Cag_i04.IGHV1-18.IGHJ2.5	1	1	0.0482625482625	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-18.IGHJ2.5.0	CAGGTTCAACTGTG	5.0.0								
3	den1Cag_i04.IGHV1-18.IGHJ4.20	1	2	0.0965250965251	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-18.IGHJ4.20.0	CAGGTTCACTGGT	5.0.0								
4	den1Cag_i04.IGHV1-18.IGHJ4.21	1	2	0.0965250965251	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-18.IGHJ4.21.0	CAGGTTCACTGGT	5.0.0								
5	den1Cag_i04.IGHV1-18.IGHJ4.4	1	4	0.19305019305	0	0	0	0	0	0	7	12	den1Cag_i04.IGHV1-18.IGHJ4.4.0	CAGGTTCACTGGT	5.0.0							
6	den1Cag_i04.IGHV1-18.IGHJ4.40	1	1	0.0482625482625	0	0	0	0	0	0	7	14	den1Cag_i04.IGHV1-18.IGHJ4.40.0	CAGGTTCAATTAGT	5.0.0							
7	den1Cag_i04.IGHV1-18.IGHJ6.1	33	59	2.84749034749	0.291730868002	1.65610510335	0.406153254541	1.45316504411	0.0096057													
8	den1Cag_i04.IGHV1-18.IGHJ6.24	2	2	0.0965250965251	0	0	0	0.301029995664	0	0	0	den1Cag_i04.IGHV1-18.IGHJ6.24.0	CAGGTTCACTGGT	5.0.0								
9	den1Cag_i04.IGHV1-18.IGHJ6.38	3	3	0.144787644788	-5.55111512313e-17	0	0	0.47712125472	0	9	13	den1Cag_i04.IGHV1-18.IGHJ6.38.0	CAGGTTCACTGGT	5.0.0								
10	den1Cag_i04.IGHV1-18.IGHJ6.40	1	1	0.0482625482625	0	0	0	0	0	8	15	den1Cag_i04.IGHV1-18.IGHJ6.40.0	CAGGTTCACTGGT	5.0.0								
11	den1Cag_i04.IGHV1-18.IGHJ6.55	1	1	0.0482625482625	0	0	0	0	0	4	6	den1Cag_i04.IGHV1-18.IGHJ6.55.0	CAGGTTCACTGGT	5.0.0								
12	den1Cag_i04.IGHV1-18.IGHJ6.65	1	1	0.0482625482625	0	0	0	0	0	4	7	den1Cag_i04.IGHV1-18.IGHJ6.65.0	CAGGTTCACTGGT	5.0.0								
13	den1Cag_i04.IGHV1-18.IGHJ6.74	1	1	0.0482625482625	0	0	0	0	0	7	12	den1Cag_i04.IGHV1-18.IGHJ6.74.0	CAGGTTCACTGGT	5.0.0								
14	den1Cag_i04.IGHV1-18.IGHJ6.76	1	1	0.0482625482625	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-18.IGHJ6.76.0	CAGGTTCACTGGT	5.0.0								
15	den1Cag_i04.IGHV1-2.IGHJ6.15	3	3	0.144787644788	-5.55111512313e-17	0	0	0.47712125472	0	0	0	den1Cag_i04.IGHV1-2.IGHJ6.15.0	CAGGTTCACTGGT	5.0.0								
16	den1Cag_i04.IGHV1-2.IGHJ6.24	1	1	0.0482625482625	0	0	0	0	0	10	15	den1Cag_i04.IGHV1-2.IGHJ6.24.0	CAGGTTCACTGGT	5.0.0								
17	den1Cag_i04.IGHV1-2.IGHJ6.25	1	1	0.0482625482625	0	0	0	0	0	13	14	den1Cag_i04.IGHV1-2.IGHJ6.25.0	CAGGTTCACTGGT	5.0.0								
18	den1Cag_i04.IGHV1-24.IGHJ4.3	1	1	0.0482625482625	0	0	0	0	0	1	31	den1Cag_i04.IGHV1-24.IGHJ4.3.0	CAGGTTCACTGGT	5.0.0								
19	den1Cag_i04.IGHV1-3.IGHJ4.1	2	0	0.0965250965251	0	0	0	0	0	9	10	den1Cag_i04.IGHV1-3.IGHJ4.1.0	CAGGTTCACTGGT	5.0.0								
20	den1Cag_i04.IGHV1-3.IGHJ6.10	1	1	0.0482625482625	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-3.IGHJ6.10.0	CAGGTTCACTGGT	5.0.0								
21	den1Cag_i04.IGHV1-3.IGHJ6.10	1	1	0.0482625482625	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-3.IGHJ6.10.0	CAGGTTCACTGGT	5.0.0								
22	den1Cag_i04.IGHV1-3.IGHJ6.10	1	1	0.0482625482625	0	0	0	0	0	0	0	den1Cag_i04.IGHV1-3.IGHJ6.10.0	CAGGTTCACTGGT	5.0.0								
20	Ctrl-06.IGHV1-2.IGHJ3.1.2																					



VDJs.txt



fasta y fastq

exec

Alleles_onlyVDJ.fa

exec

CDR3.fa

exec

config.out

exec

consensusfullVJ.fa

exec

consensusVJ.fa

exec

den1Cag_i04_filtere
d.fastq

exec

den1Cag_i04_VDJ.f
astq

exec

den1Cag_i04.fastq

exec

den1Cag_i04.fna

exec

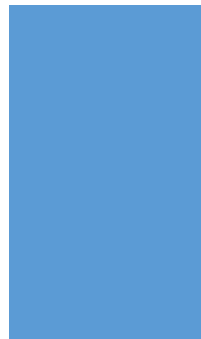
log.out

exec

onlyVDJ.fa

exec

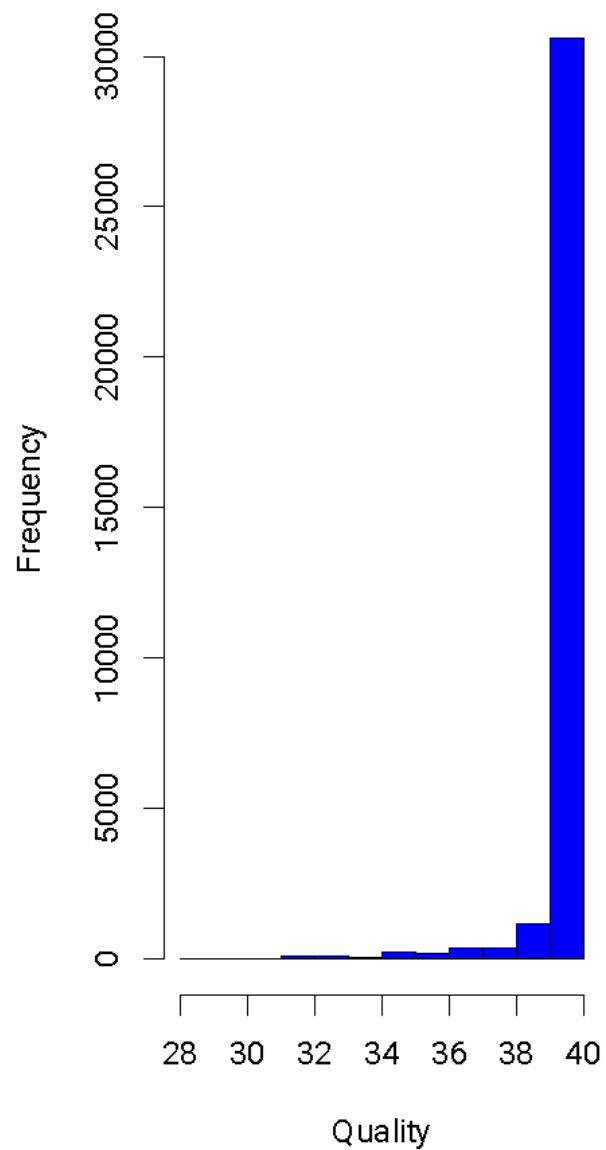
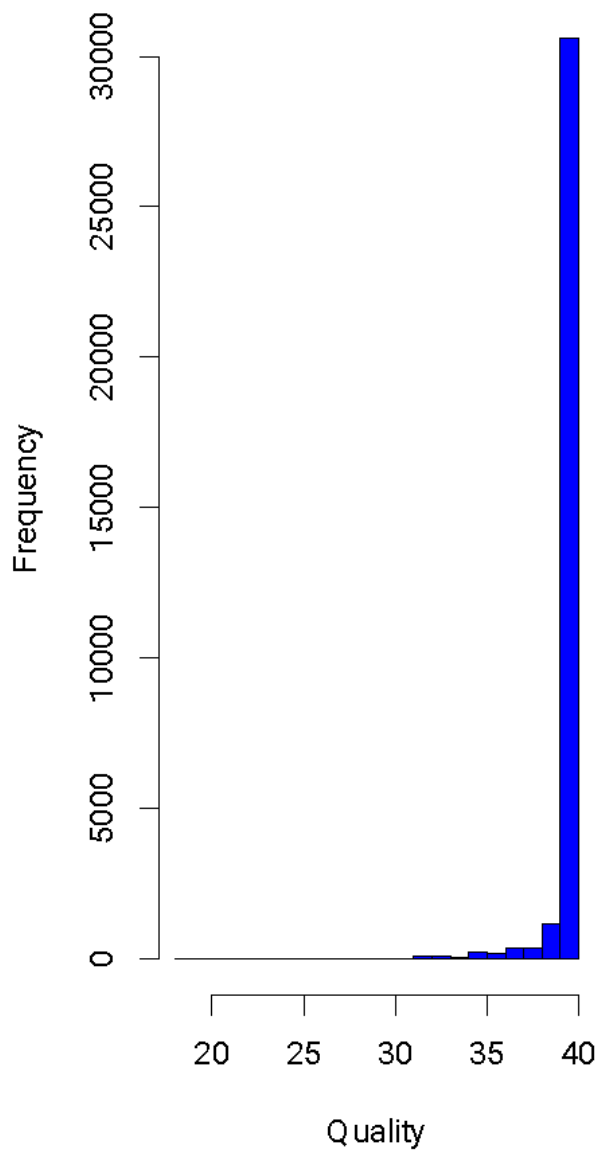
reads.fa





Gráficas

Unfiltered reads median quality per read Filtered reads median quality per read

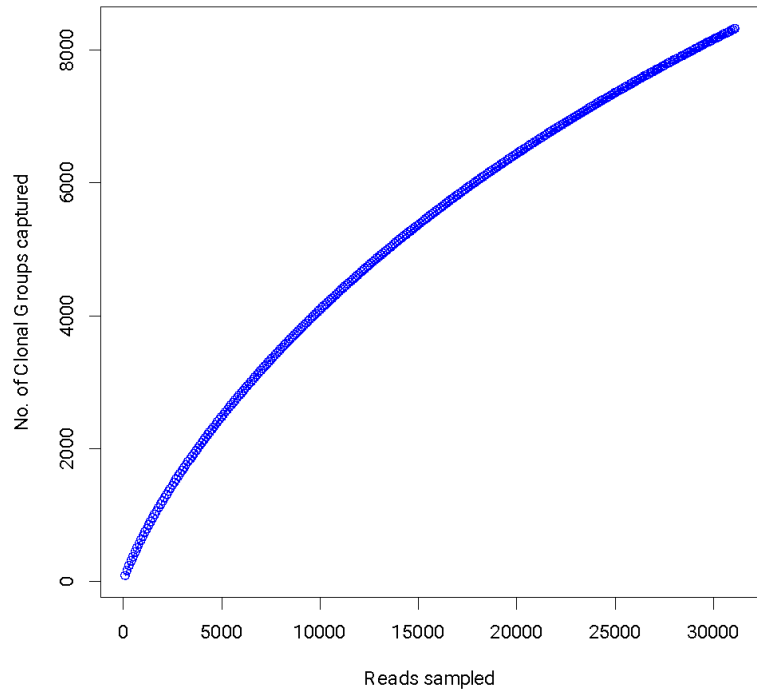




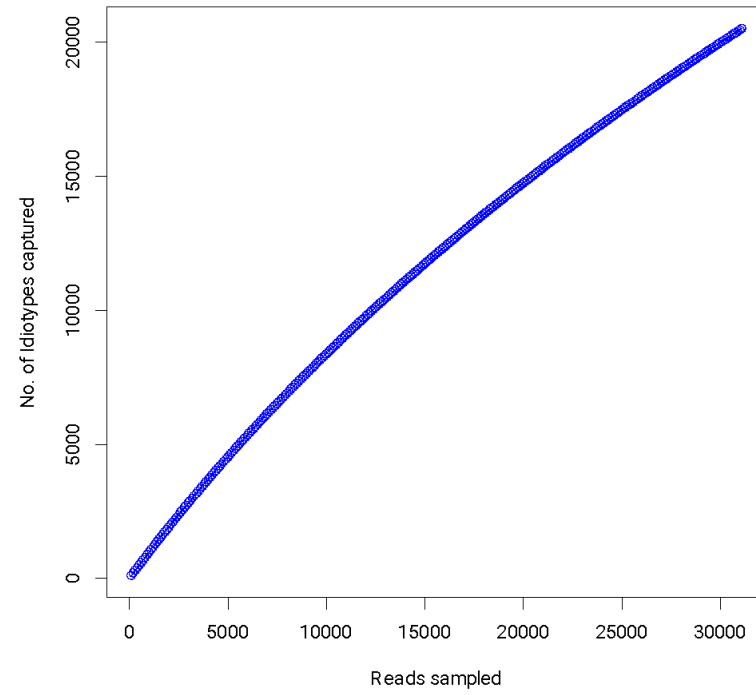
Gráficas



Clonal Group Rarefaction Curve
den1Cag_i04

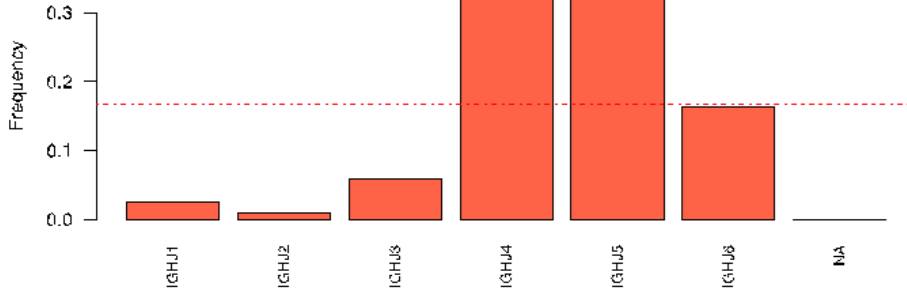
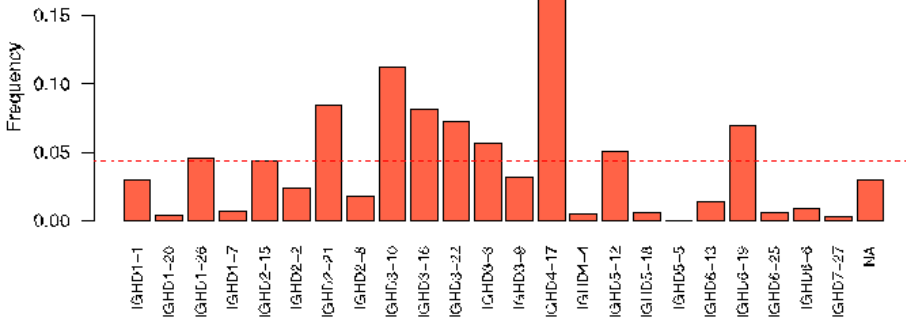
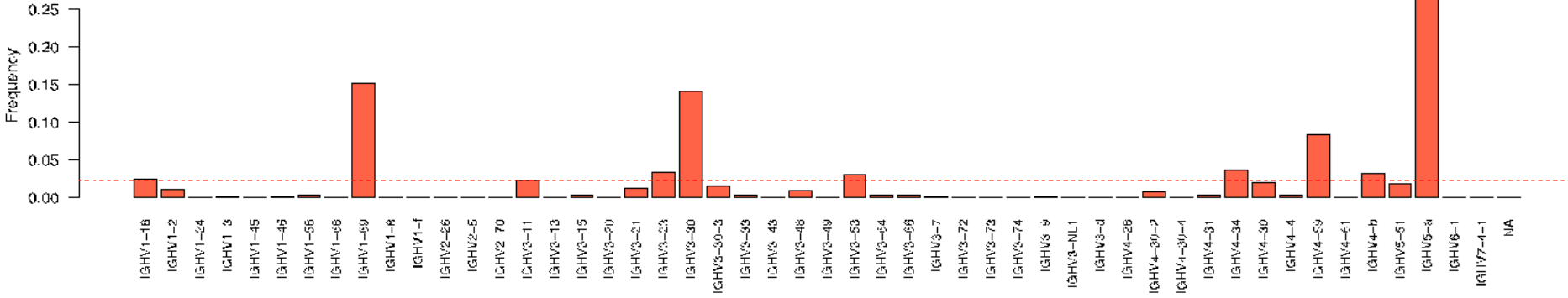


