Oxford Vaccine Group
Department of Paediatrics
Clinical Vaccine Research and Immunisation Education
Medical Sciences Division

UNIVERSITY OF OXFORD

Methodologies to assess transcriptome, LongRNome, miRNome

Andre Goncalves
Postdoctoral Researcher
30th October 2023

IESAM
INSTITUTO DE ESTUDOS SUPERIORES DA AMAZÔNIA
Excelência em Meio Ambiente e Tecnologia

Computer Engineering
Graduated

Computational base
Algorithm and programming language

UNIVERSIDADE FEDERAL DO PARÁ

Genetic and Molecular Biology
Master

miRNA Expression Profile for the
Human Gastric Antrum Region

UNIVERSIDADE FEDERAL DE MINAS GERAIS
INCIPIT VITA NOVA
7 DE SETEMBRO DE 1927

Bioinformatic
PhD

miRNA and mRNA Expression Profile for the
Macrophages and *T. cruzi* interaction

UNIVERSIDADE FEDERAL
DE SANTA CATARINA

LABORATÓRIO DE IMUNOLOGIA APLICADA A AQUICULTURA
LIAA

Postdoc

Shrimp intestinal immunity

USP
Universidade de São Paulo

CSBL

Postdoc

System biology/immunology
in infectious diseases

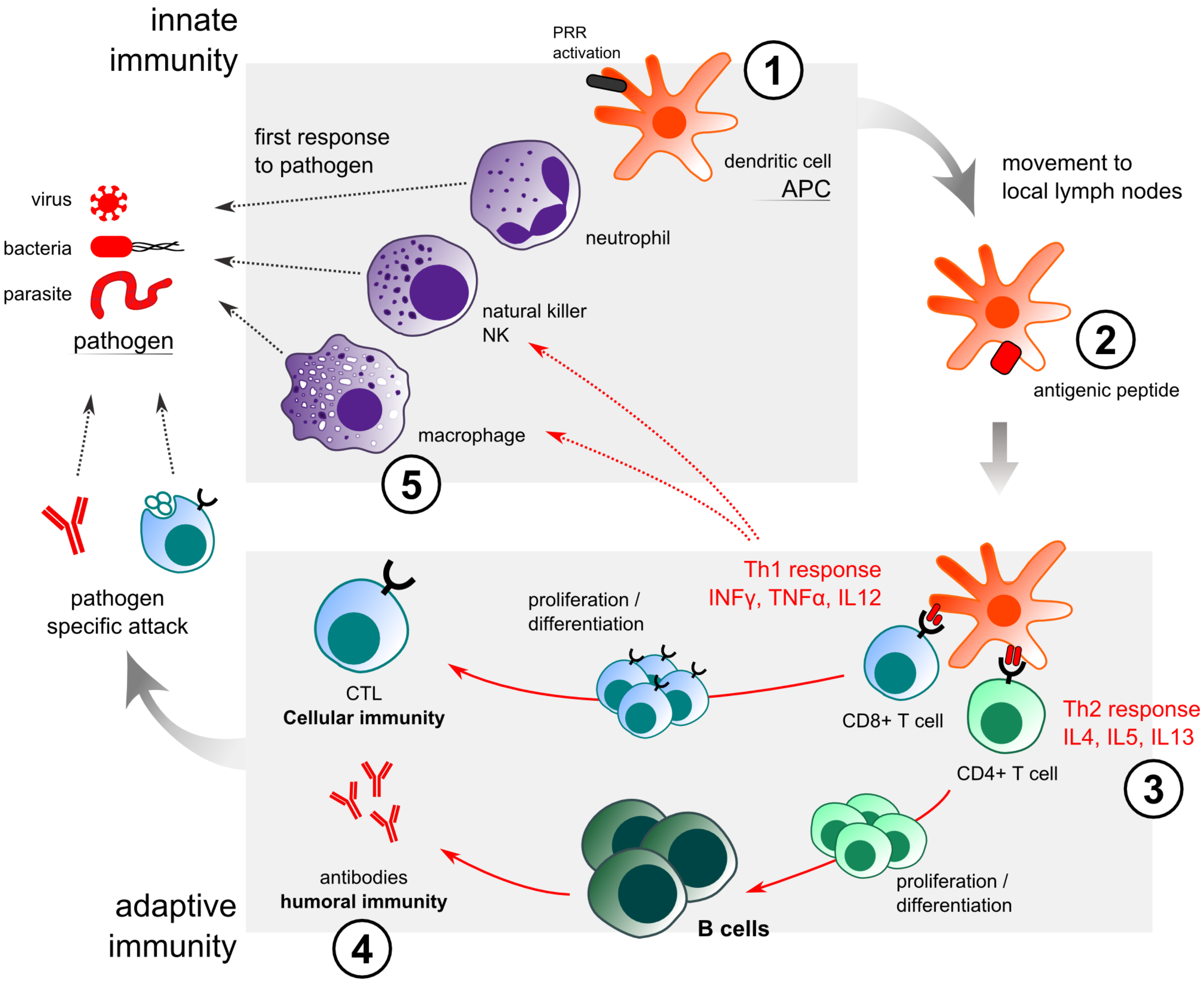
LSTM
LIVERPOOL SCHOOL
OF TROPICAL MEDICINE

Postdoc

Nasal mucosa immunity during
Streptococcus pneumoniae colonization

UNIVERSITY OF
OXFORD

Innate and immune response

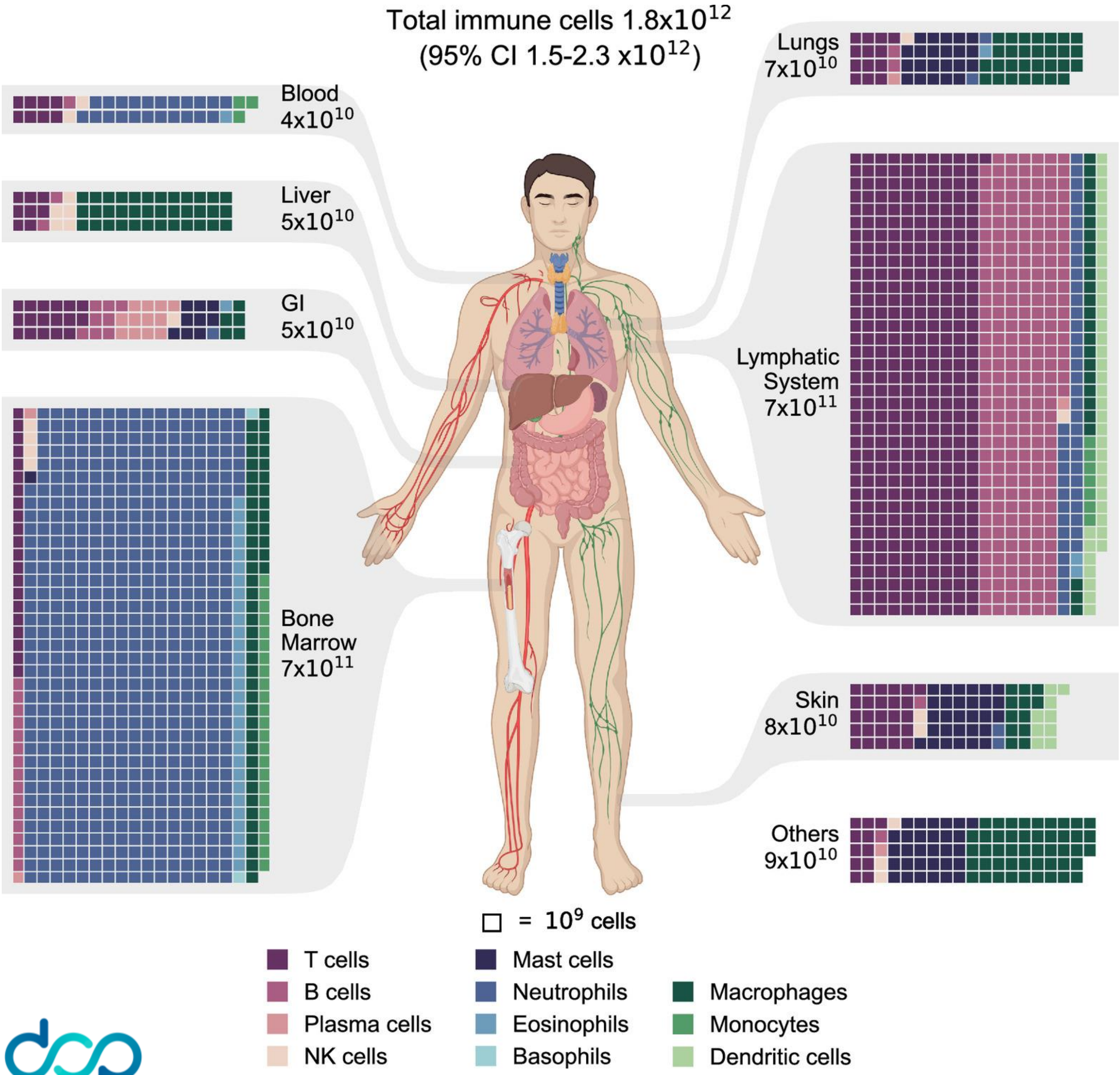


oxford vaccine group

DEPARTMENT OF PAEDIATRICS

UNIVERSITY OF OXFORD

Immune cells distributed in the human body




PNAS RESEARCH ARTICLE | SYSTEMS BIOLOGY OPEN ACCESS

The total mass, number, and distribution of immune cells in the human body

Ron Sender[a], Yarden Weiss[b], Yoav Navon[a], Idan Milo[c], Nofar Azulay[c], Leeat Keren[c], Shai Fuchs[d], Danny Ben-Zvi[e], Elad Noor[a], and Ron Milo[a,1]