Idiotype Rarefaction Curve
den1Cag_i04



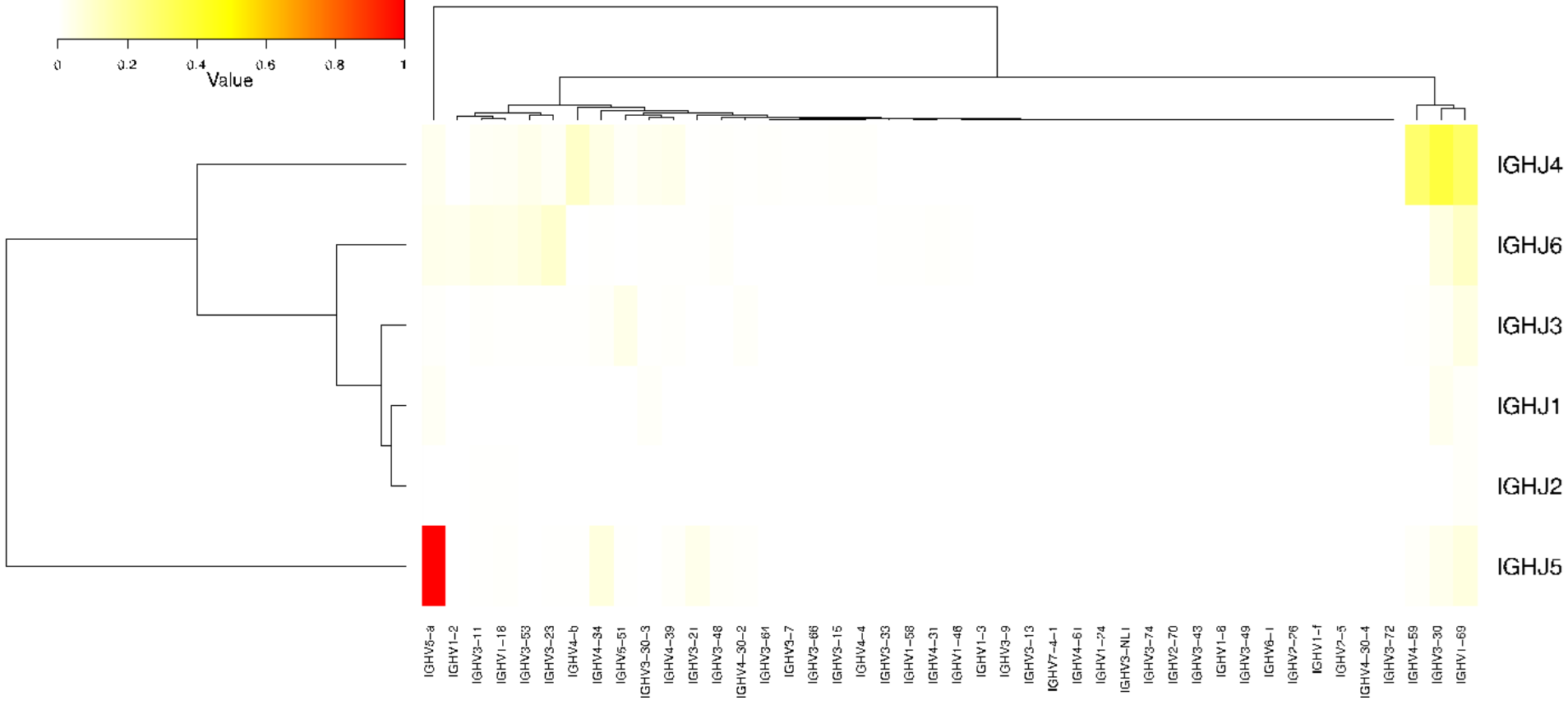
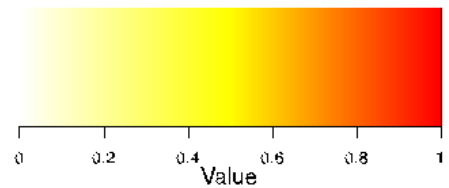


Gráficas





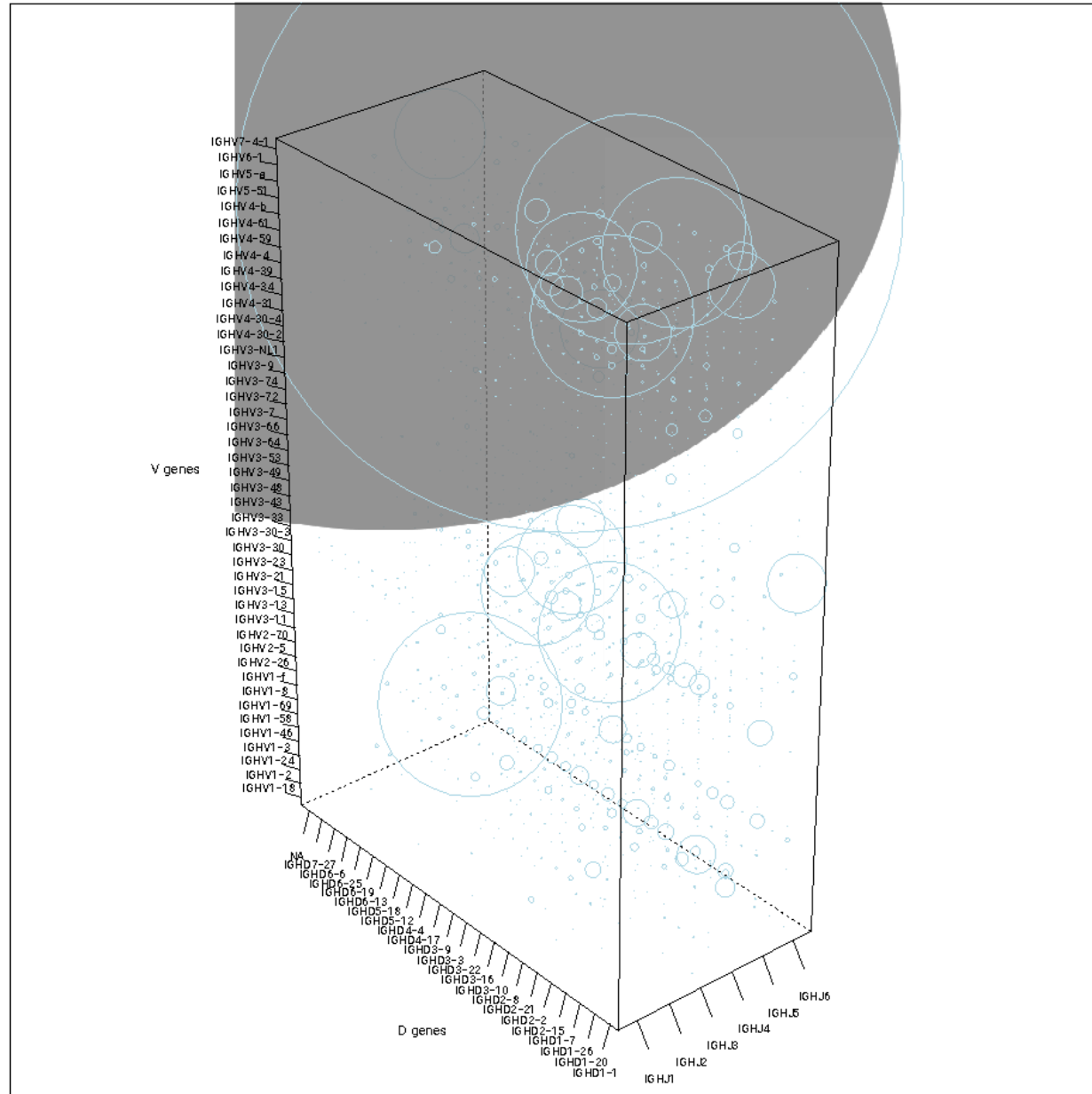
Gráficas





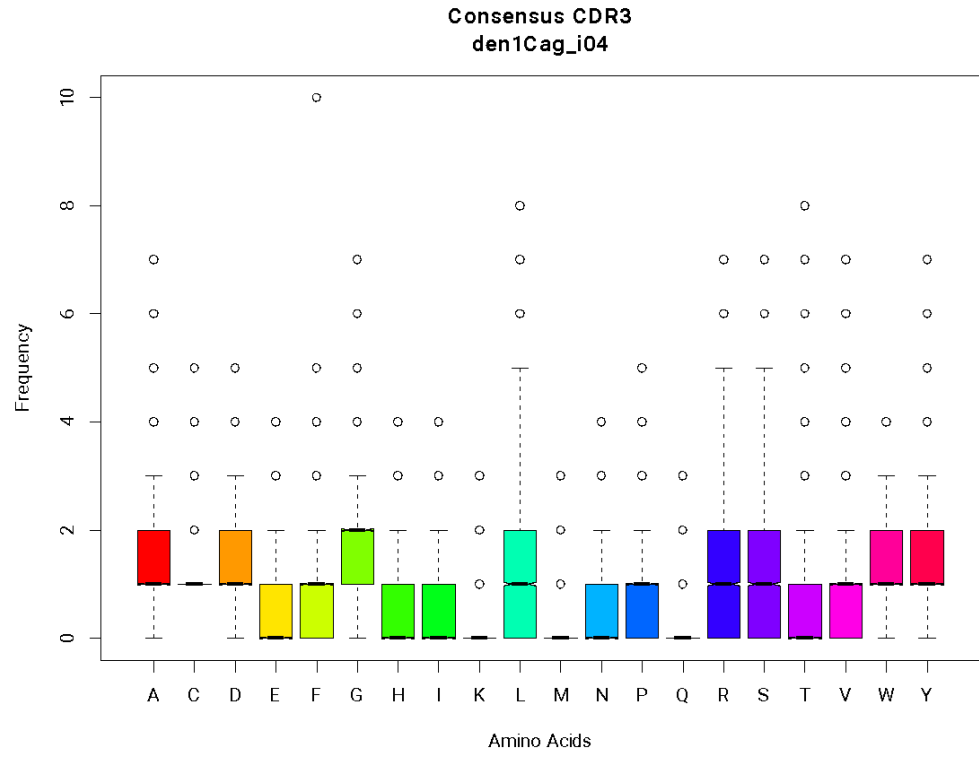
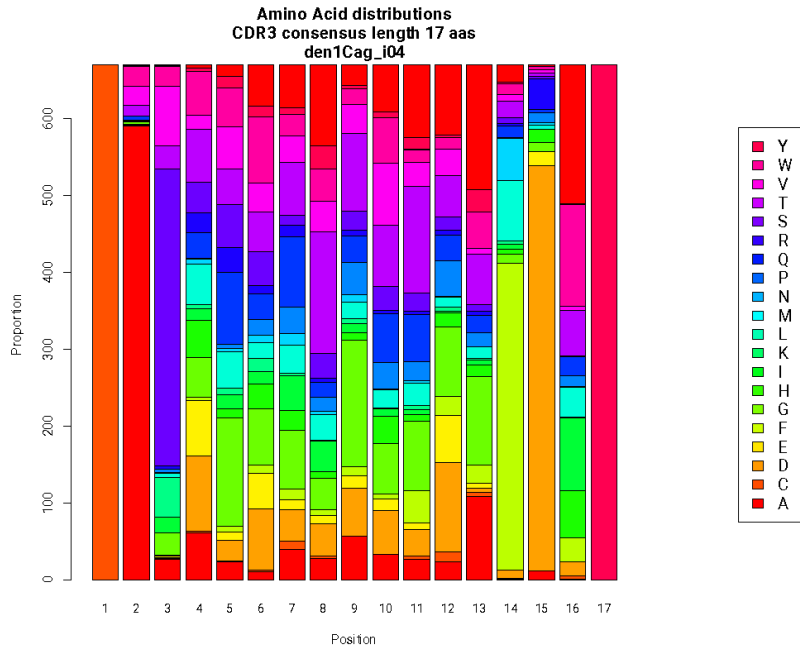
Gráficas

VD J rearrangements
den1Cag_i04



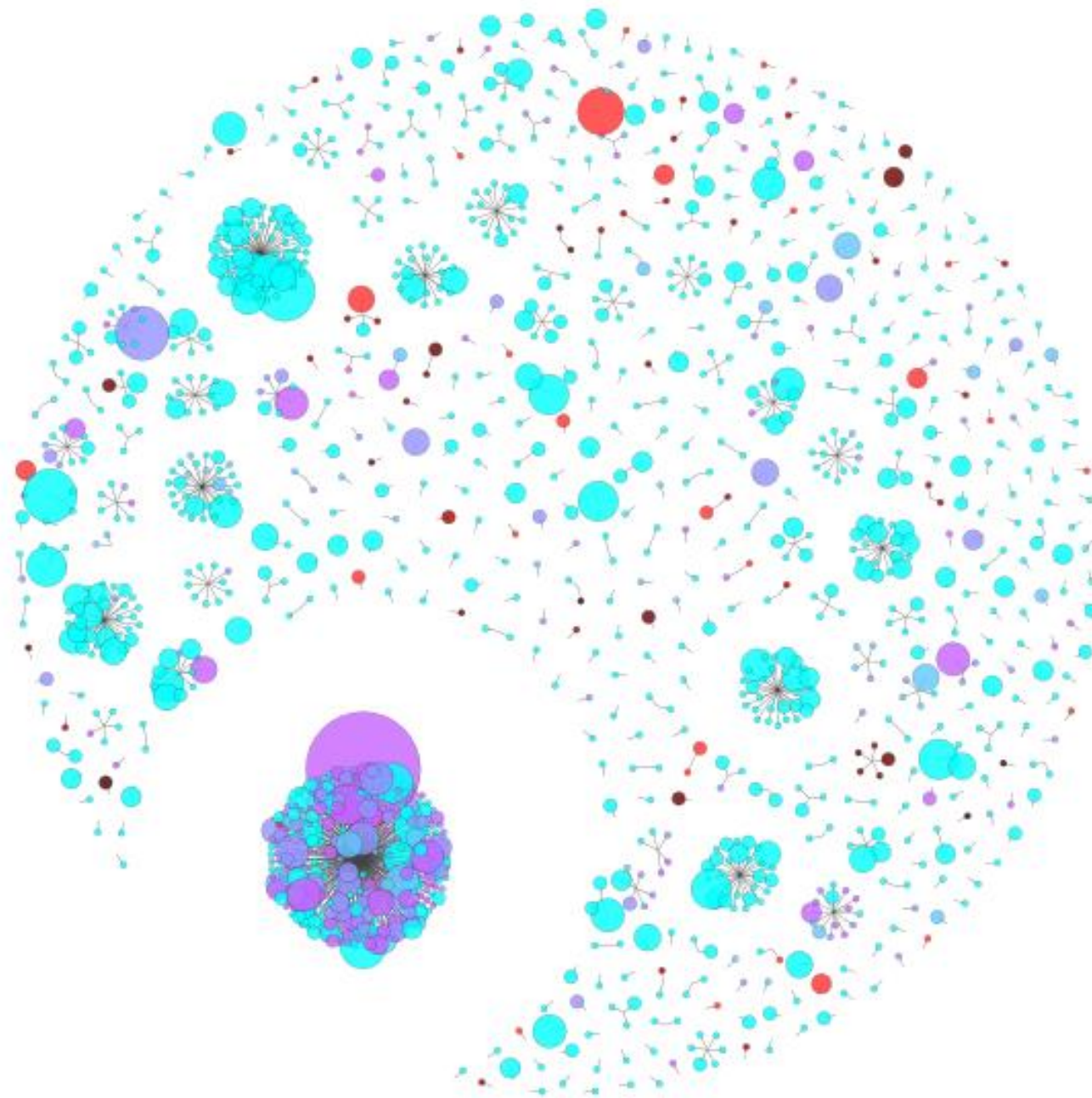
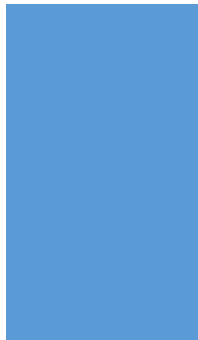


Gráficas



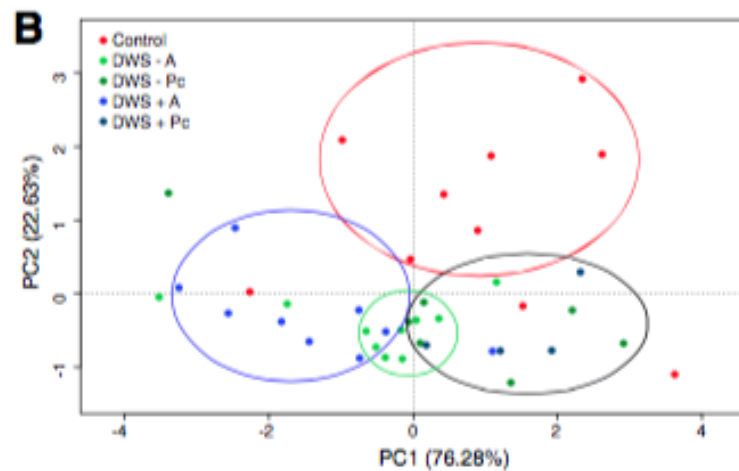
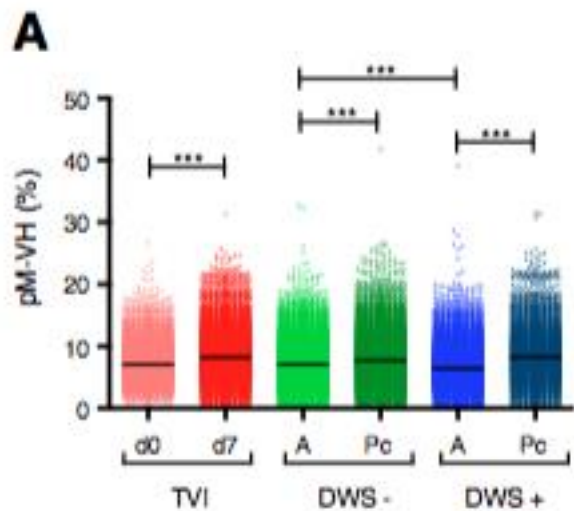
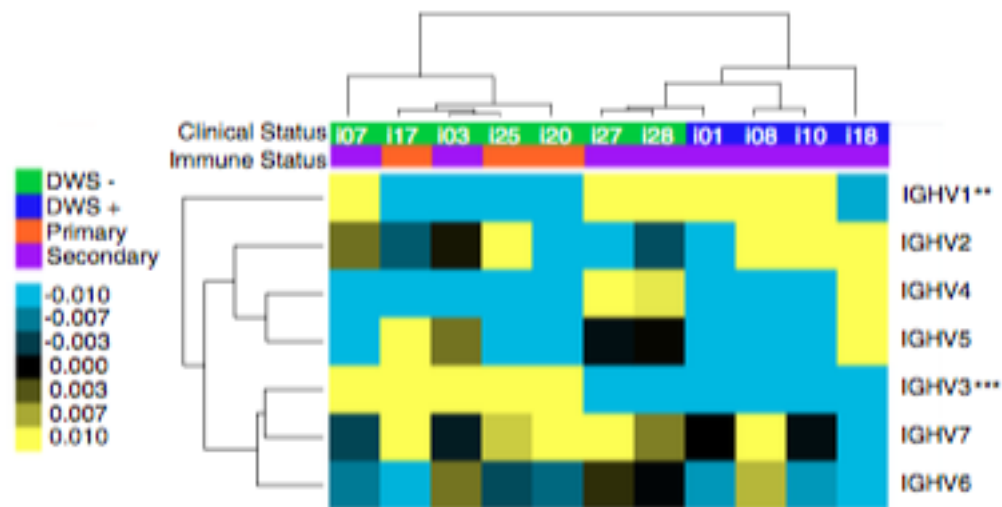
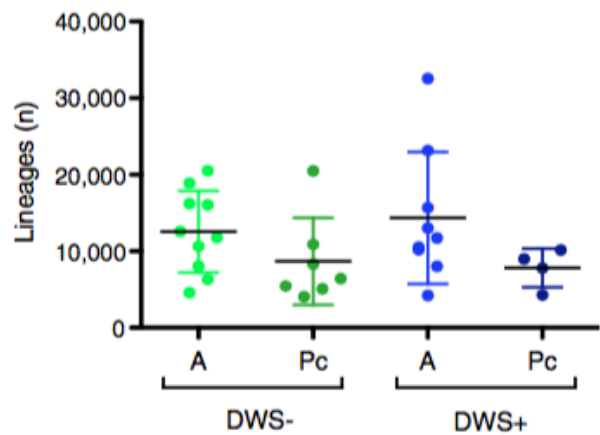


Gráficas



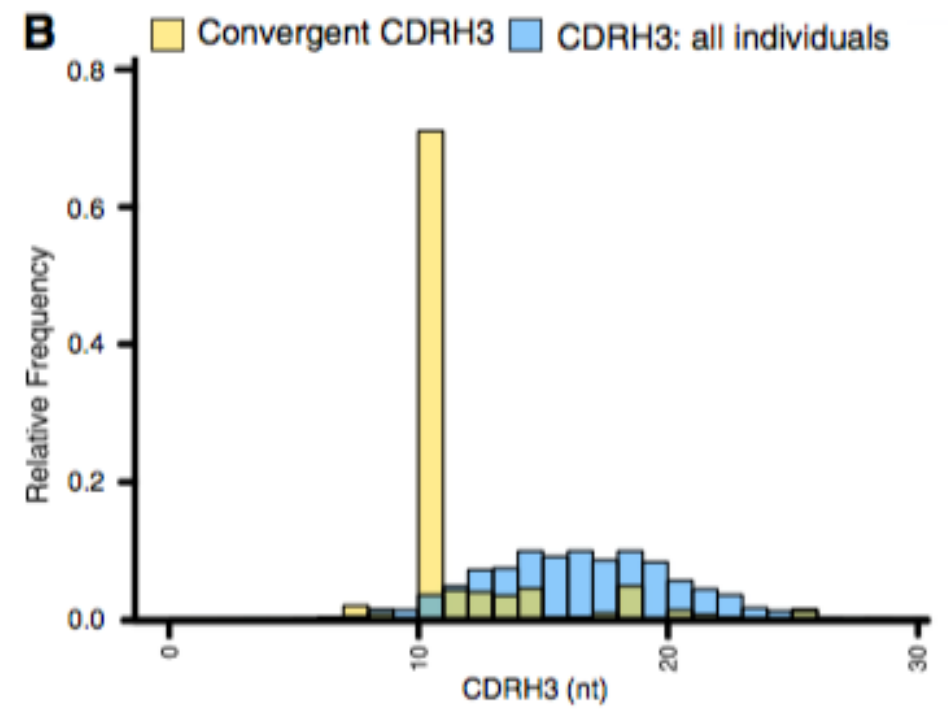
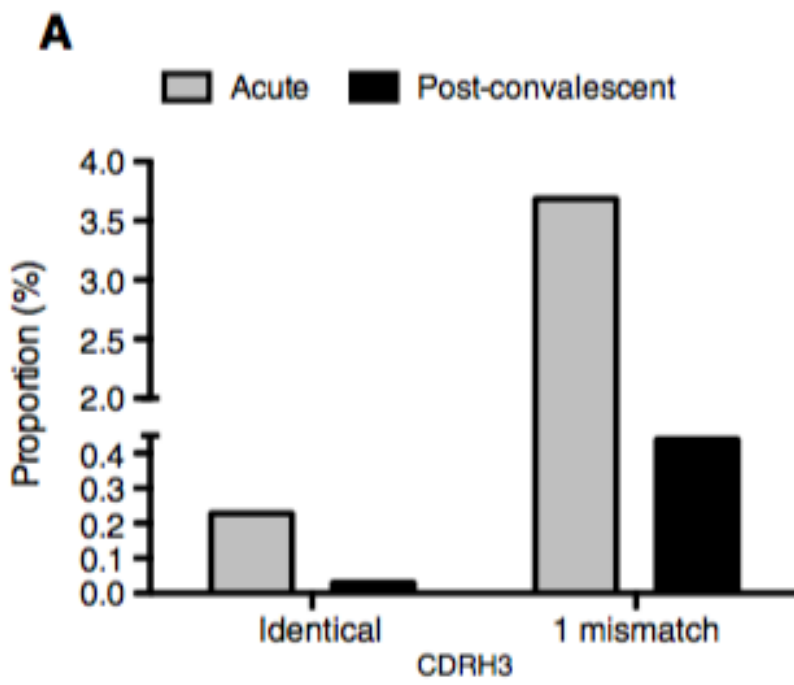


Comparaciones entre diferentes genotecas





Comparaciones entre diferentes genotecas



Modelo de estudio: intestino de pez lacustre

**Importancia ecológica:
Consumidor secundario
(Zooplanctófago)**

**Modelo sin estudio:
Intestino corto sin estómago**

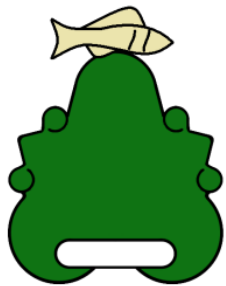
**Importancia nutricional:
30% de DHA de todas las grasas que
posee**



Chirostoma estor

OMEGA 3 FACTORY

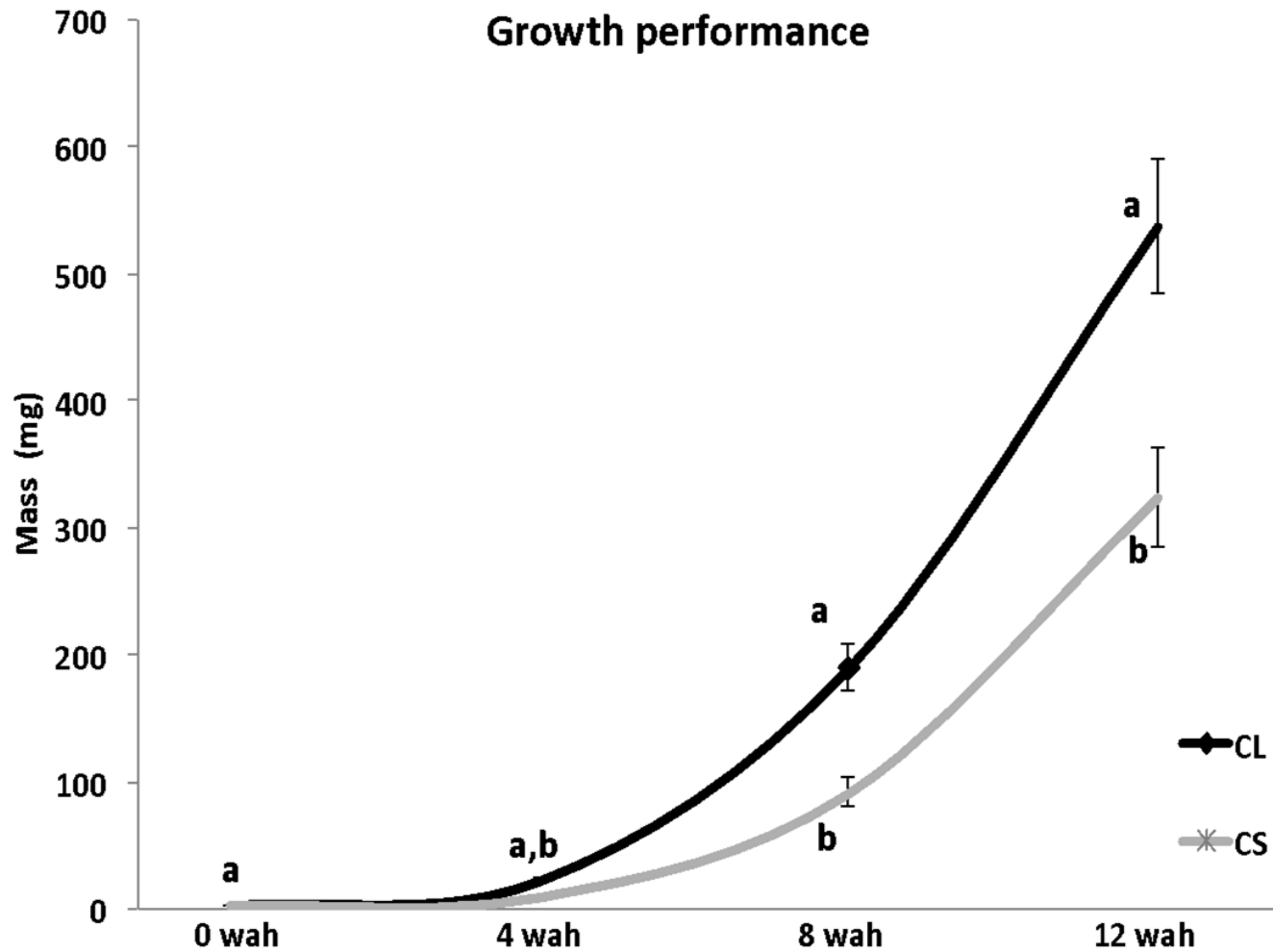
	%EPA	%DHA	Ratio DHA/EPA
Salmon	7.1	15.7	2.2
Tuna	6	22	3.7
Liver Cod oil	8	11	1.4
Pike silverside	2	29.2	14.2



Cultural:
Emblema de la región
Purépecha

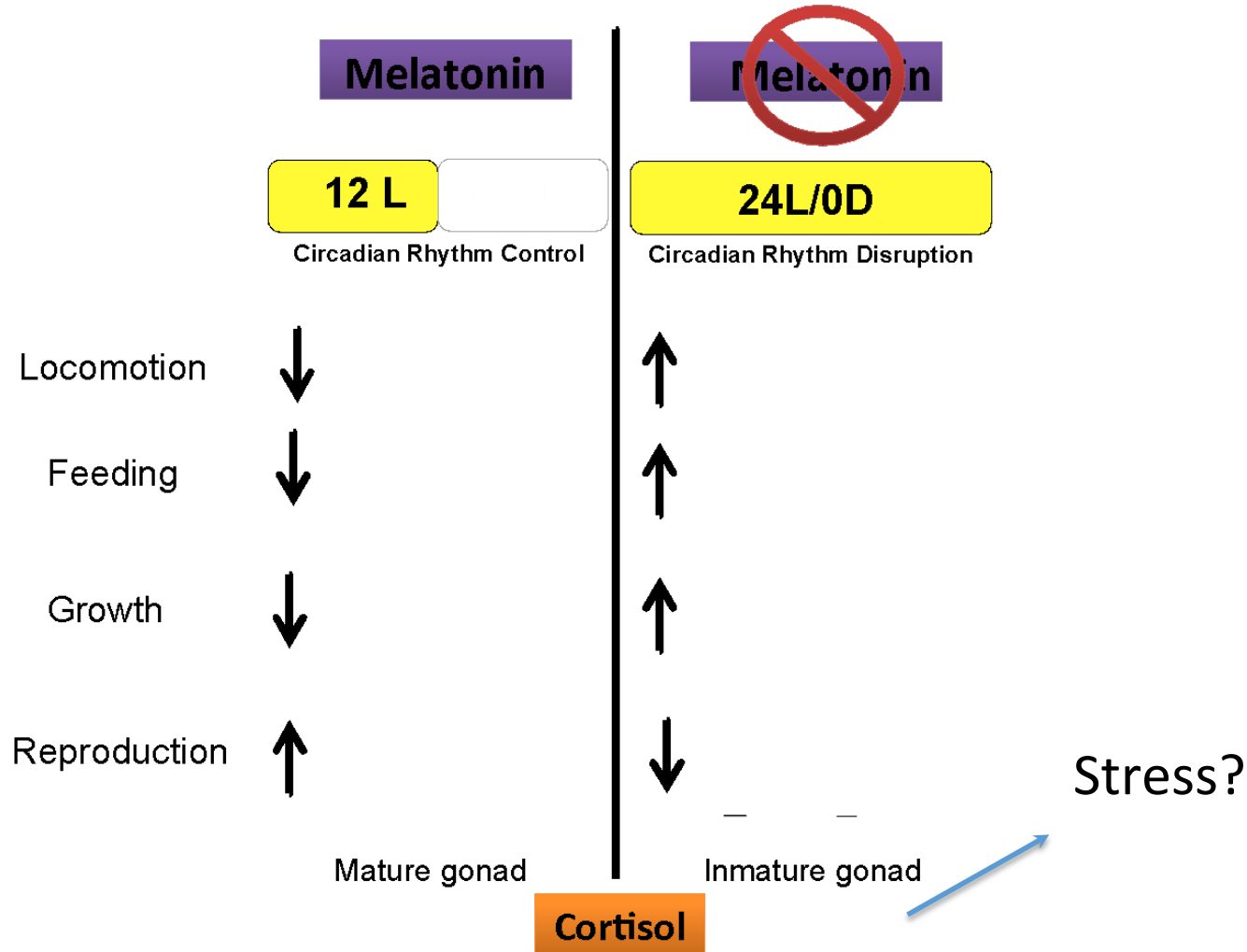
**Alto valor comercial:
\$200-800/kg**

Fonseca-Madrugal *et al.*, (2014)
Journal of lipid research



- Nutrition?
- Space?
- Stress?
- Immune system?
- Microbiota health?

Circadian light disruption in fish to promote growth



Photoperiod effects positive ✓

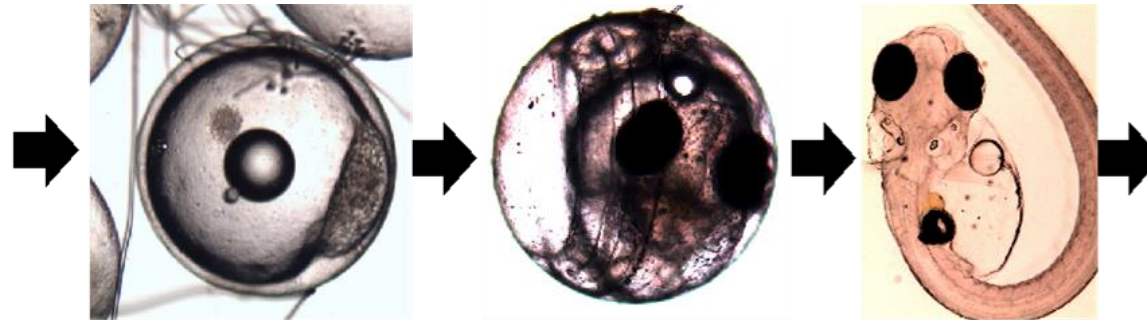
Mechanisms not fully understood X

What are the source mechanisms behind photoperiod effects?