Edited by David Baker, University of Washington, Seattle, WA; received May 21, 2023; accepted September 11, 2023

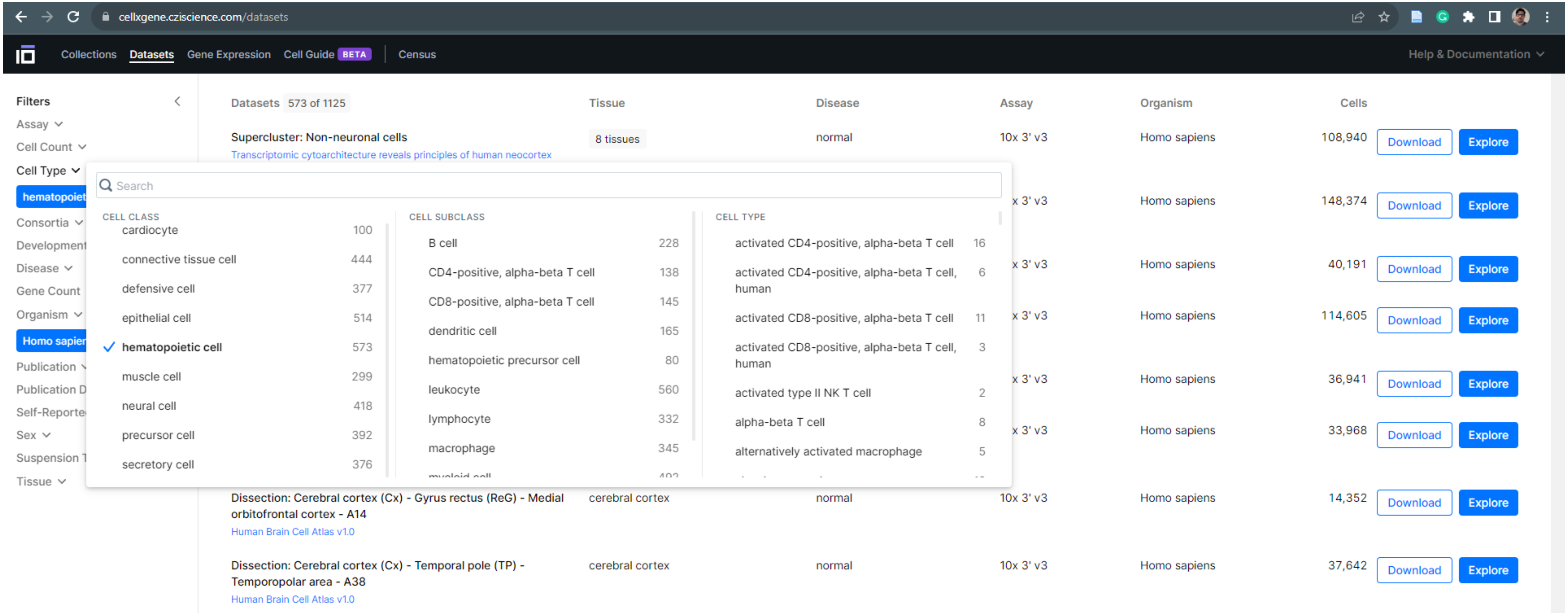
CELLxGENE database

Discover the mechanisms of human health

Download and visually explore reference-quality data to understand the functionality of human tissues at the cellular level with Chan Zuckerberg CELL by GENE Discover (CZ CELLxGENE Discover).

UNIQUE CELLS	DATASETS	CELL TYPES
57.4M	1125	738

CELLxGENE database



cellxgene.cziscience.com/datasets

Filters

Datasets 573 of 1125	Tissue	Disease	Assay	Organism	Cells		
Supercluster: Non-neuronal cells Transcriptomic cytoarchitecture reveals principles of human neocortex	8 tissues	normal	10x 3' v3	Homo sapiens	108,940	Download	Explore
			x 3' v3	Homo sapiens	148,374	Download	Explore
			x 3' v3	Homo sapiens	40,191	Download	Explore
			x 3' v3	Homo sapiens	114,605	Download	Explore
			x 3' v3	Homo sapiens	36,941	Download	Explore
			x 3' v3	Homo sapiens	33,968	Download	Explore
Dissection: Cerebral cortex (Cx) - Gyrus rectus (ReG) - Medial orbitofrontal cortex - A14 Human Brain Cell Atlas v1.0	cerebral cortex	normal	10x 3' v3	Homo sapiens	14,352	Download	Explore
Dissection: Cerebral cortex (Cx) - Temporal pole (TP) - Temporopolar area - A38 Human Brain Cell Atlas v1.0	cerebral cortex	normal	10x 3' v3	Homo sapiens	37,642	Download	Explore

Search

CELL CLASS	
cardiocyte	100
connective tissue cell	444
defensive cell	377
epithelial cell	514
✓ hematopoietic cell	573
muscle cell	299
neural cell	418
precursor cell	392
secretory cell	376

CELL SUBCLASS	
B cell	228
CD4-positive, alpha-beta T cell	138
CD8-positive, alpha-beta T cell	145
dendritic cell	165
hematopoietic precursor cell	80
leukocyte	560
lymphocyte	332
macrophage	345

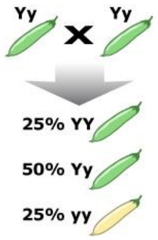
CELL TYPE	
activated CD4-positive, alpha-beta T cell	16
activated CD4-positive, alpha-beta T cell, human	6
activated CD8-positive, alpha-beta T cell	11
activated CD8-positive, alpha-beta T cell, human	3
activated type II NK T cell	2
alpha-beta T cell	8
alternatively activated macrophage	5

oxford vaccine group

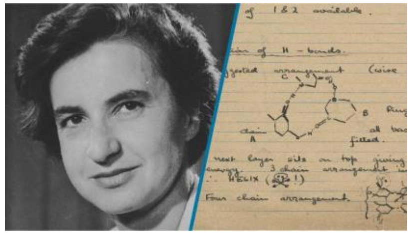
DEPARTMENT OF PAEDIATRICS

UNIVERSITY OF OXFORD

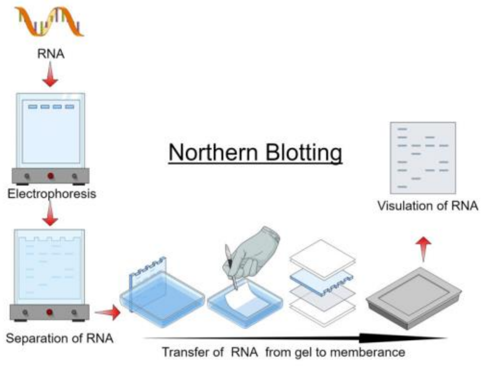
Evolution of Gene Expression Analysis: A Century of Discovery