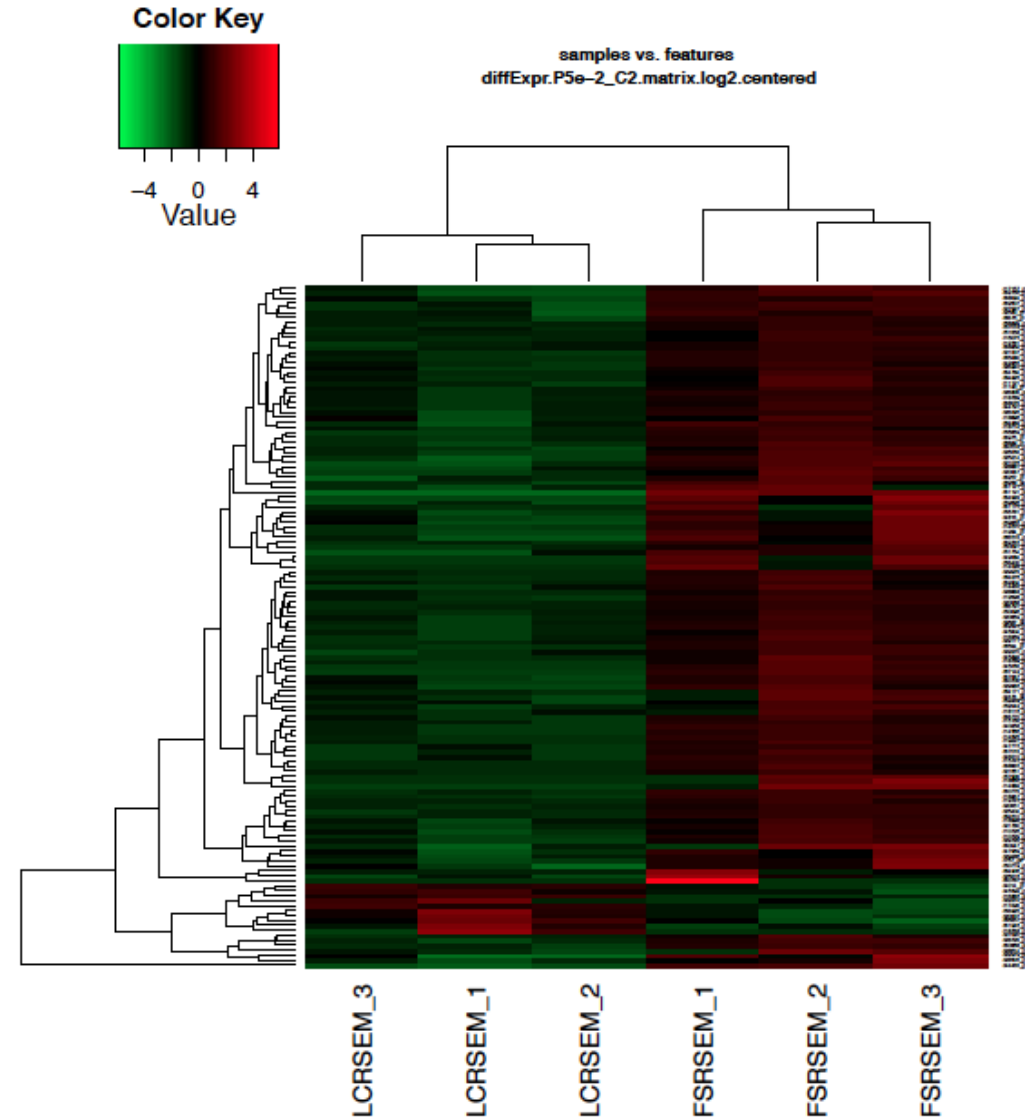
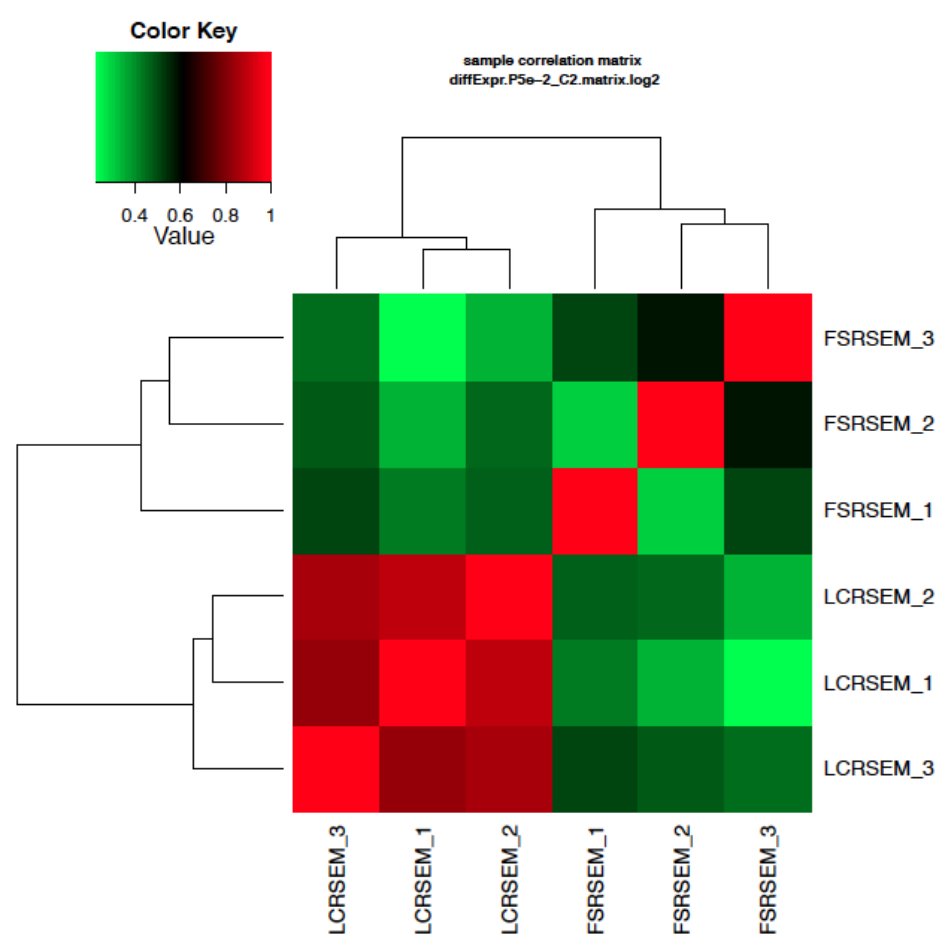
Integrative study

“What does not kill you, makes you stronger” – F. Nietzsche: Hormesis

De Novo Transcriptome assembly and differential gene expression analysis of *Chirostoma estor* under continuous LED lighting and 12L:12D photoperiod



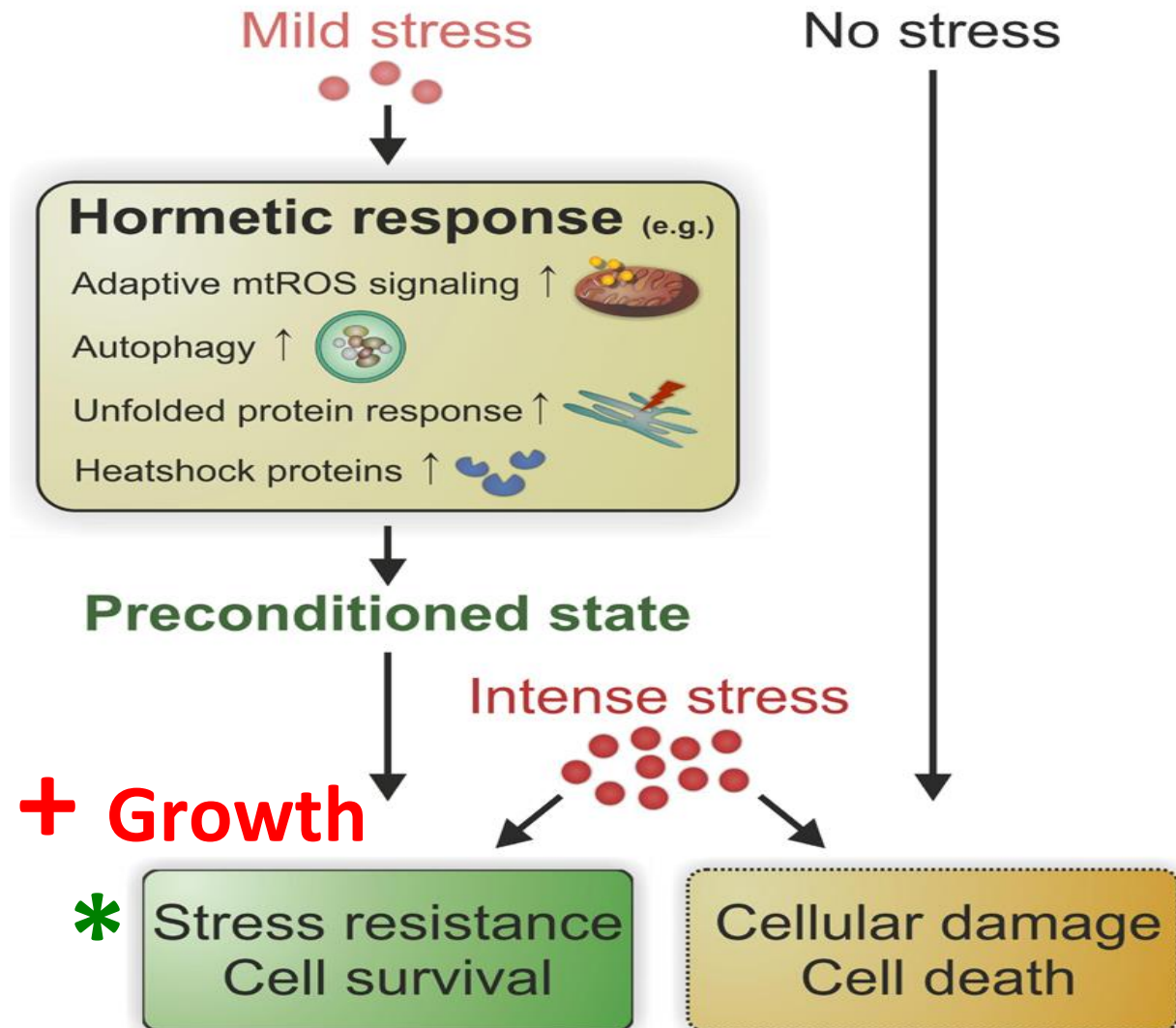
Differential expression and correlation between samples under continuous light vs control



Overregulated key genes in CL.	
Genes	Principal know functions
UCP2_3	Control mitochondria derived ROS
AQP3	H ₂ O ₂ intake, cell migration
LACTB2	ROS response, cell proliferation
KLF5, ANKRD9, PDK2	Mitochondrial fission, antiapoptotic, oxidative stress resistance
H1_5	H3K9 Metilation, Response to ROS, growth
CYP2K	Lipid peroxidation repair
EGFR, EGR1, HGS , CTGF, RASSF9_10, SCEL	Growth and cell differentiation related factors
LAM3B, COL1A, CTHR1, CGNL1	Extracellualr matrix development
SLC432A, KLF5	Aminoacid intestinal absortion, Activation of intestinal alkaline phosphatase
CMA1, FCGBP, RAG1, TRIB1,GIMA4, PELI2 NATTERIN-3,MUC2A	Imnue system response, VDJ recombination, mucus production
TEF (Thyrotrop embryonic factor)	Response direct to light, Activate DNA repair genes, antiapoptotic in presence of ROS

How the fish improved growth and feed efficiency under LED-CL? Hypothesis....

Overregulated key genes in CL.		
Genes	Principal know functions	
UCP2_3	Control mitochondria derived ROS	* +
AQP3	H ₂ O ₂ intake, cell migration	* +
LACTB2	ROS response, cell proliferation	* +
KLF5, ANKRD9, PDK2	Mitochondrial fission, antiapoptotic, oxidative stress resistance	* +
H1_5	H3K9 Metilation, Response to ROS, growth	* +
CYP2K	Lipid peroxidation repair	* +
EGFR, EGR1, HGS , CTGF, RASSF9_10, SCEL	Growth and cell differentiation related factors	* +
LAM3B, COL1A, CTHR1, CGNL1	Extracellulr matrix development	* +
SLC432A, KLF5	Aminoacid intestinal absortion, Activation of intestinal alkaline phosphatase	* +
CMA1, FCGBP, RAG1, TRIB1,GIMA4, PELI2 NATTERIN-3,MUC2A	Inmune system response, VDJ recombination, mucus production	* +
TEF (Thyrotrop embryonic factor)	Response direct to light, Activate DNA repair genes antiapoptotic in presence of ROS	* +



Over and downregulated genes in CL

Overregulated key genes in CL.

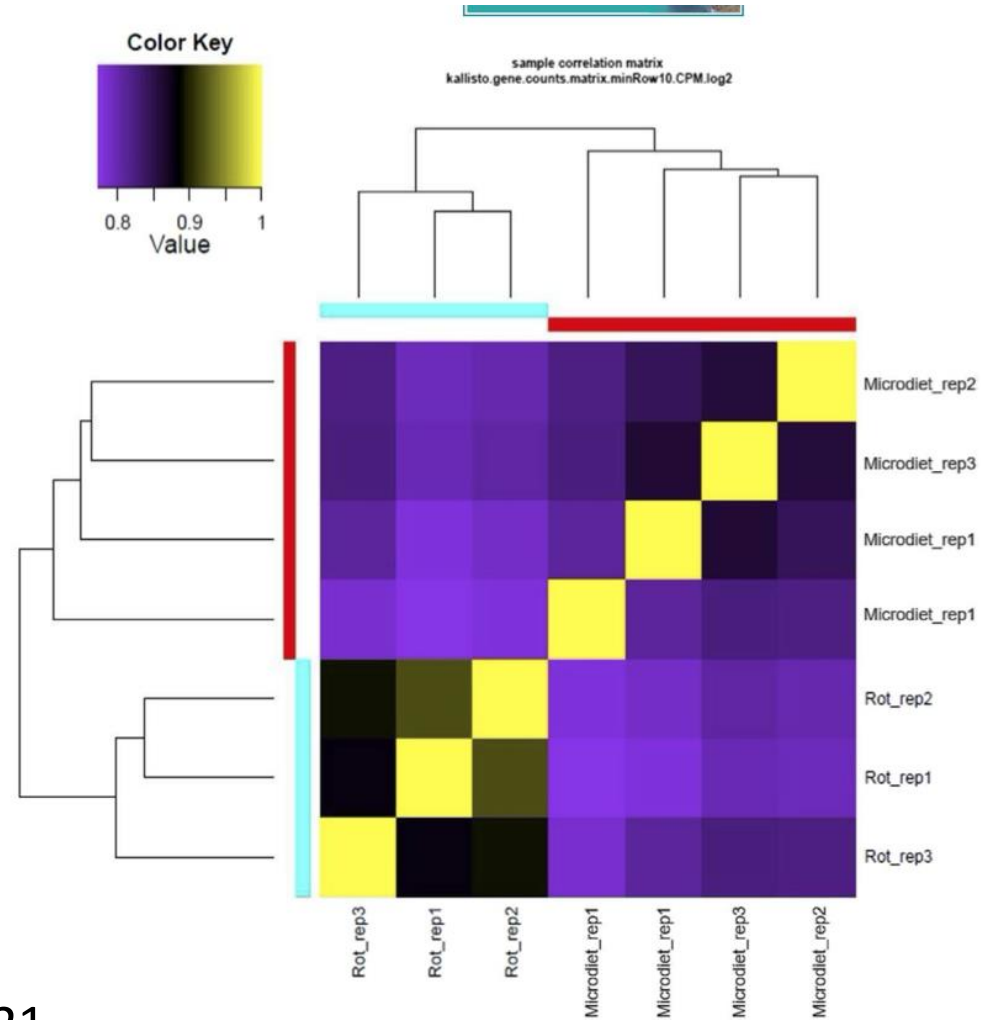
Genes	Principal know functions
UCP2_3	ROS diminishing, regulated by food intake, cell proliferation
AQP3	H ₂ O ₂ intake, cell migration
LACTB2	ROS response, cell proliferation
KLF5, ANKRD9, PDK2	Mitochondrial fission, antiapoptotic, oxidative stress resistance
H1_5	H3K9 Metilation, Response to ROS, growth
CYP2K	Lipid peroxidation repair
EGFR, EGR1, HGS , CTGF, RASSF9_10, SCEL	Growth and cell differentiation related factors
LAM3B, COL1A, CTHR1, CGNL1	Extracellualr matrix development
SLC432A, KLF5	Aminoacid intestinal absortion, Activation of intestinal alkaline phosphatase
CMA1, FCGBP, RAG1, TRIB1,GIMA4, PELI2 NATTERIN-3,MUC2A	Inmune system response, VDJ recombination, mucus production
TEF (Thyrotrop embryonic factor)	Response direct to light, Activate DNA repair genes, antiapoptotic in presence of ROS

Downregulated key genes in CL.