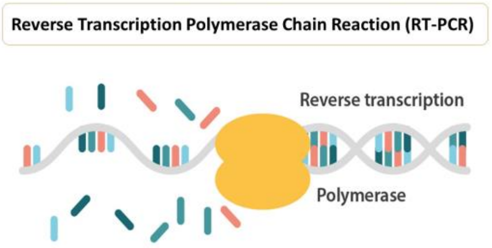



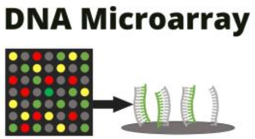
Gregor Mendel



Rosalind Franklin







1900s

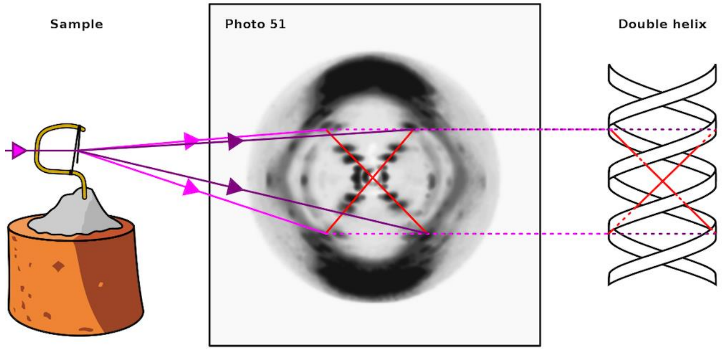
1940s-1950s

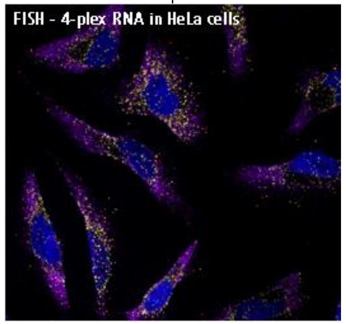
1970s

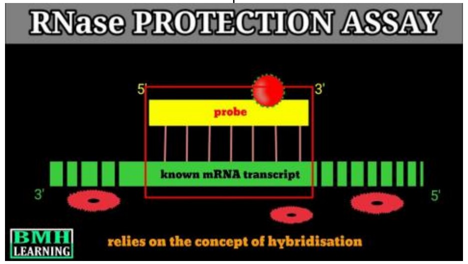
1980s

1990s

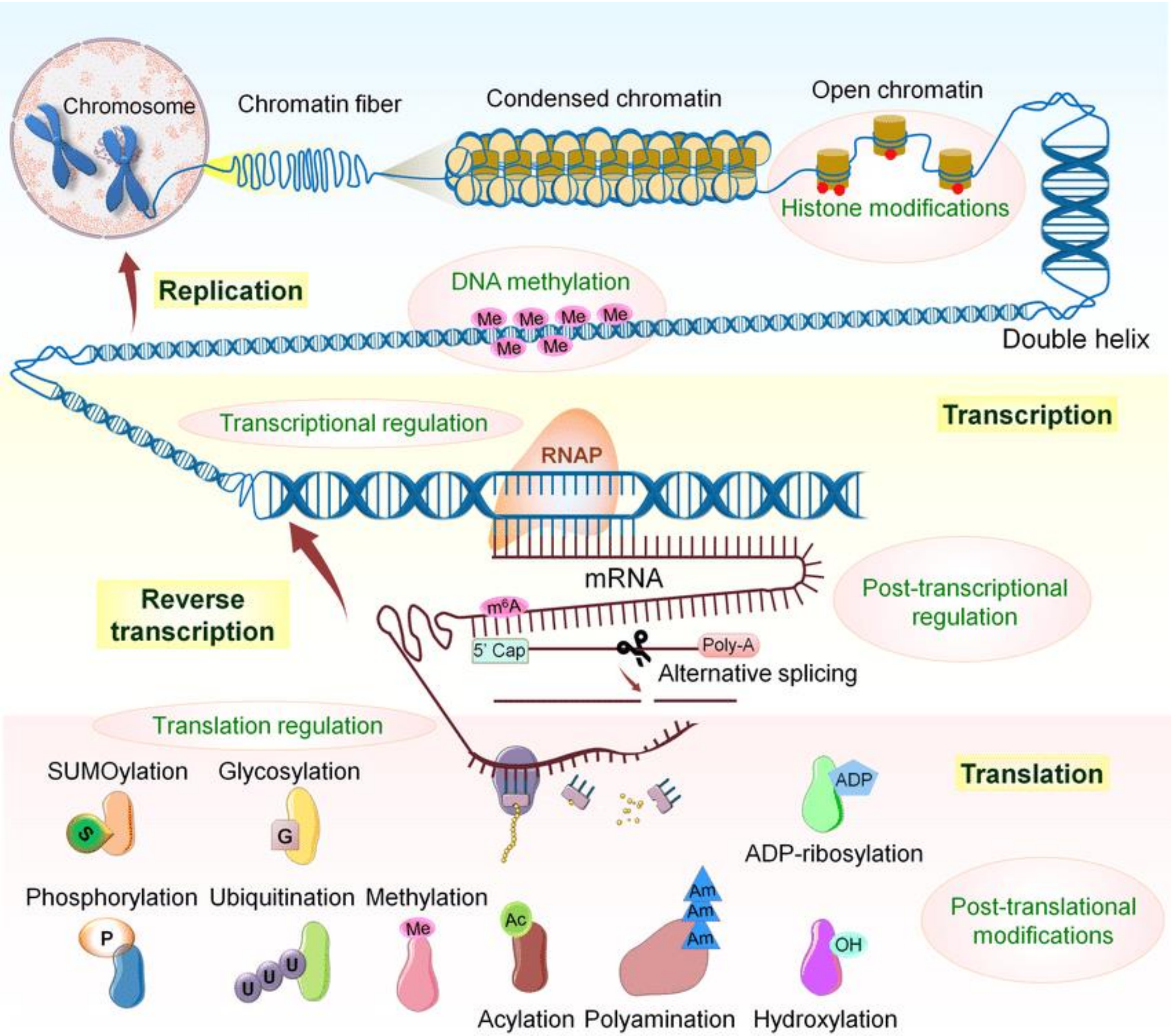
2000s







Overview of the Central Dogma of Molecular Biology and Processes Involved the Flow of Genetic Information



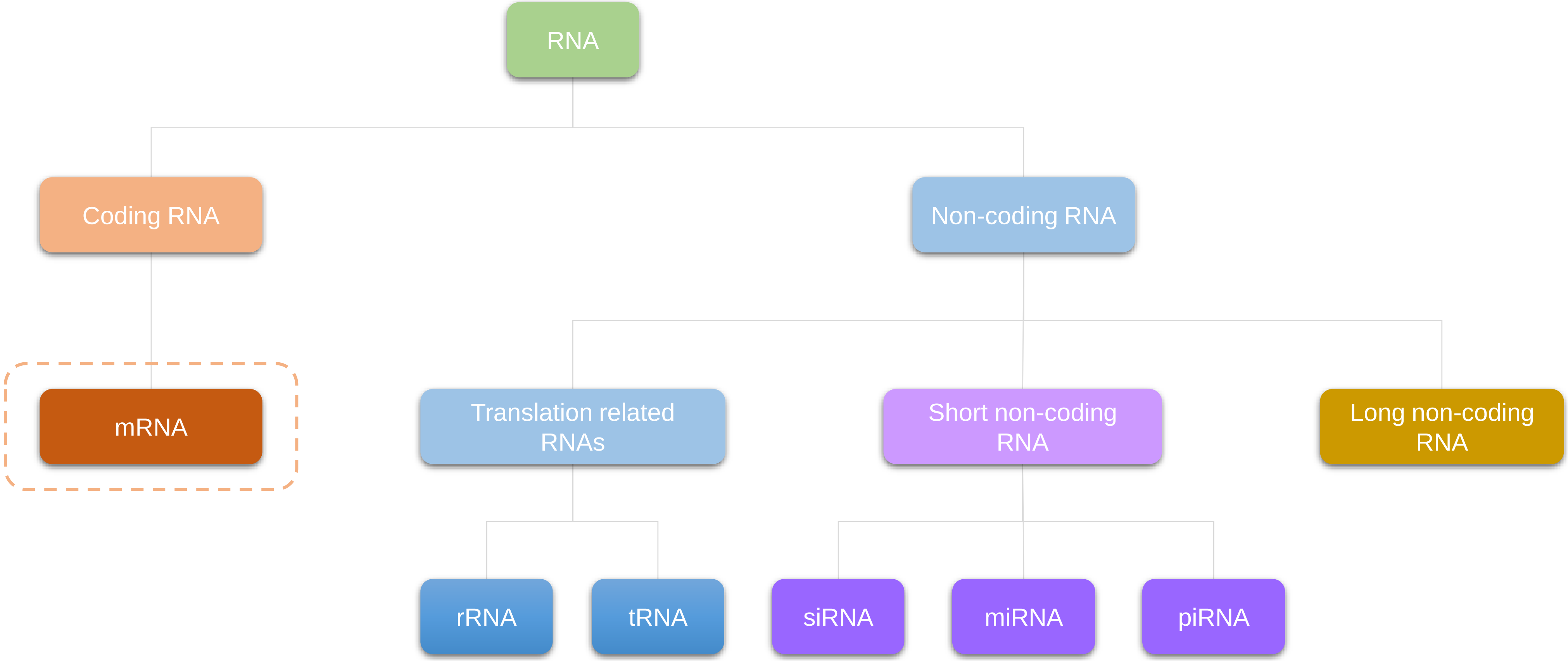
Pramanik et al., 2020
DOI: 10.1016/j.molp.2020.11.002

oxford vaccine group

DEPARTMENT OF PAEDIATRICS

UNIVERSITY OF OXFORD

Coding and non-coding RNAs



RNA

Coding RNA

Non-coding RNA

mRNA

Translation related RNAs

Short non-coding RNA

Long non-coding RNA

rRNA

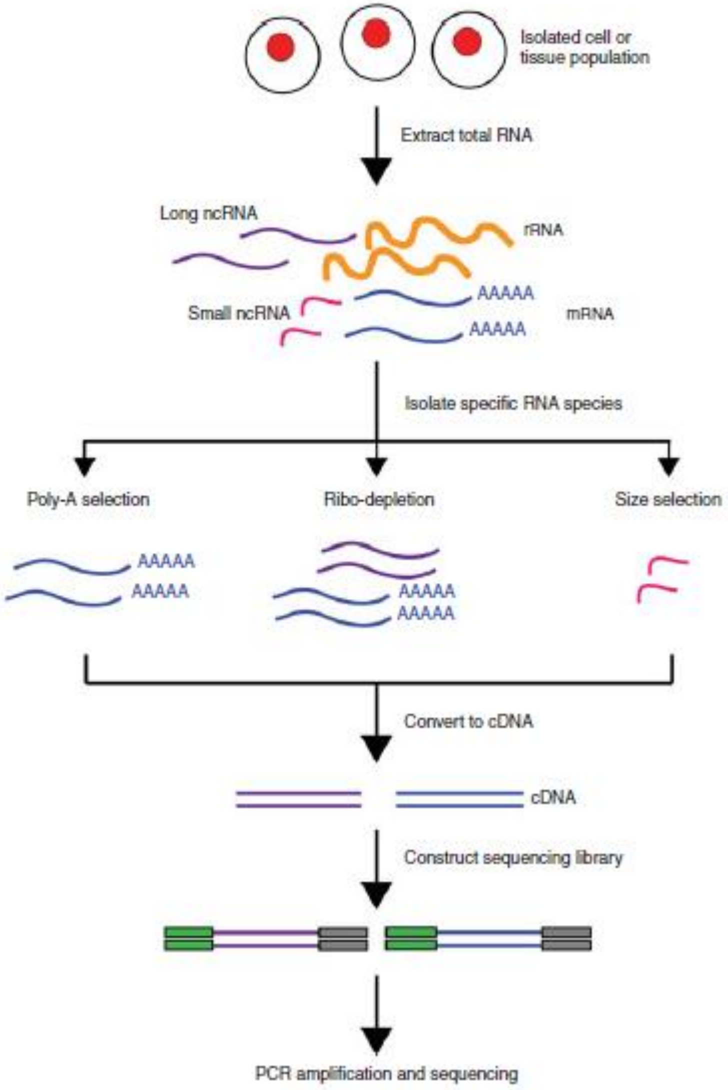
tRNA

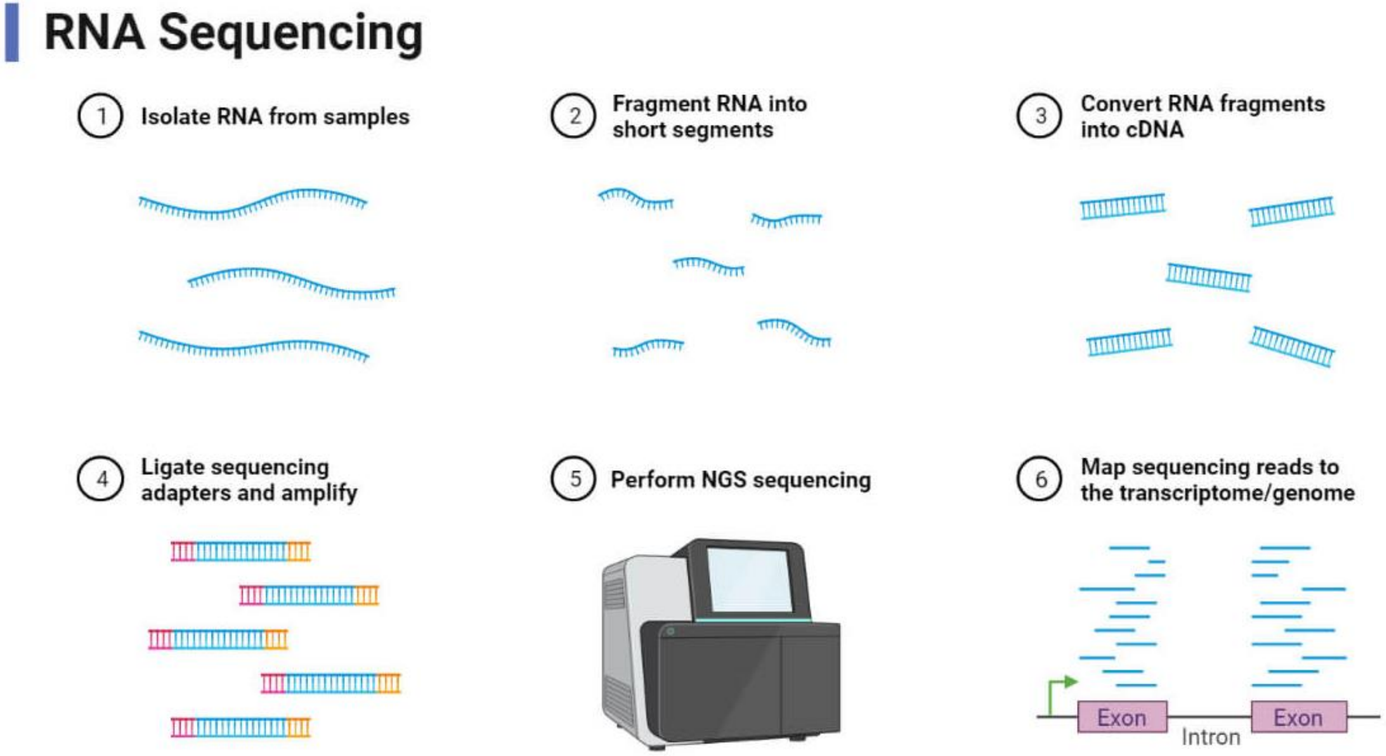
siRNA

miRNA

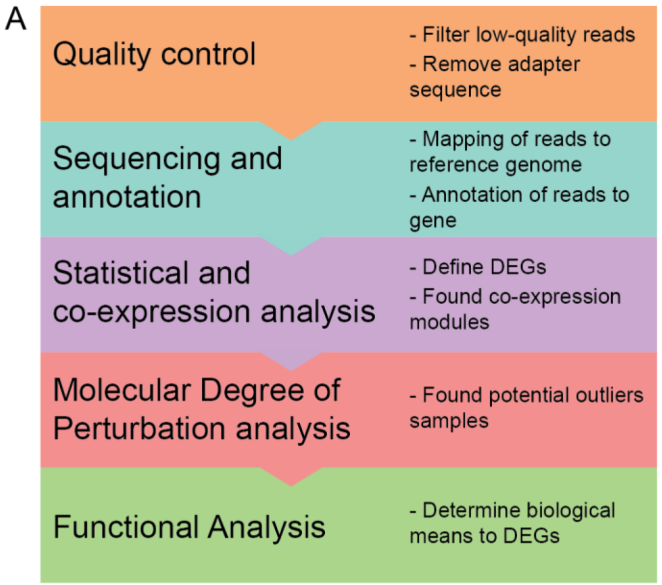
piRNA

How to select and sequence specific RNA classes?