Genes	Principal know functions
MUC2B, MARCH 1 CHAT, LYZC, HSPA1_8,RGS, GIMAP5,	Inmune system response
MYO7, OPNS1W, GUCY2F, PRPH2, PROM1, GPR98, RGS9, RPE65, GRK1_7, CNGB1, ABCA4, ARL6, PDE6A, NR1F2, GUCA1, CNGB1, LRIT1,LRIT3, RP1	Retinal functions and photoreception
STX3, PCLO, KCNMA1, CACNA2D1, KCNV2,PLCH2,FMN1,AMPH, UNC13, GDI1_2,ANK2, EPS8, ATCAY, SV2A, PIP5K, CPLX3_4, SLC17A6_7_8, GRIN2A, ARL13B, SLC6A1,SYNE1, ADCY8, SYT1, CRTAC1	Synaptic processes, Neurotransmission, Neurogenesis,
SLC4A10, SLC5A1, SLC22A13, SLC24A2, SLC34A,	Bicarbonate, glucose, urate, sodium, potasio, calcium, organic acid transporters
SGK1,	Cellular Stress response
NDUFAF3, BCDO2	Assembly of mitochondrial complex I, Response to ROS
USP2_21	Circadian clock regulation
PTN13,	Apoptotic processes
URGCP,	Abnormal cell proliferation (cancer)
OSBL7	Lipid (cholesterol) transport
SCN2A	Cardiovascular function

Evaluando el efecto de dos dietas con transcriptómica

FIGURE 3 Correlation matrix with the respective replicates of the treatments. The colour palette indicates a high correlation between replicates of the same treatment. The yellow colour indicates a correlation of 1 (comparison between the same replicate)



Evaluando el efecto de dos dietas con transcriptómica

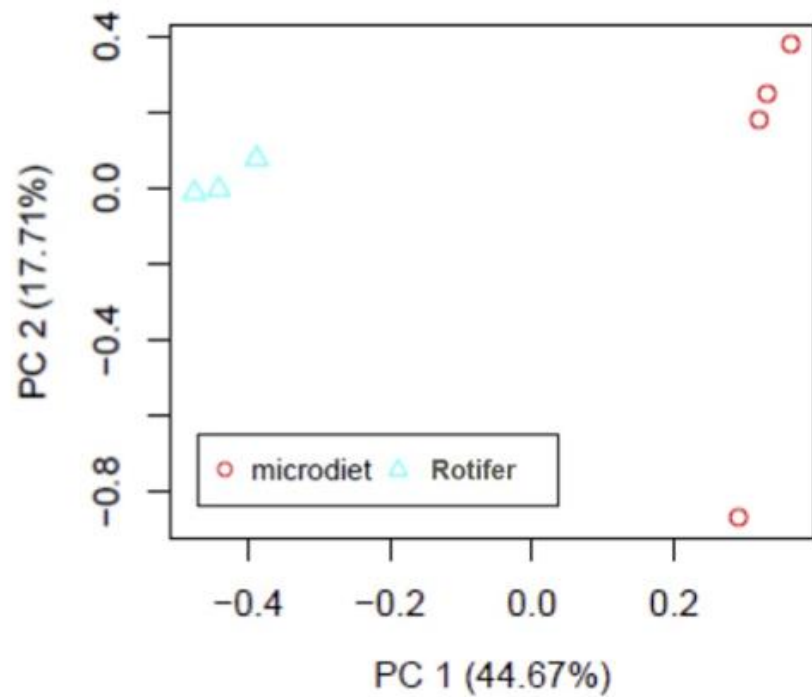


FIGURE 4 Principal component analysis of treatments (diets) based on the overall transcriptomic variance. Clustered replicates indicate similar transcriptional expression profiles. Distant replicates indicate different transcriptional profiles

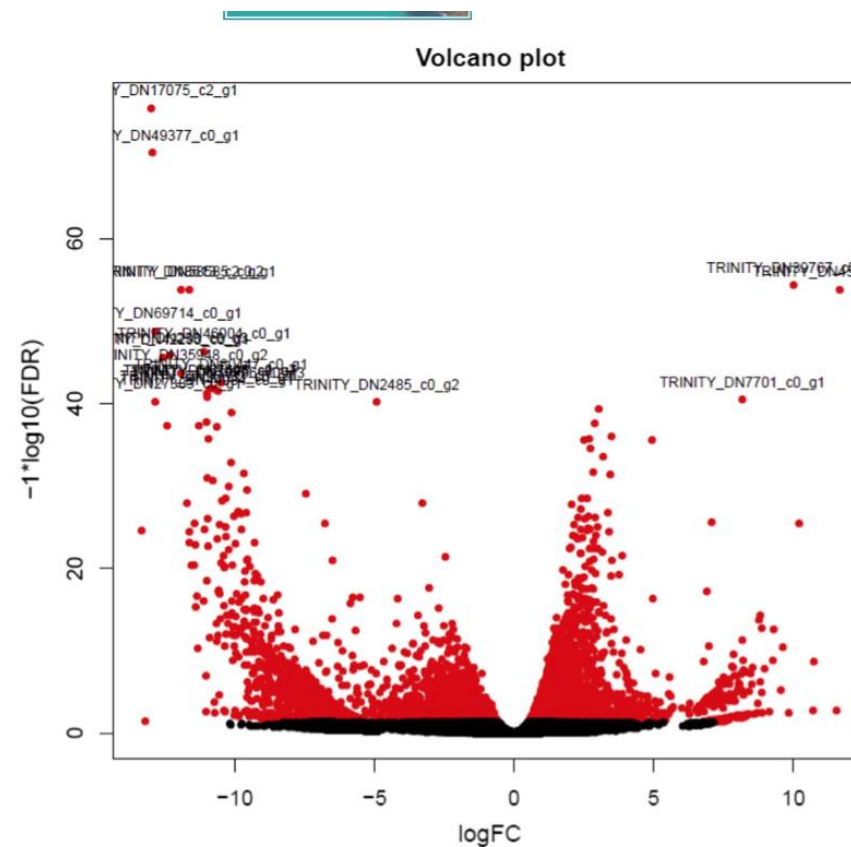


FIGURE 5 The volcano plot represents the results of differential expression analysis between both treatments. Genes with FDR (false discovery rate) <.01 are red-coloured

Evaluando el efecto de dos dietas con transcriptómica

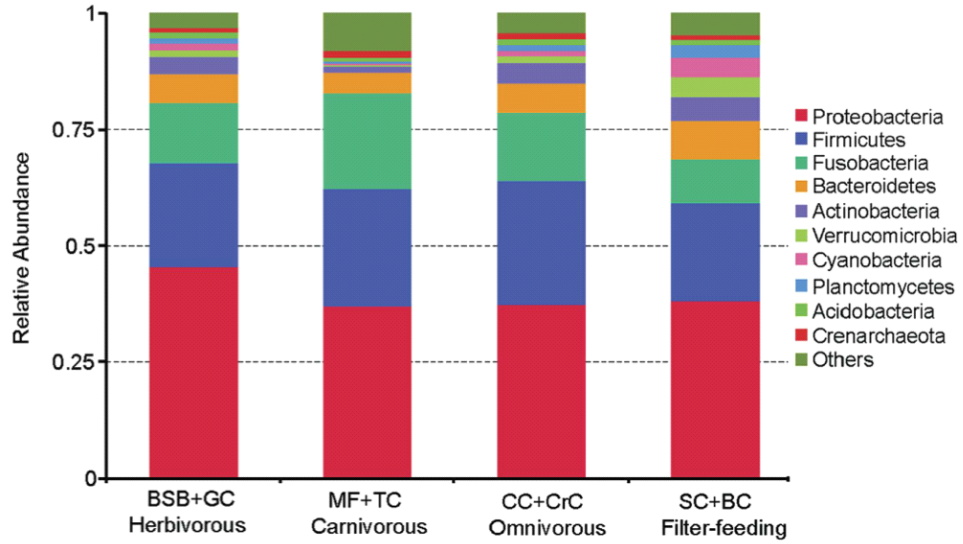
	Biological process	Genes
Cartilage	Chondrocyte development	<ul style="list-style-type: none"> Leukocyte cell-derived chemotaxin 1 Collagen alpha-1(I) chain
	Cartilage condensation	<ul style="list-style-type: none"> Collagen alpha-2(XI) chain
Bone	Mineralization	<ul style="list-style-type: none"> Tetranectin Osteocalcin Osteocalcin 2 Unique cartilage matrix-associated protein
	Ossification	<ul style="list-style-type: none"> Extracellular calcium-sensing receptor Four and a half LIM domains protein 2 Ryanodine receptor 1 PDZ and LIM domain protein 7 Unique cartilage matrix-associated protein
	Remodelling	<ul style="list-style-type: none"> Secreted phosphoprotein 24
Muscle	Binding of actin, myosin and tropomyosin filaments and their motor action	<ul style="list-style-type: none"> Myosin heavy chain fast skeletal muscle Immunoglobulin-like and fibronectin type III domain-containing protein 1 Myosin-binding protein C, fast-type
	Regulation of muscle contraction	<ul style="list-style-type: none"> Troponin T Troponin I, fast skeletal muscle Sarcoplasmic reticulum histidine-rich calcium-binding protein Triadin Sodium channel subunit beta-3 Myosin light chain 1 Troponin C, skeletal muscle
	Assembly and development of muscle fibres	<ul style="list-style-type: none"> Tropomodulin-4 Myosin-6 Myosin regulatory light chain 2, ventricular/cardiac muscle isoform Myosin light chain 3 Myosin regulatory light chain 2 Collagen alpha-3(VI) chain Myozenin-1
	Sarcomere organization	<ul style="list-style-type: none"> Calsequestrin-1 M-protein, striated muscle, Myozenin-2
	Muscular cell homeostasis	<ul style="list-style-type: none"> ATP-dependent 6-phosphofructokinase, muscle type Tropomyosin alpha-1 chain
	Epigenetic mark: Transcription regulation in myogenesis	<ul style="list-style-type: none"> SMYD1 Muscle-specific histone lysine (H3K4) methyltransferase
Digestion	Digestive enzymes or precursors	<ul style="list-style-type: none"> Elastase-1 Chymotrypsin C Chymotrypsin-like elastase family member 2a Chymotrypsinogen B Phospholipase A2
	Morphogenesis of the digestive tract, liver, and development of the exocrine pancreas	<ul style="list-style-type: none"> Probable ATP-dependent RNA helicase DDX46

TABLE 5 Microdiet-fed pike silverside larvae overexpressed genes and their biological processes

	Biological process	Genes
Cell instability	Apoptosis	<ul style="list-style-type: none"> Cellular tumour antigen p53 Fas-activated serine/threonine kinase BCL2/adenovirus E1B 19 kDa protein-interacting protein 3-like ANP32B SAFB-like transcription modulator PCBP4, Poly(rC)-binding protein 4
Cell stress	Oxidative stress	<ul style="list-style-type: none"> Lon protease homolog, mitochondrial PRODH proline dehydrogenase1 RalA-binding protein 1 F-box only protein 7
DNA damage	DNA repair (NHEJ) pathway	<ul style="list-style-type: none"> JMJD1C RBM14 SETD1A YY1
Epigenetic marks	Chromatin dynamic Transcription regulation	<ul style="list-style-type: none"> SET protein SETD1A HMGB1 YY1 Pax7/Pax3 Hypomethylation of CpG islands
	miRNAs regulation	<ul style="list-style-type: none"> N-myc proto-oncogene protein MYCN YY1

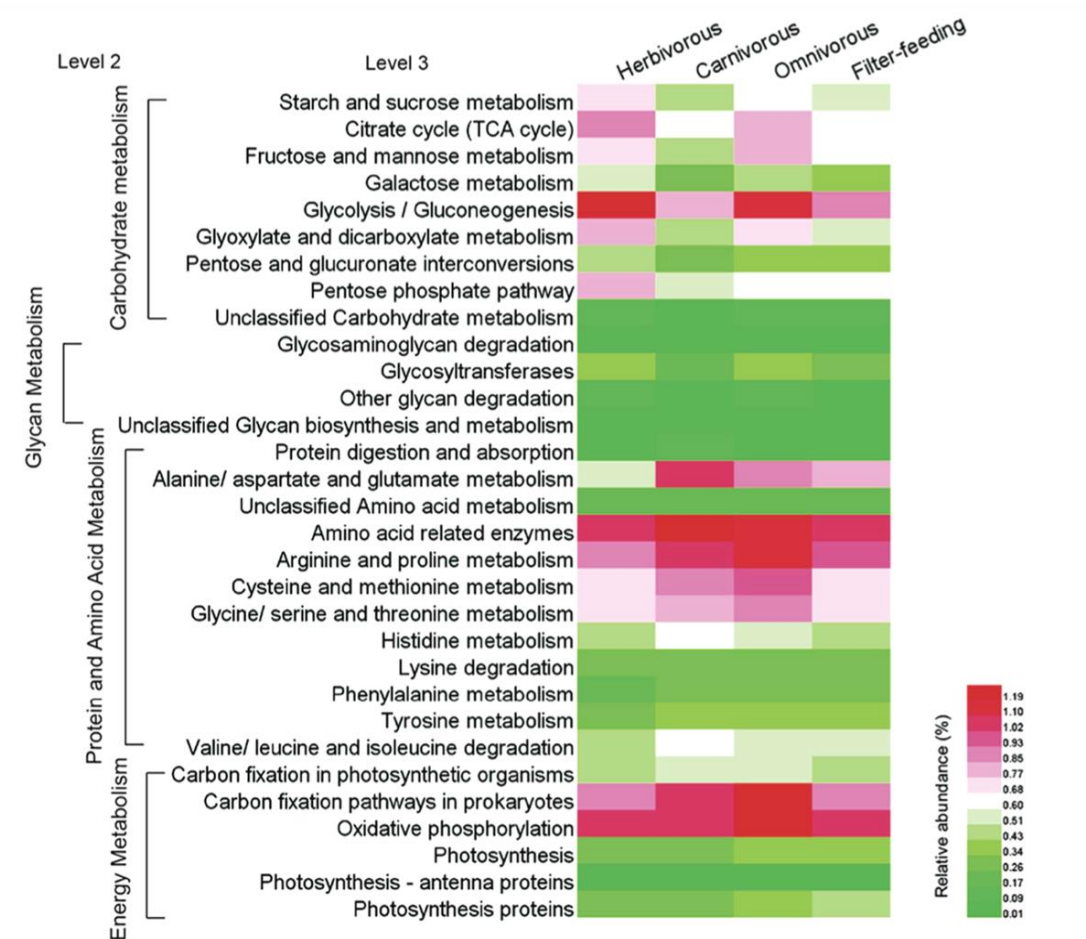
	Biological process	Genes	
Cartilage	Chondrocyte development	<ul style="list-style-type: none"> ● <i>Leukocyte cell-derived chemotaxin 1</i> ● <i>Collagen alpha-1(I) chain</i> 	
	Cartilage condensation	<ul style="list-style-type: none"> ● <i>Collagen alpha-2(XI) chain</i> 	
Bone	Mineralization	<ul style="list-style-type: none"> ● <i>Tetranectin</i> ● <i>Osteocalcin</i> ● <i>Osteocalcin 2</i> ● <i>Unique cartilage matrix-associated protein</i> 	
	Ossification	<ul style="list-style-type: none"> ● <i>Extracellular calcium-sensing receptor</i> ● <i>Four and a half LIM domains protein 2</i> ● <i>Ryanodine receptor 1</i> ● <i>PDZ and LIM domain protein 7</i> ● <i>Unique cartilage matrix-associated protein</i> 	
	Remodelling	<ul style="list-style-type: none"> ● <i>Secreted phosphoprotein 24</i> 	
Muscle	Binding of actin, myosin and tropomyosin filaments and their motor action	<ul style="list-style-type: none"> ● <i>Myosin heavy chain fast skeletal muscle</i> ● <i>Immunoglobulin-like and fibronectin type III domain-containing protein 1</i> ● <i>Myosin-binding protein C, fast-type</i> 	
	Regulation of muscle contraction	<ul style="list-style-type: none"> ● <i>Troponin T</i> ● <i>Troponin I, fast skeletal muscle</i> ● <i>Sarcoplasmic reticulum histidine-rich calcium-binding protein</i> ● <i>Triadin</i> ● <i>Sodium channel subunit beta-3</i> ● <i>Myosin light chain 1</i> ● <i>Troponin C, skeletal muscle</i> 	
	Assembly and development of muscle fibres	<ul style="list-style-type: none"> ● <i>Tropomodulin-4</i> ● <i>Myosin-6</i> ● <i>Myosin regulatory light chain 2, ventricular/cardiac muscle isoform</i> ● <i>Myosin light chain 3</i> ● <i>Myosin regulatory light chain 2</i> ● <i>Collagen alpha-3(VI) chain</i> ● <i>Myozenin-1</i> 	
	Sarcomere organization	<ul style="list-style-type: none"> ● <i>Calsequestrin-1</i> ● <i>M-protein, striated muscle, Myozenin-2</i> 	
	Muscular cell homeostasis	<ul style="list-style-type: none"> ● <i>ATP-dependent 6-phosphofructokinase, muscle type</i> ● <i>Tropomyosin alpha-1 chain</i> 	
	Epigenetic mark: Transcription regulation in myogenesis	<ul style="list-style-type: none"> ● <i>SMYD1 Muscle-specific histone lysine (H3K4) methyltransferase</i> 	
	Digestion	Digestive enzymes or precursors	<ul style="list-style-type: none"> ● <i>Elastase-1</i> ● <i>Chymotrypsin C</i> ● <i>Chymotrypsin-like elastase family member 2a</i> ● <i>Chymotrypsinogen B</i> ● <i>Phospholipase A2</i>
		Morphogenesis of the digestive tract, liver, and development of the exocrine pancreas	<ul style="list-style-type: none"> ● <i>Probable ATP-dependent RNA helicase DDX46</i>

Estudios metagenómicos de la microbiota



Carpas con cuatro hábitos alimenticios
V4 del 16S rRNA (Illumina)

(Liu et al. 2016)



Metodología

Microbiota



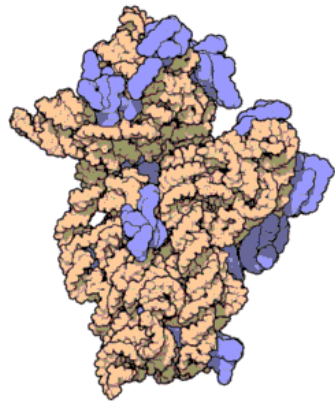
Extracción de ADN
(CTAB)