Transcriptome pipeline of analysis



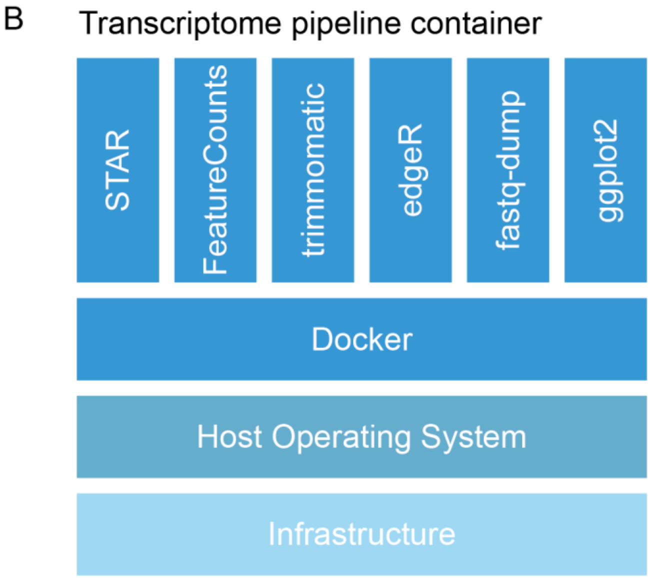
Quality control
- Filter low-quality reads
- Remove adapter sequence

Sequencing and annotation
- Mapping of reads to reference genome
- Annotation of reads to gene

Statistical and co-expression analysis
- Define DEGs
- Found co-expression modules

Molecular Degree of Perturbation analysis
- Found potential outliers samples

Functional Analysis
- Determine biological means to DEGs



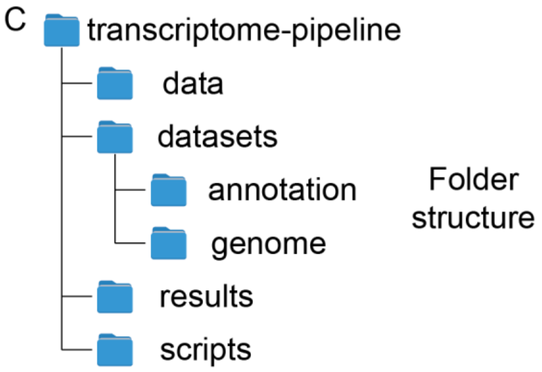
Transcriptome pipeline container

STAR
FeatureCounts
trimmomatic
edgeR
fastq-dump
ggplot2

Docker

Host Operating System

Infrastructure



transcriptome-pipeline
- data
- datasets
  - annotation
  - genome
- results
- scripts

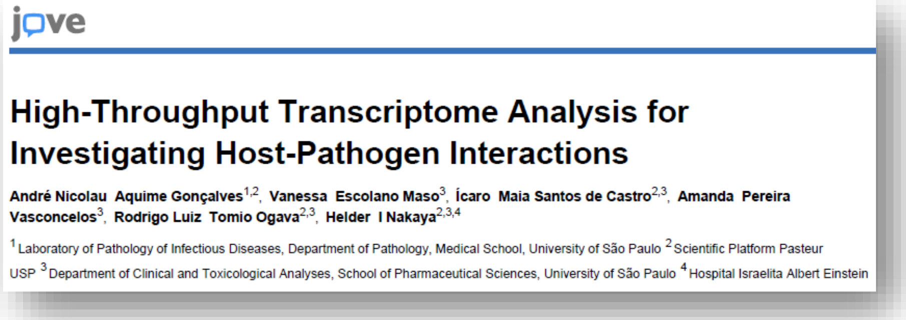
Folder structure

Transcriptome libraries from Soares-Schanoski A et al., 2019

x20 healthy people

x39 CHIKV-infected people

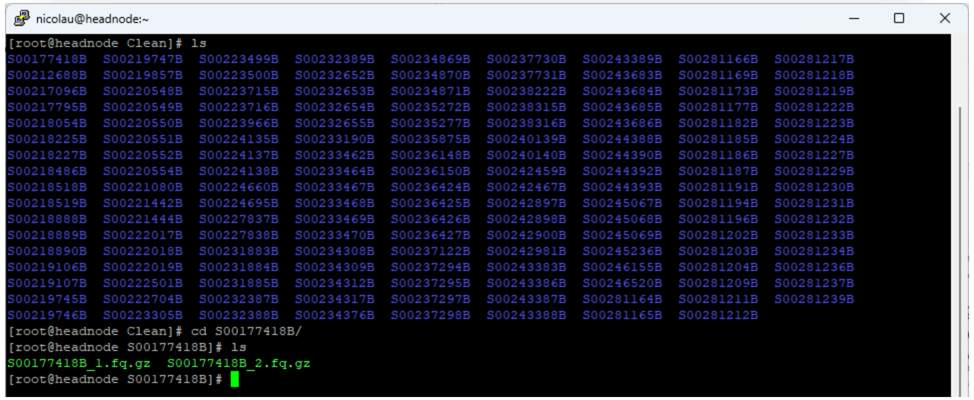
Blood samples were collected and sequencing was done on the Illumina platform



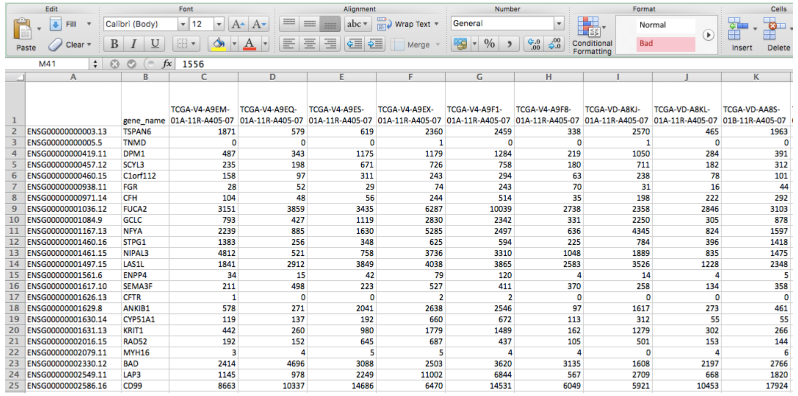
High-Throughput Transcriptome Analysis for Investigating Host-Pathogen Interactions

André Nicolau Aquime Gonçalves[1,2], **Vanessa Escolano Maso**[3], **Ícaro Maia Santos de Castro**[2,3], **Amanda Pereira Vasconcelos**[3], **Rodrigo Luiz Tomio Ogava**[2,3], **Helder I Nakaya**[2,3,4]

[1] Laboratory of Pathology of Infectious Diseases, Department of Pathology, Medical School, University of São Paulo [2] Scientific Platform Pasteur USP [3] Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo [4] Hospital Israelita Albert Einstein