Secuenciación
Illumina

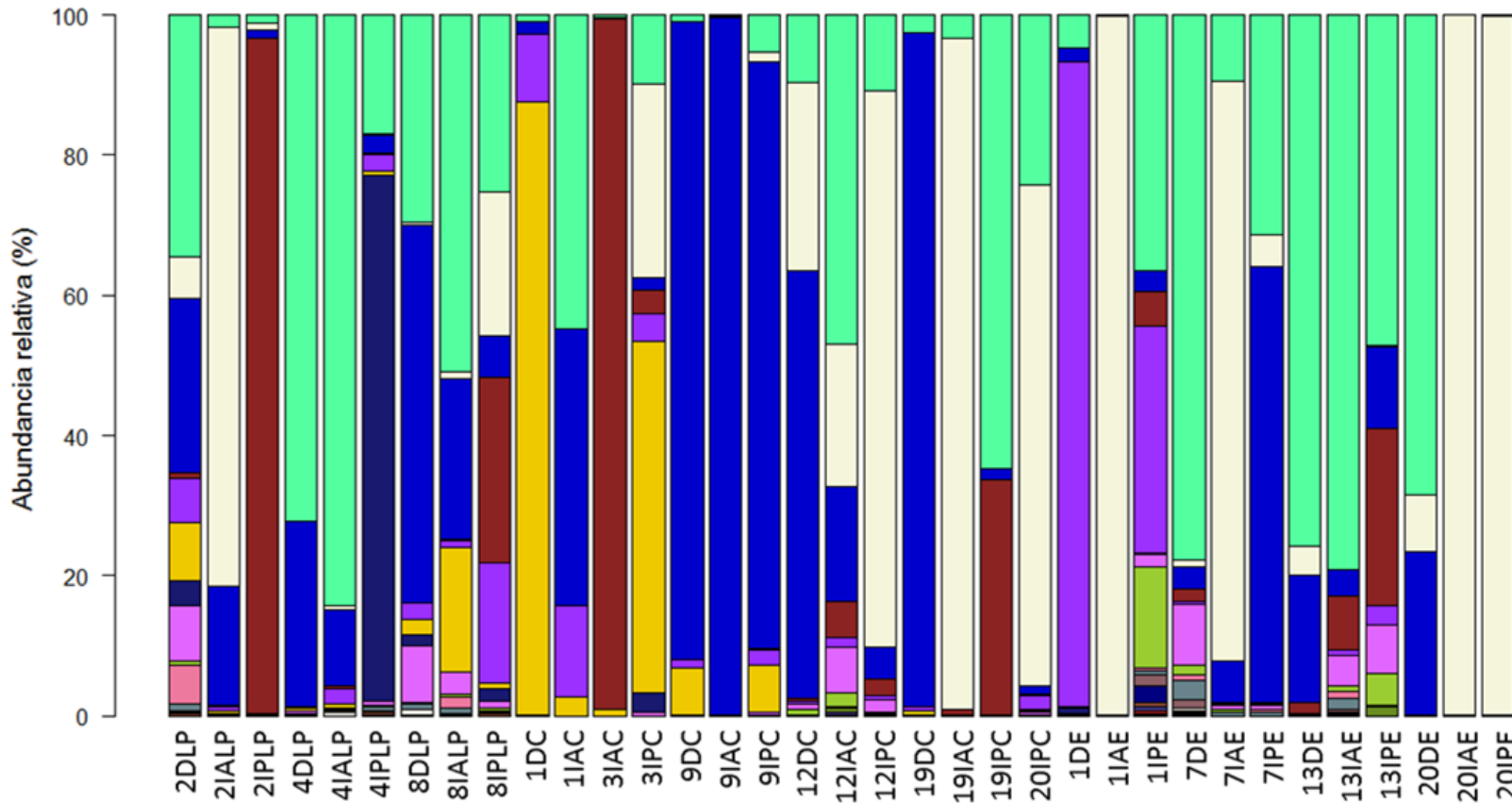
Gen ARNr 16S (V3 y V4)

Análisis bioinformático

Análisis estadístico
NMDS y ANOSIM con R



Perfiles de microbiota intestinal (Phylum) del Pez blanco *C. estor*



Nivel taxonómico	Promedio
Phylum	
Proteobacteria	27.76
Tenericutes	23.79
Firmicutes	22.77
Unclassified_p	8.89
Actinobacteria	5.6
Cyanobacteria	5.35
Fusobacteria	2.46
Bacteroidetes	1.58
Acidobacteria	0.75
Planctomycetes	0.31

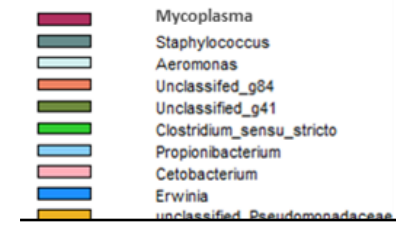
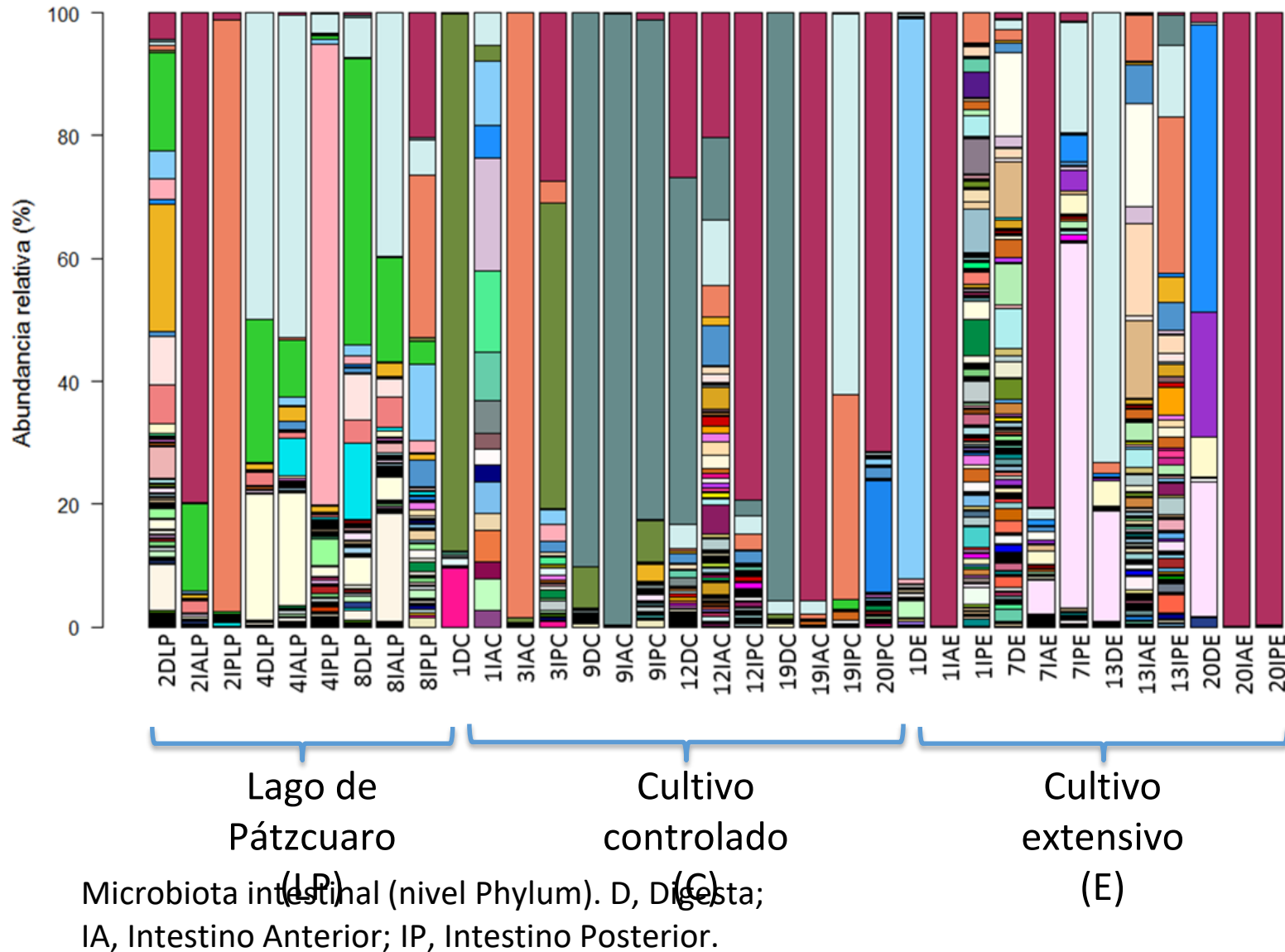
Lago de Pátzcuaro (LP)

Cultivo controlado (C)

Cultivo extensivo (E)

Microbiota intestinal (nivel Phylum). D, Digesta; IA, Intestino Anterior; IP, Intestino Posterior.

Perfiles de microbiota intestinal (Género) del Pez blanco *C. estor*



Nivel taxonómico		Promedio
Género		
	<i>Mycoplasma</i>	23.24188
	<i>Staphylococcus</i>	12.7433
	<i>Aeromonas</i>	10.12672
	<i>Unclassified_g84</i>	8.888901
	<i>Unclassified_g41</i>	4.474305
	<i>Clostridium_sensu_stricto</i>	3.81236
	<i>Propionibacterium</i>	3.626169
	<i>Unclassified_g4</i>	3.19765
	<i>Cetobacterium</i>	2.452709
	<i>Erwinia</i>	1.732638
	<i>Unclassified_g25</i>	1.554027
	<i>unclassified_Pseudomonadaceae</i>	1.147998
	<i>Acinetobacter</i>	1.034033
	<i>Azonexus</i>	0.915969
	<i>Unclassified_g39</i>	0.81418
	<i>Exiguobacterium</i>	0.727521
	<i>Pantoea</i>	0.720089
	<i>Massilia</i>	0.700117
	<i>Flavobacterium</i>	0.683944
	<i>Dechloromonas</i>	0.673887

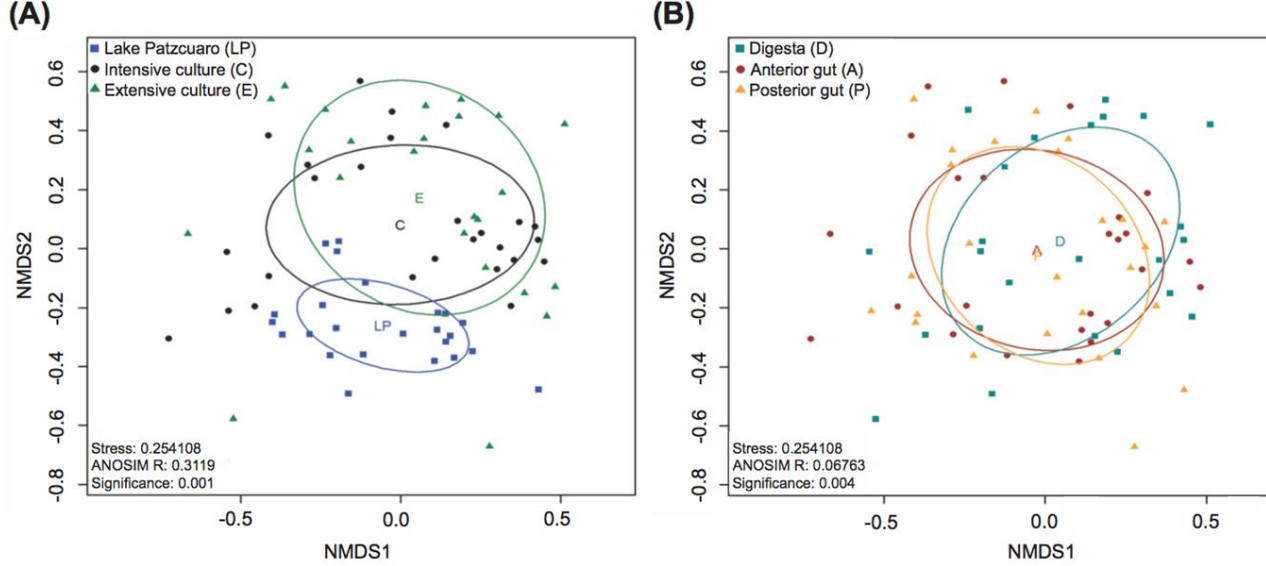
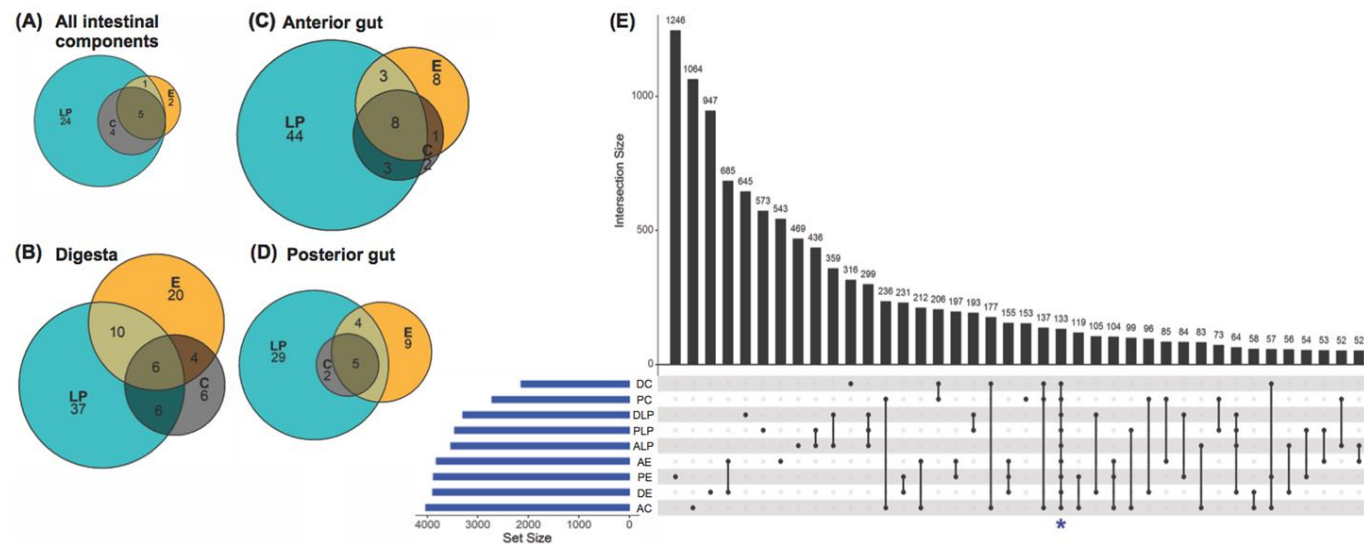
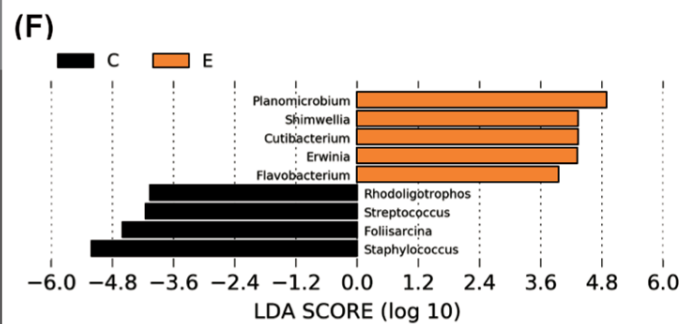
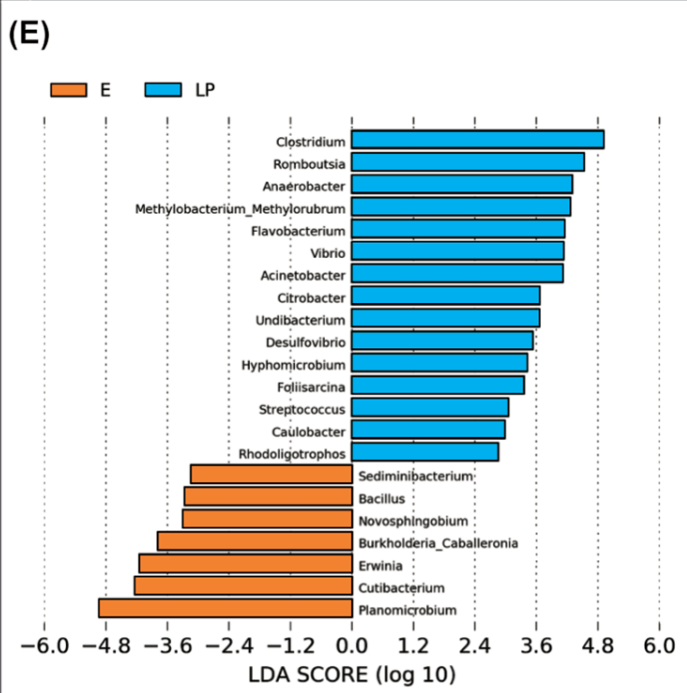
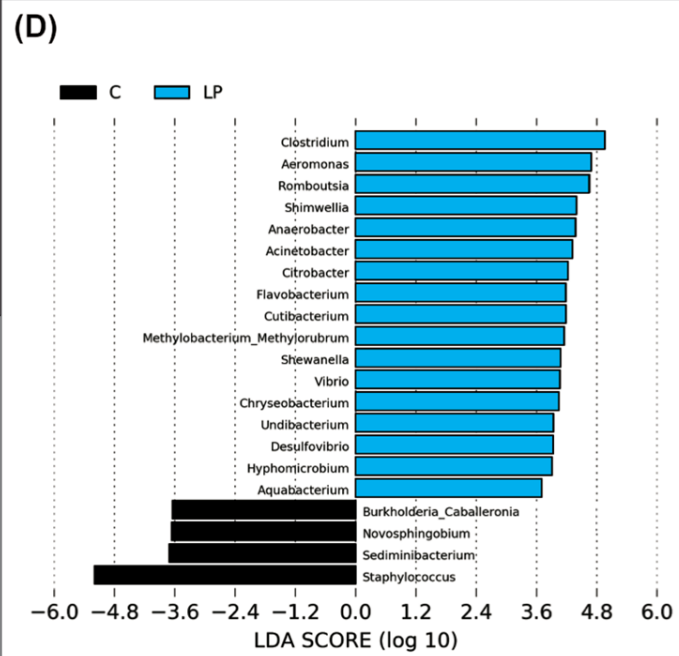
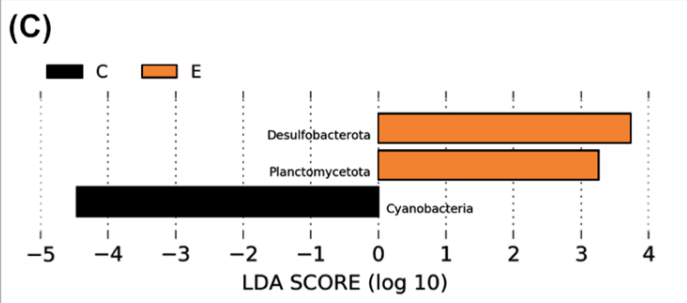
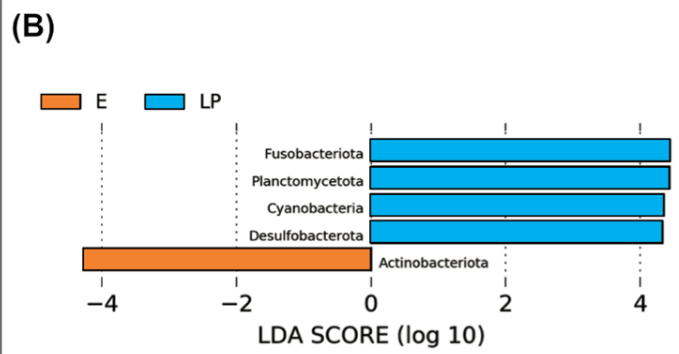
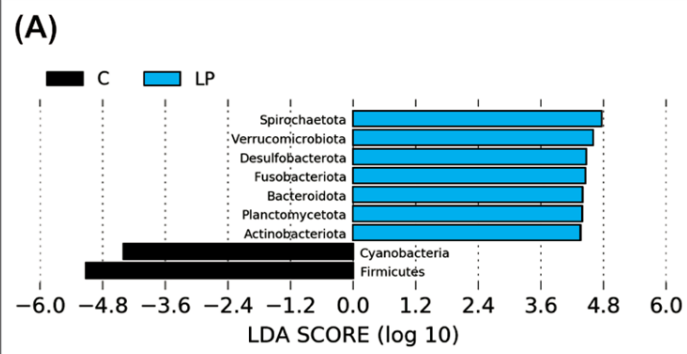