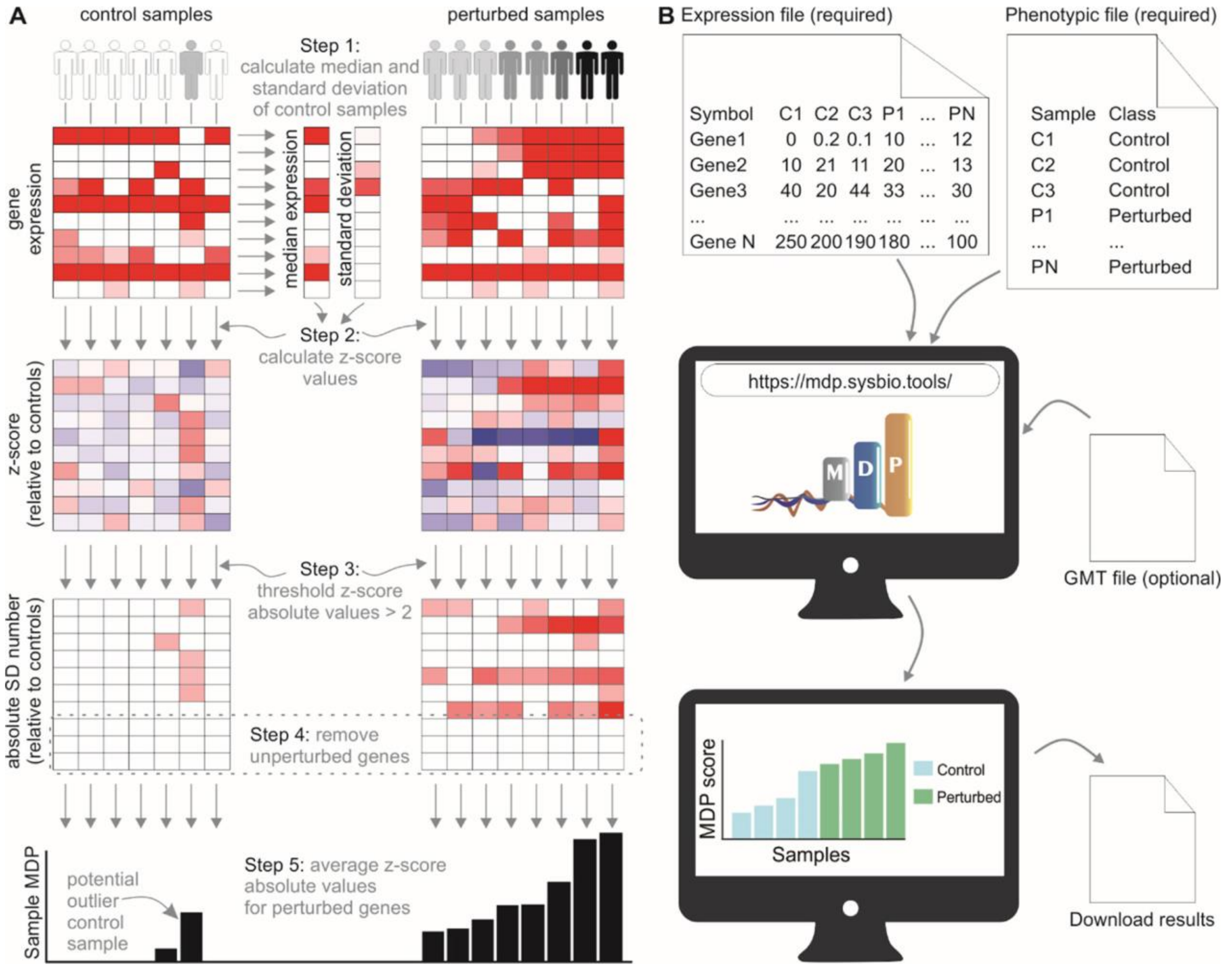
```
[root@headnode Clean]# ls
S00177418B  S00219747B  S00223499B  S00232389B  S00234869B  S00237730B  S00243389B  S00281166B  S00281217B
S00212688B  S00219857B  S00223500B  S00232652B  S00234870B  S00237731B  S00243683B  S00281169B  S00281218B
S00217096B  S00220548B  S00223715B  S00232653B  S00234871B  S00238222B  S00243684B  S00281173B  S00281219B
S00217795B  S00220549B  S00223716B  S00232654B  S00235272B  S00238315B  S00243685B  S00281177B  S00281222B
S00218054B  S00220550B  S00223966B  S00232655B  S00235277B  S00238316B  S00243686B  S00281182B  S00281223B
S00218225B  S00220551B  S00224135B  S00233190B  S00235875B  S00240139B  S00244388B  S00281185B  S00281224B
S00218227B  S00220552B  S00224137B  S00233462B  S00236148B  S00240140B  S00244390B  S00281186B  S00281227B
S00218486B  S00220554B  S00224138B  S00233464B  S00236150B  S00242459B  S00244392B  S00281187B  S00281229B
S00218518B  S00221080B  S00224660B  S00233467B  S00236424B  S00242467B  S00244393B  S00281191B  S00281230B
S00218519B  S00221442B  S00224695B  S00233468B  S00236425B  S00242897B  S00245067B  S00281194B  S00281231B
S00218888B  S00221444B  S00227837B  S00233469B  S00236426B  S00242898B  S00245068B  S00281196B  S00281232B
S00218889B  S00222017B  S00227838B  S00233470B  S00236427B  S00242900B  S00245069B  S00281202B  S00281233B
S00218890B  S00222018B  S00231883B  S00234308B  S00237122B  S00242981B  S00245236B  S00281203B  S00281234B
S00219106B  S00222019B  S00231884B  S00234309B  S00237294B  S00243383B  S00246155B  S00281204B  S00281236B
S00219107B  S00222501B  S00231885B  S00234312B  S00237295B  S00243386B  S00246520B  S00281209B  S00281237B
S00219745B  S00222704B  S00232387B  S00234317B  S00237297B  S00243387B  S00281164B  S00281211B  S00281239B
S00219746B  S00223305B  S00232388B  S00234376B  S00237298B  S00243388B  S00281165B  S00281212B
[root@headnode Clean]# cd S00177418B/
[root@headnode S00177418B]# ls
S00177418B_1.fq.gz  S00177418B_2.fq.gz
[root@headnode S00177418B]# 
```



M41 1556

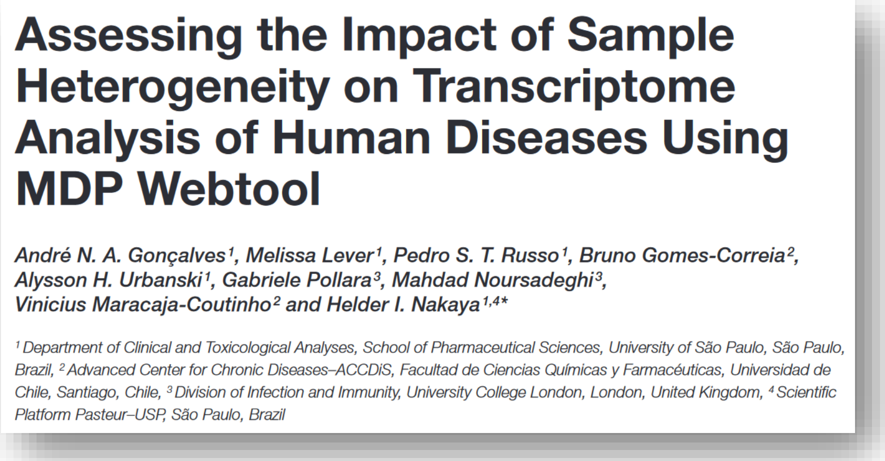
	A	B	C	D	E	F	G	H	I	J	K
1		gene_name	TCGA-V4-A9EM-01A-11R-A405-07	TCGA-V4-A9EQ-01A-11R-A405-07	TCGA-V4-A9ES-01A-11R-A405-07	TCGA-V4-A9EX-01A-11R-A405-07	TCGA-V4-A9F1-01A-11R-A405-07	TCGA-V4-A9F8-01A-11R-A405-07	TCGA-VD-A8KJ-01A-11R-A405-07	TCGA-VD-A8KL-01A-11R-A405-07	TCGA-VD-AA8S-01B-11R-A405-07
2	ENSG00000000003.13	TSPAN6	1871	579	619	2360	2459	338	2570	465	1963
3	ENSG00000000005.5	TNMD	0	0	0	1	0	0	1	0	0
4	ENSG00000000419.11	DPM1	487	343	1175	1179	1284	219	1050	284	391
5	ENSG00000000457.12	SCYL3	235	198	671	726	758	180	711	182	312
6	ENSG00000000460.15	C1orf112	158	97	311	243	294	63	238	78	101
7	ENSG00000000938.11	FGR	28	52	29	74	243	70	31	16	44
8	ENSG00000000971.14	CFH	104	48	56	244	514	35	198	222	292
9	ENSG00000001036.12	FUCA2	3151	3859	3435	6287	10039	2738	2358	2846	3103
10	ENSG00000001084.9	GCLC	793	427	1119	2830	2342	331	2250	305	878
11	ENSG00000001167.13	NFYA	2239	885	1630	5285	2497	636	4345	824	1597
12	ENSG00000001460.16	STPG1	1383	256	348	625	594	225	784	396	1418
13	ENSG00000001461.15	NIPAL3	4812	521	758	3736	3310	1048	1889	835	1475
14	ENSG00000001497.15	LAS1L	1841	2912	3849	4038	3865	2583	3526	1228	2348
15	ENSG00000001561.6	ENPP4	34	15	42	79	120	4	14	4	5
16	ENSG00000001617.10	SEMA3F	211	498	223	527	411	370	258	134	358
17	ENSG00000001626.13	CFTR	1	0	0	2	2	0	0	0	0
18	ENSG00000001629.8	ANKIB1	578	271	2041	2638	2546	97	1617	273	461
19	ENSG00000001630.14	CYP51A1	119	137	192	660	672	113	312	55	55
20	ENSG00000001631.13	KRIT1	442	260	980	1779	1489	162	1279	302	266
21	ENSG00000002016.15	RAD52	192	152	645	687	437	105	501	153	144
22	ENSG00000002079.11	MYH16	3	4	5	5	4	4	0	4	6
23	ENSG00000002330.12	BAD	2414	4696	3088	2503	3620	3135	1608	2197	2766
24	ENSG00000002549.11	LAP3	1145	978	2249	11002	6844	567	2709	668	1820
25	ENSG00000002586.16	CD99	8663	10337	14686	6470	14531	6049	5921	10453	17924

Outlier identification based on class sample expression



Symbol	C1	C2	C3	P1	...	PN
Gene1	0	0.2	0.1	10	...	12
Gene2	10	21	11	20	...	13
Gene3	40	20	44	33	...	30
...	...	...	...	...	...	...
Gene N	250	200	190	180	...	100

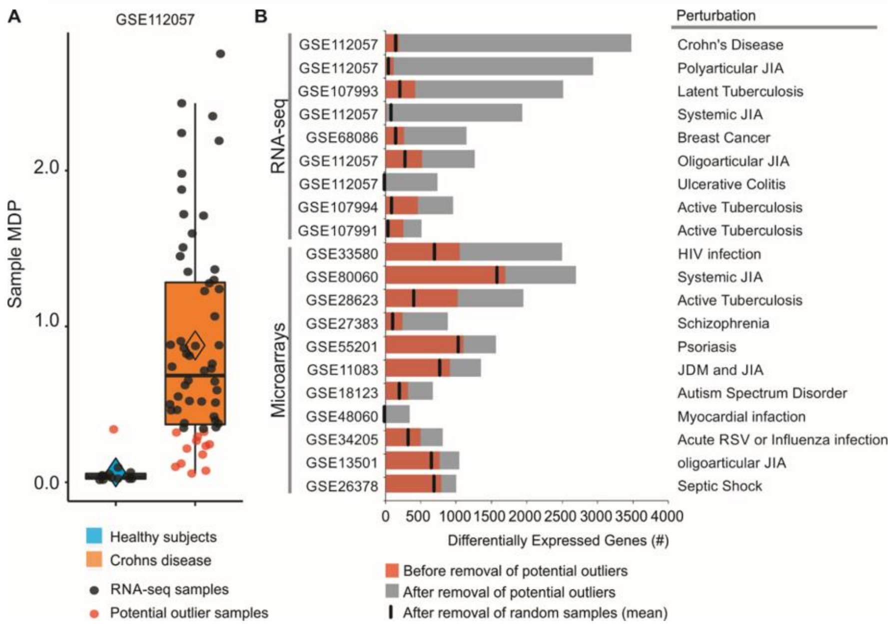
Sample	Class
C1	Control
C2	Control
C3	Control
P1	Perturbed
...	...
PN	Perturbed



Assessing the Impact of Sample Heterogeneity on Transcriptome Analysis of Human Diseases Using MDP Webtool

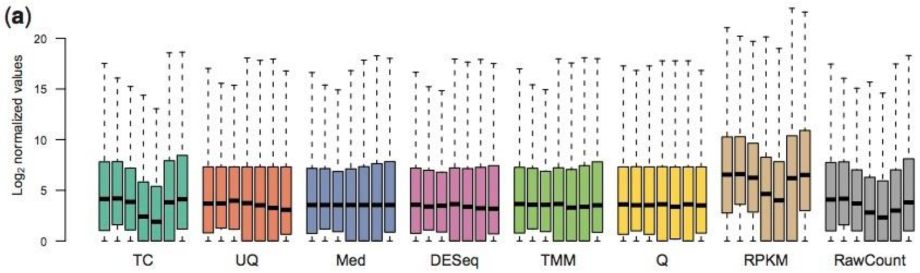
André N. A. Gonçalves[1], Melissa Lever[1], Pedro S. T. Russo[1], Bruno Gomes-Correia[2], Alysson H. Urbanski[1], Gabriele Pollara[3], Mahdad Noursadeghi[3], Vinicius Maracaja-Coutinho[2] and Helder I. Nakaya[1,4]*

[1] Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil, [2] Advanced Center for Chronic Diseases–ACCDiS, Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, Santiago, Chile, [3] Division of Infection and Immunity, University College London, London, United Kingdom, [4] Scientific Platform Pasteur–USP, São Paulo, Brazil

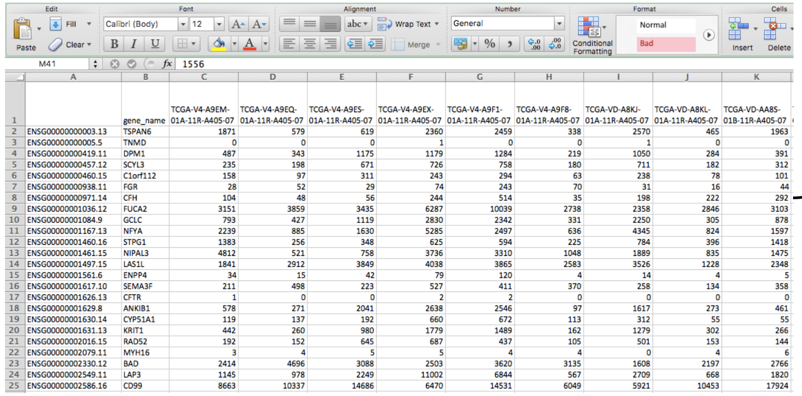


$$Z = \frac{x - \mu}{\sigma}$$

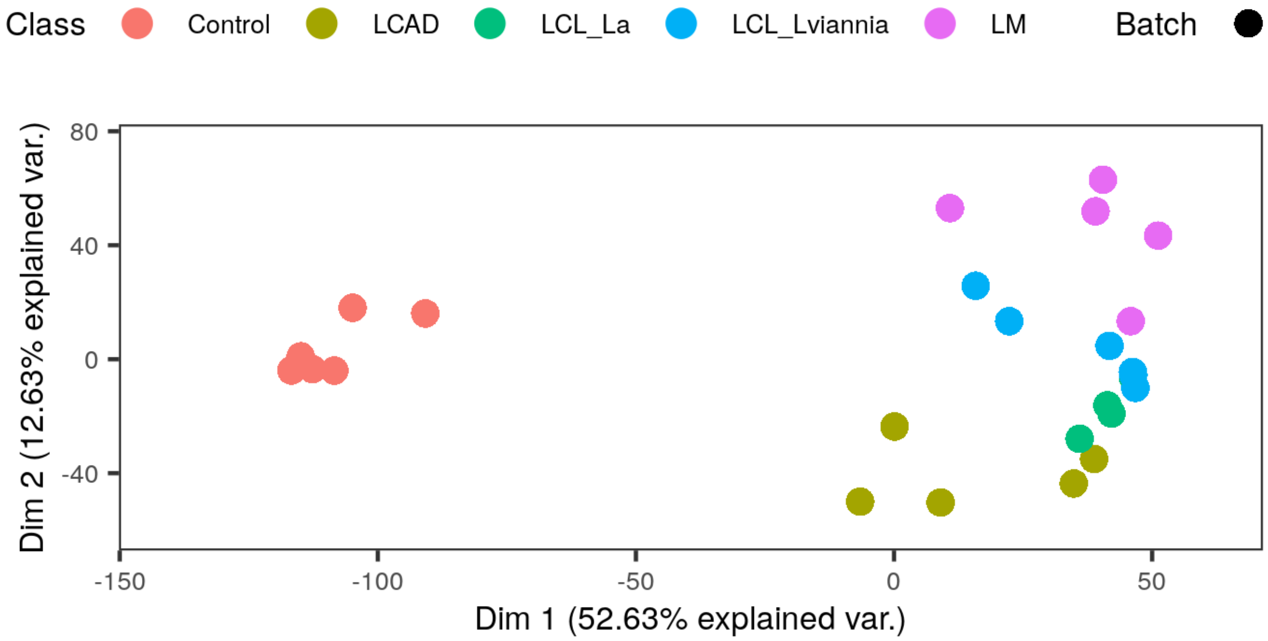
Principal Component Analysis (PCA)