Figure 2 Beta diversity analysis. Nonmetric multidimensional scaling (NMDS) plot of the Br beta-diversity from *C. estor* microbiota profiles, estimated from the OTU abundance matrix. The groupings were by environment (A) and by intestinal component (B).

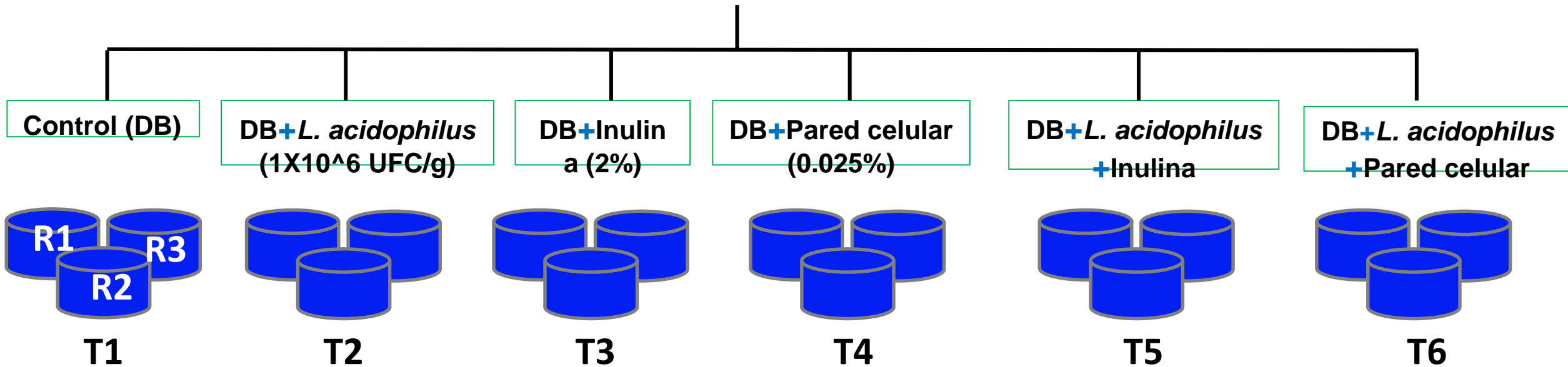
Full-size DOI: 10.7717/peerj





Experimento de suplementos probióticos y prebióticos para determinar su efecto en el crecimiento, la supervivencia y la modulación de la microbiota y metaboloma intestinal.

Dietas con suplementos alimenticios (pro y prebióticos) en *C. estor* en cautiverio



DB=Dieta base administrada en la planta

Duración de experimento: 12 semanas

No. de peces por tanque: 113 peces de 7 meses

Peso prom.: 5.6 g
Talla prom.: 9.2 cm

Resultados en crecimiento (peso)

Dif. significativas (ANOVA) en crecimiento a las 12 semanas.

Mejor crecimiento en peso en combinación de probióticos + prebióticos.

Tratamientos

1: DB = Dieta base

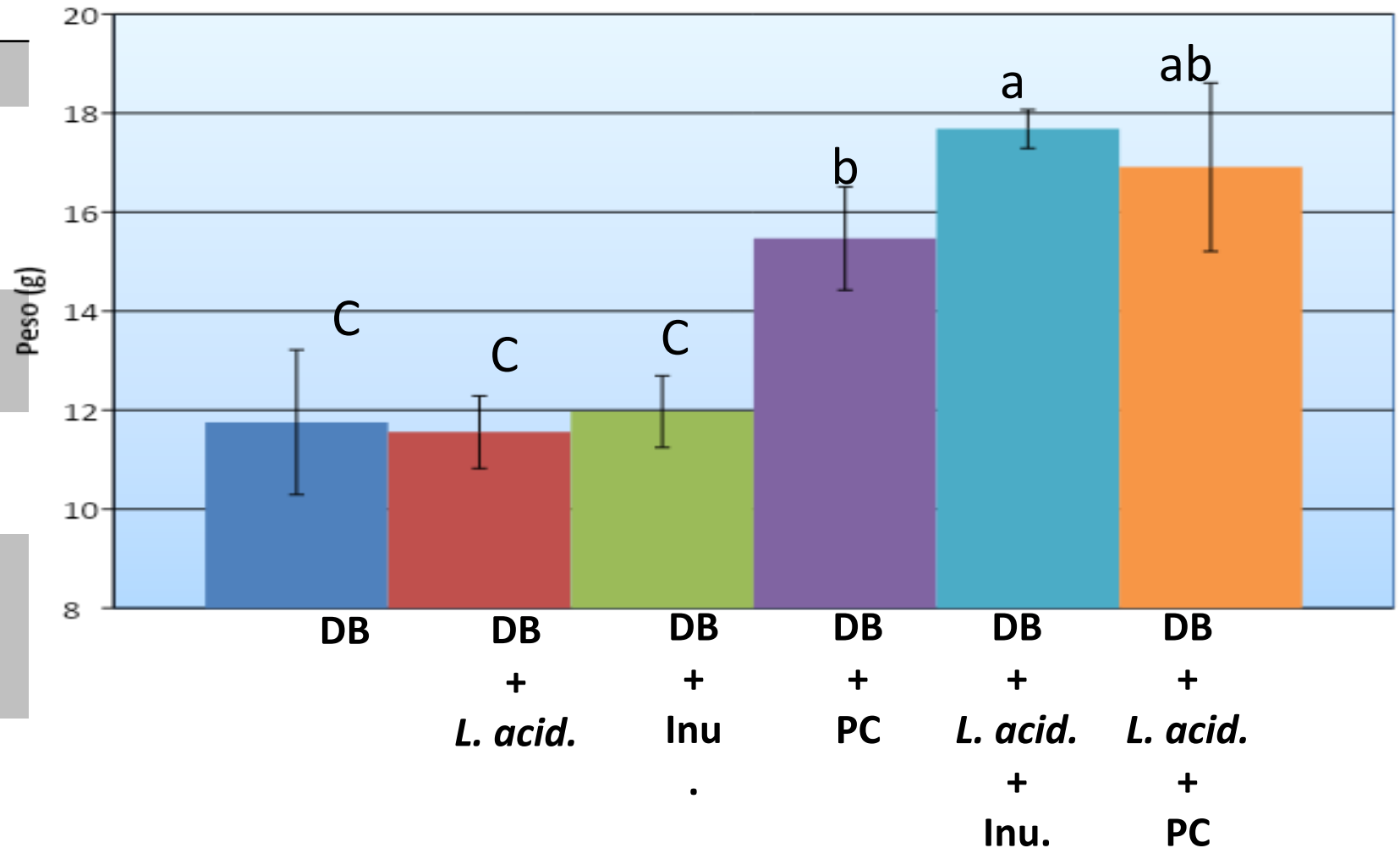
2: DB + *Lactobacillus acidophilus*

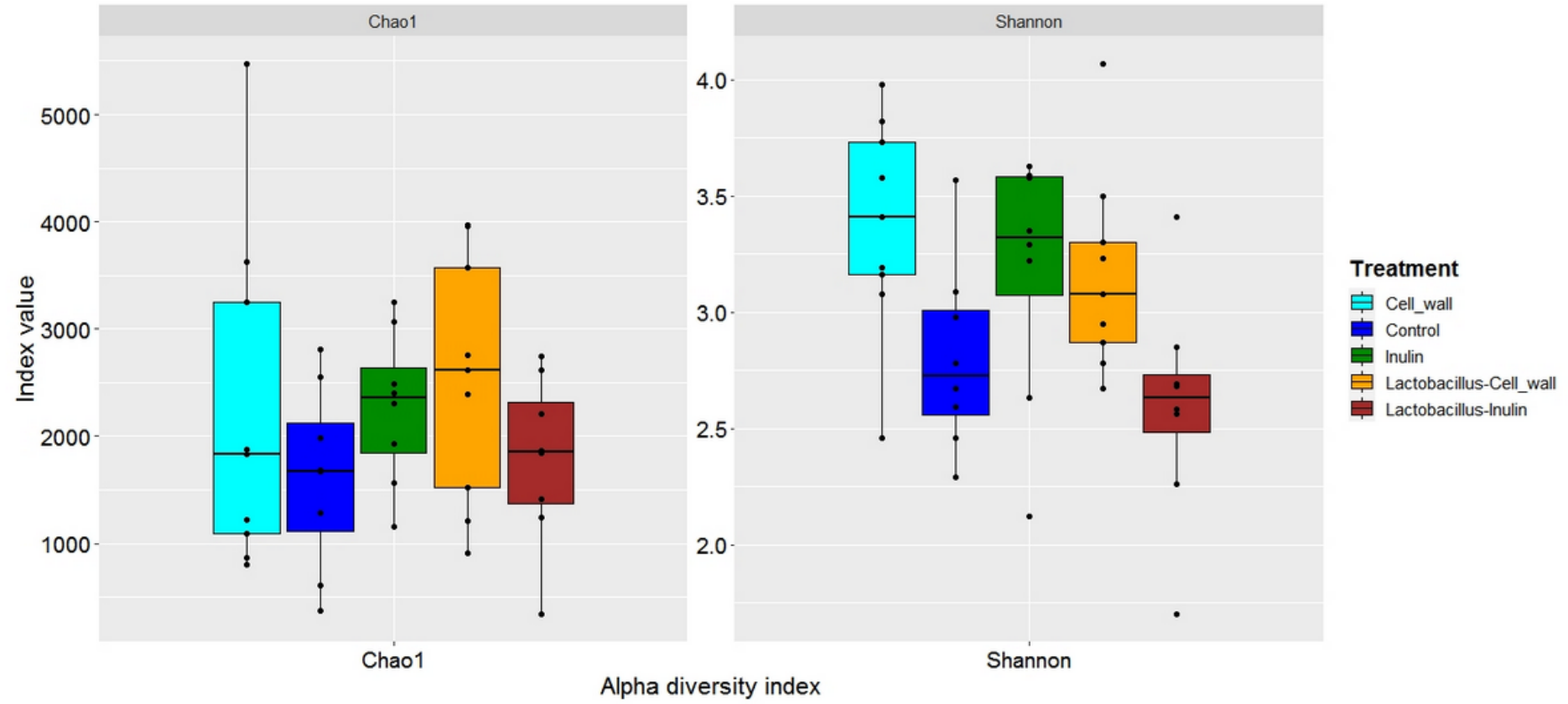
3: DB + Inulina

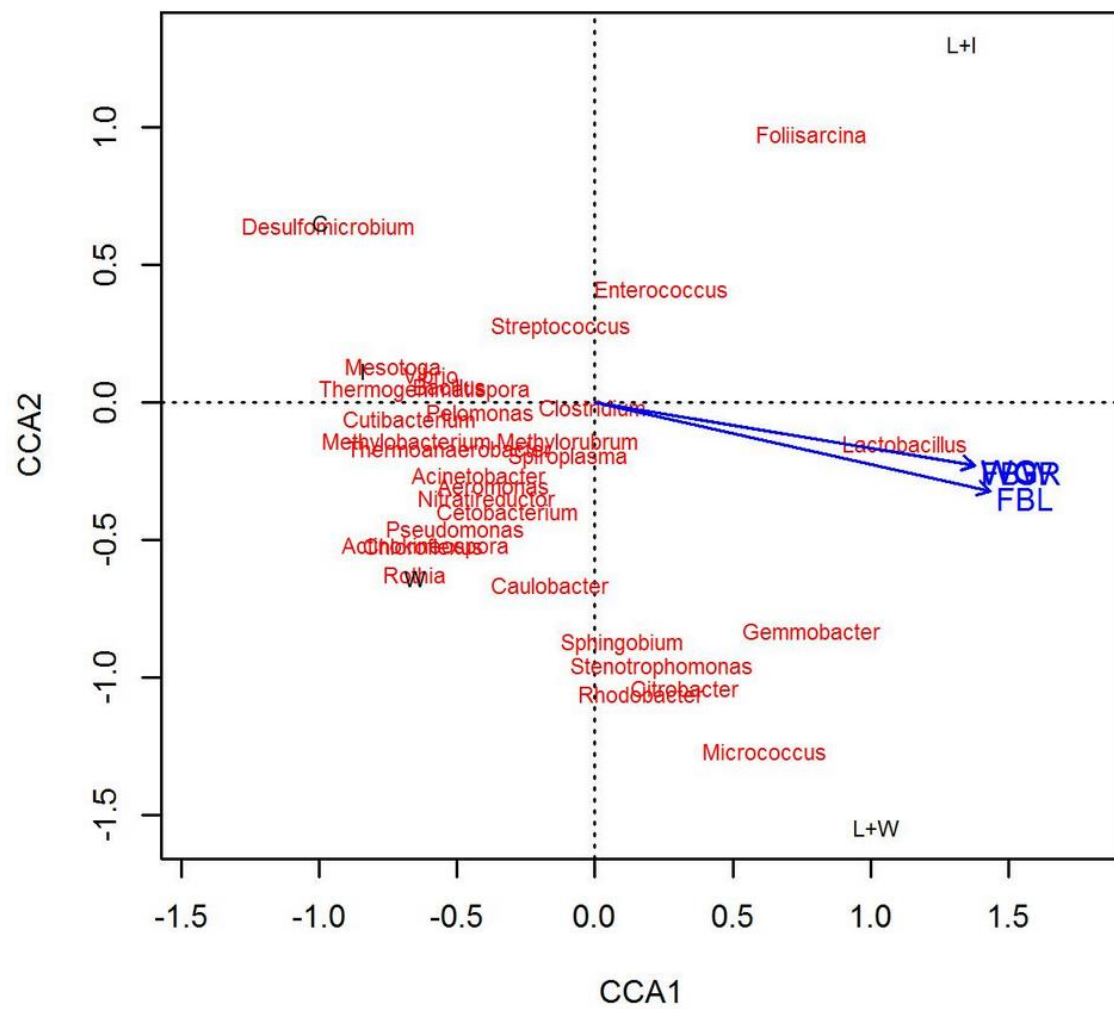
4: DB + Pared celular

5: DB + *L. acidophilus* + Inulina

6: DB + *L. acidophilus* + Pared celular







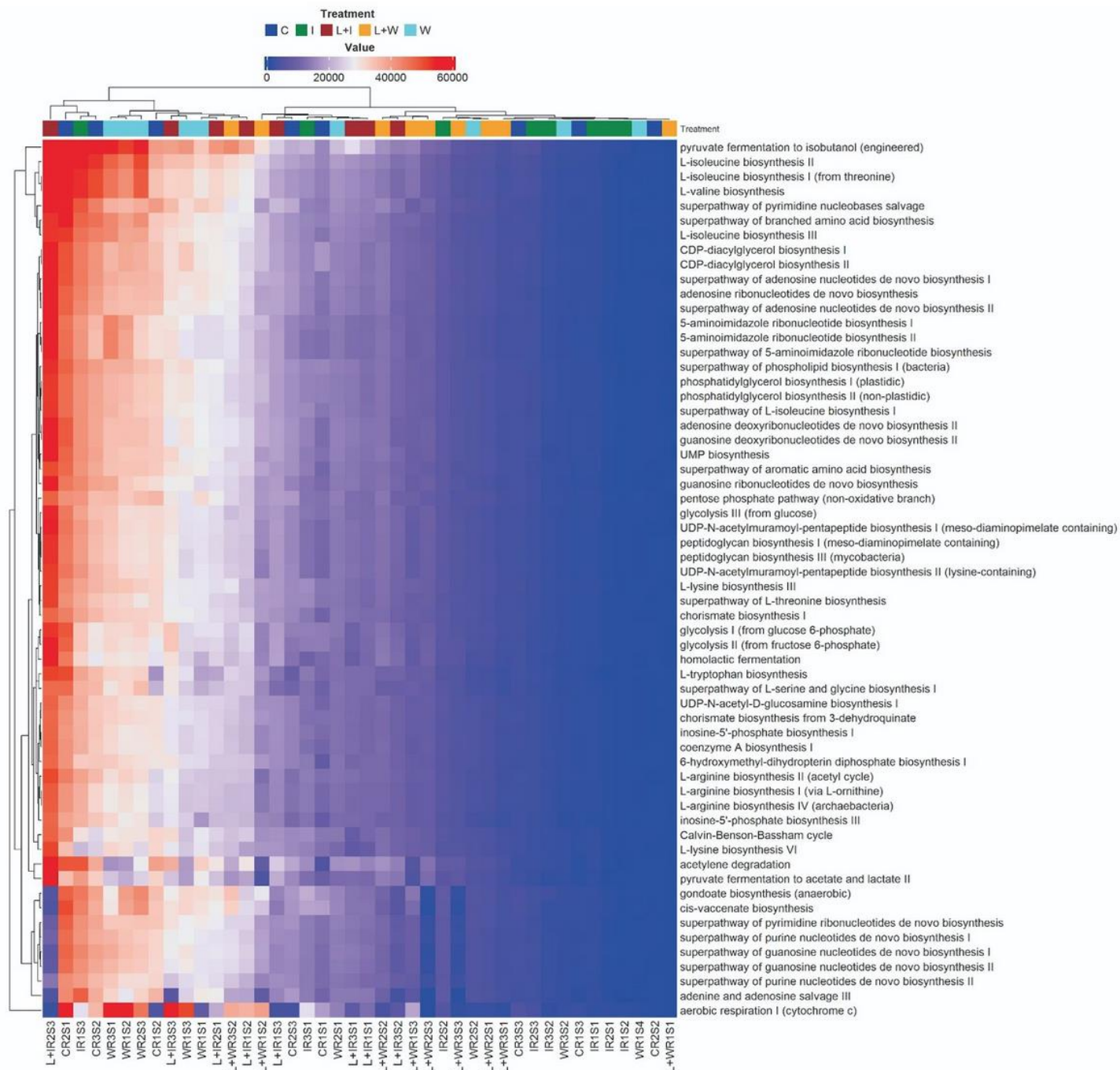
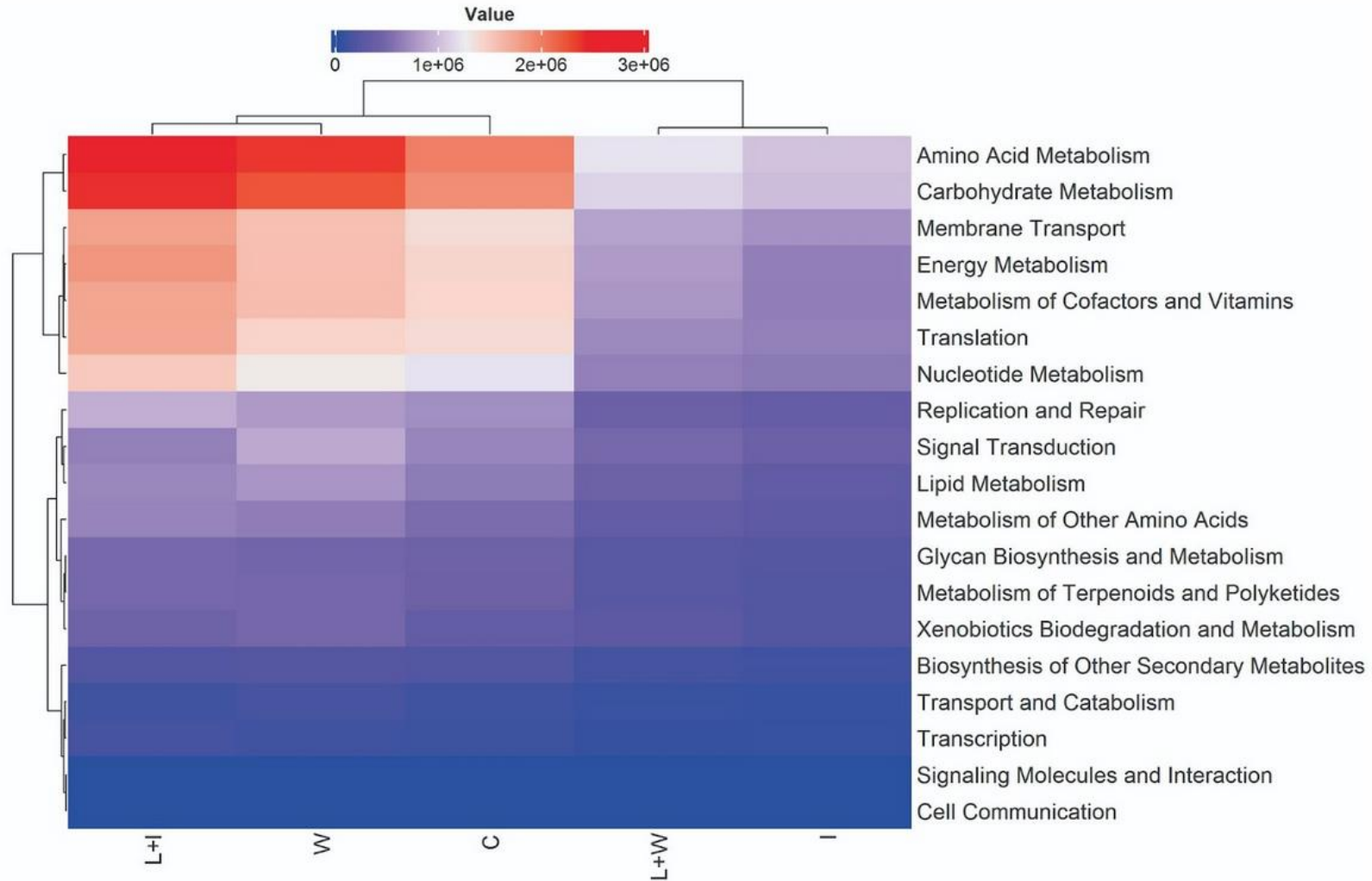


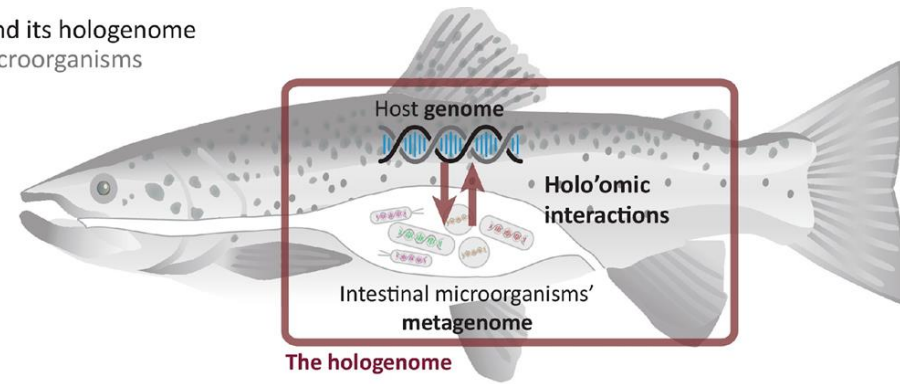
Figure S2 60 vías más abundantes de PICRUSt2. Heatmap of the 60 most abundant functional pathways

Predictive functional of KOs metagenomics analysis by PICRUSt2

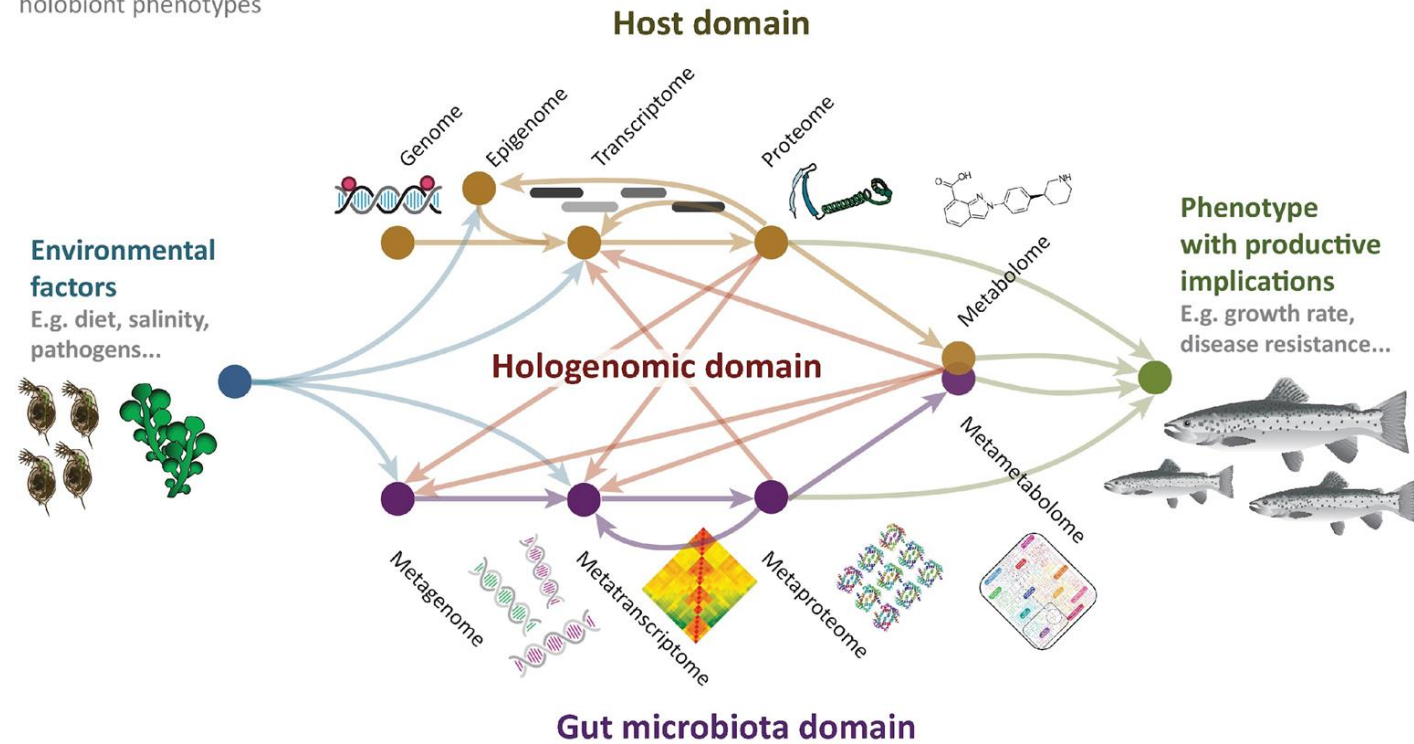


Conclusions

(A) The holobiont and its hologenome
Host + intestinal microorganisms



(B) Holo'omic interactions
Biomolecular interactions between hosts and symbiotic microorganisms triggered by environmental factors yield different holobiont phenotypes



Exposures, Lifestyle and Clinical Factors, Phenotypes



Transcriptome
Metabolome
Epigenome
Microbiome
Proteome
Genome

Findable, Accessible, Interoperable, and Reusable **Data**

Findable, Accessible, Interoperable, and Reusable **Software**



Derived from multi-omics data analysis methods

Use and analysis of curated, community, and in-silico resources



Analyze and Interpret multi-omics data in the context of biology, chemistry, and disease

Conclusions and Perspectives

Bioinformatics has brought about a paradigm shift across diverse domains, wielding transformative influence in biotechnology, medicine, environmental science, and agriculture.

In biotechnology, its progress is evident in the accelerated design of genetically modified organisms, fostering innovations in enzymes, biofuels, and biomaterials with far-reaching implications for sustainable practices.

Meanwhile, in the medical area, bioinformatics stands as a guiding tool, shaping personalized healthcare through in-depth genomics, nutrigenomics, pharmacogenomics, clinical data analysis, and immune repertoire analysis. By pinpointing disease biomarkers, it facilitates early diagnosis and tailored treatments, and its role in drug and vaccine discovery expedites the possibilities to improve the treatment of diseases.

Acknowledgements

*CONAHCYT - Instituto de Investigaciones Agropecuarias y Forestales, Universidad Michoacana de San Nicolás de Hidalgo
Centro de Investigación Sobre Enfermedades Infecciosas, Instituto Nacional de Salud Pública, Cuernavaca, Morelos, Mexico
Unidad Universitaria de Secuenciación Masiva y Bioinformática, Instituto de Biotecnología, Universidad Nacional Autónoma
de México, Morelos, Mexico.*

Microbial Informatics Team, European Bioinformatics Institute (EMBL-EBI), Cambridge, UK

*Laboratorio Nacional de Nutrigenómica y Microbiómica Digestiva Animal and Consejo Nacional de Humanidades, Ciencia y
Tecnología.*

References

- Amillano Cisneros, Jesus Mateo, Luciana Raggi Hoyos, y Carlos A. Martínez Palacios. 2019. “Implementación de probióticos en la acuicultura de peces con importancia comercial: una revisión.” *Revista Mexicana de Industria y Salud* 2 (9): 1–11. <http://fcqgp.ujed.mx/docs/posgrado/remdis-v2-9.pdf>.
- Amillano-Cisneros, Jesús Mateo, et al. “Loss of Gut Microbial Diversity in the Cultured, Agastric Fish, Mexican Pike Silverside (*Chirostoma Estor*: Atherinopsidae).” *PeerJ*, vol. 10, 2022, p. e13052, <https://doi.org/10.7717/peerj.13052>.
- Escobar-Zepeda, A., Godoy-Lozano, E. E., Raggi, L., Segovia, L., Merino, E., Gutiérrez-Rios, R. M., Juarez, K., Licea-Navarro, A. F., Pardo-Lopez, L., & Sanchez-Flores, A. (2018). Analysis of sequencing strategies and tools for taxonomic annotation: Defining standards for progressive metagenomics. *Scientific Reports*, 8. <https://doi.org/10.1038/s41598-018-30515-5>
- Escobar-Zepeda, A., Sanchez-Flores, A., & Quirasco Baruch, M. (2016). Metagenomic analysis of a Mexican ripened cheese reveals a unique complex microbiota. *Food Microbiology*, 57, 116–127. <https://doi.org/10.1016/j.fm.2016.02.004>
- Godoy-Lozano, E. E., Téllez-Sosa, J., Sánchez-González, G., Sámano-Sánchez, H., Aguilar-Salgado, A., Salinas-Rodríguez, A., Cortina-Ceballos, B., Vivanco-Cid, H., Hernández-Flores, K., Pfaff, J. M., Kahle, K. M., Doranz, B. J., Gómez-Barreto, R. E., Valdovinos-Torres, H., López-Martínez, I., Rodriguez, M. H., & Martínez-Barnetche, J. (2016). Lower IgG somatic hypermutation rates during acute dengue virus infection is compatible with a germinal center-independent B cell response. *Genome Medicine*, 8, 23. <https://doi.org/10.1186/s13073-016-0276-1>
- Jiménez-Jacinto, V., Sanchez-Flores, A., & Vega-Alvarado, L. (2019). Integrative Differential Expression Analysis for Multiple EXperiments (IDEAMEX): A Web Server Tool for Integrated RNA-Seq Data Analysis. *Frontiers in Genetics*, 10, 279. <https://doi.org/10.3389/fgene.2019.00279>



© MARVID-Mexico

No part of this document covered by the Federal Copyright Law may be reproduced, transmitted or used in any form or medium, whether graphic, electronic or mechanical, including but not limited to the following: Citations in articles and comments Bibliographical, compilation of radio or electronic journalistic data. For the effects of articles 13, 162,163 fraction I, 164 fraction I, 168, 169,209 fraction III and other relative of the Federal Law of Copyright. Violations: Be forced to prosecute under Mexican copyright law. The use of general descriptive names, registered names, trademarks, in this publication do not imply, uniformly in the absence of a specific statement, that such names are exempt from the relevant protector in laws and regulations of Mexico and therefore free for General use of the international scientific community. VCIERMMI is part of the media of MARVID-Mexico., E: 94-443.F: 008- (www.marvid.org/booklets)