Normalized gene expression matrix

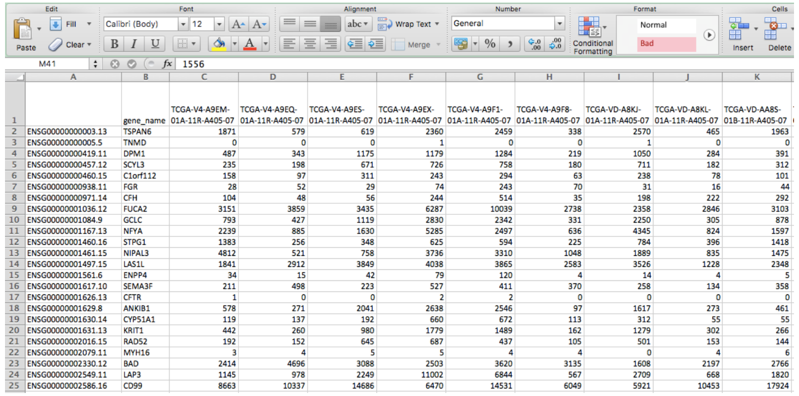


	A	B	C	D	E	F	G	H	I	J	K
1		gene_name	TCGA-V4-A9EM-01A-11R-A405-07	TCGA-V4-A9EQ-01A-11R-A405-07	TCGA-V4-A9ES-01A-11R-A405-07	TCGA-V4-A9EX-01A-11R-A405-07	TCGA-V4-A9F1-01A-11R-A405-07	TCGA-V4-A9F8-01A-11R-A405-07	TCGA-VD-A8KJ-01A-11R-A405-07	TCGA-VD-A8KL-01A-11R-A405-07	TCGA-VD-AA8S-01B-11R-A405-07
2	ENSG00000000003.13	TSPAN6	1871	579	619	2360	2459	338	2570	465	1963
3	ENSG00000000005.5	TNMD	0	0	0	1	0	0	1	0	0
4	ENSG00000000419.11	DPM1	487	343	1175	1179	1284	219	1050	284	391
5	ENSG00000000457.12	SCYL3	235	198	671	726	758	180	711	182	312
6	ENSG00000000460.15	C1orf112	158	97	311	243	294	63	238	78	101
7	ENSG00000000938.11	FGR	28	52	29	74	243	70	31	16	44
8	ENSG00000000971.14	CFH	104	48	56	244	514	35	198	222	292
9	ENSG00000001036.12	FUCA2	3151	3859	3435	6287	10039	2738	2358	2846	3103
10	ENSG00000001084.9	GCLC	793	427	1119	2830	2342	331	2250	305	878
11	ENSG00000001167.13	NFYA	2239	885	1630	5285	2497	636	4345	824	1597
12	ENSG00000001460.16	STPG1	1383	256	348	625	594	225	784	396	1418
13	ENSG00000001461.15	NIPAL3	4812	521	758	3736	3310	1048	1889	835	1475
14	ENSG00000001497.15	LAS1L	1841	2912	3849	4038	3865	2583	3526	1228	2348
15	ENSG00000001561.6	ENPP4	34	15	42	79	120	4	14	4	5
16	ENSG00000001617.10	SEMA3F	211	498	223	527	411	370	258	134	358
17	ENSG00000001626.13	CFTR	1	0	0	2	2	0	0	0	0
18	ENSG00000001629.8	ANKIB1	578	271	2041	2638	2546	97	1617	273	461
19	ENSG00000001630.14	CYP51A1	119	137	192	660	672	113	312	55	55
20	ENSG00000001631.13	KRIT1	442	260	980	1779	1489	162	1279	302	266
21	ENSG00000002016.15	RAD52	192	152	645	687	437	105	501	153	144
22	ENSG00000002079.11	MYH16	3	4	5	5	4	4	0	4	6
23	ENSG00000002330.12	BAD	2414	4696	3088	2503	3620	3135	1608	2197	2766
24	ENSG00000002549.11	LAP3	1145	978	2249	11002	6844	567	2709	668	1820
25	ENSG00000002586.16	CD99	8663	10337	14686	6470	14531	6049	5921	10453	17924



Batch Effect

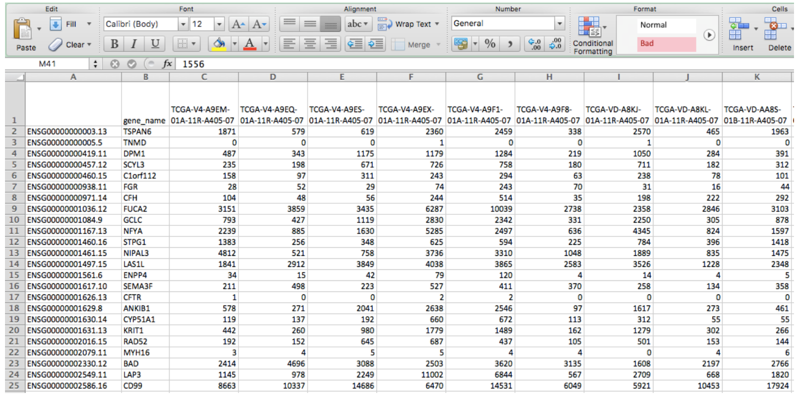
Normalized gene expression matrix



	A	B	C	D	E	F	G	H	I	J	K
1		gene_name	TCGA-V4-A9EM-01A-11R-A405-07	TCGA-V4-A9EQ-01A-11R-A405-07	TCGA-V4-A9ES-01A-11R-A405-07	TCGA-V4-A9EX-01A-11R-A405-07	TCGA-V4-A9F1-01A-11R-A405-07	TCGA-V4-A9F8-01A-11R-A405-07	TCGA-VD-A8KJ-01A-11R-A405-07	TCGA-VD-A8KL-01A-11R-A405-07	TCGA-VD-AA8S-01B-11R-A405-07
2	ENSG00000000003.13	TSPAN6	1871	579	619	2360	2459	338	2570	465	1963
3	ENSG00000000005.5	TNMD	0	0	0	1	0	0	1	0	0
4	ENSG00000000419.11	DPM1	487	343	1175	1179	1284	219	1050	284	391
5	ENSG00000000457.12	SCYL3	235	198	671	726	758	180	711	182	312
6	ENSG00000000460.15	C1orf112	158	97	311	243	294	63	238	78	101
7	ENSG00000000938.11	FGR	28	52	29	74	243	70	31	16	44
8	ENSG00000000971.14	CFH	104	48	56	244	514	35	198	222	292
9	ENSG00000001036.12	FUCA2	3151	3859	3435	6287	10039	2738	2358	2846	3103
10	ENSG00000001084.9	GCLC	793	427	1119	2830	2342	331	2250	305	878
11	ENSG00000001167.13	NFYA	2239	885	1630	5285	2497	636	4345	824	1597
12	ENSG00000001460.16	STPG1	1383	256	348	625	594	225	784	396	1418
13	ENSG00000001461.15	NIPAL3	4812	521	758	3736	3310	1048	1889	835	1475
14	ENSG00000001497.15	LAS1L	1841	2912	3849	4038	3865	2583	3526	1228	2348
15	ENSG00000001561.6	ENPP4	34	15	42	79	120	4	14	4	5
16	ENSG00000001617.10	SEMA3F	211	498	223	527	411	370	258	134	358
17	ENSG00000001626.13	CFTR	1	0	0	2	2	0	0	0	0
18	ENSG00000001629.8	ANKIB1	578	271	2041	2638	2546	97	1617	273	461
19	ENSG00000001630.14	CYP51A1	119	137	192	660	672	113	312	55	55
20	ENSG00000001631.13	KRIT1	442	260	980	1779	1489	162	1279	302	266
21	ENSG00000002016.15	RAD52	192	152	645	687	437	105	501	153	144
22	ENSG00000002079.11	MYH16	3	4	5	5	4	4	0	4	6
23	ENSG00000002330.12	BAD	2414	4696	3088	2503	3620	3135	1608	2197	2766
24	ENSG00000002549.11	LAP3	1145	978	2249	11002	6844	567	2709	668	1820
25	ENSG00000002586.16	CD99	8663	10337	14686	6470	14531	6049	5921	10453	17924

ComBat and ComBat-Seq

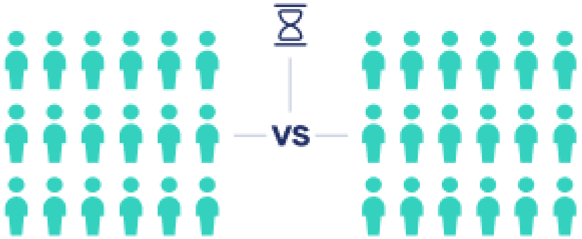
After batch effect



	A	B	C	D	E	F	G	H	I	J	K
1		gene_name	TCGA-V4-A9EM-01A-11R-A405-07	TCGA-V4-A9EQ-01A-11R-A405-07	TCGA-V4-A9ES-01A-11R-A405-07	TCGA-V4-A9EX-01A-11R-A405-07	TCGA-V4-A9F1-01A-11R-A405-07	TCGA-V4-A9F8-01A-11R-A405-07	TCGA-VD-A8KJ-01A-11R-A405-07	TCGA-VD-A8KL-01A-11R-A405-07	TCGA-VD-AA8S-01B-11R-A405-07
2	ENSG00000000003.13	TSPAN6	1871	579	619	2360	2459	338	2570	465	1963
3	ENSG00000000005.5	TNMD	0	0	0	1	0	0	1	0	0
4	ENSG00000000419.11	DPM1	487	343	1175	1179	1284	219	1050	284	391
5	ENSG00000000457.12	SCYL3	235	198	671	726	758	180	711	182	312
6	ENSG00000000460.15	C1orf112	158	97	311	243	294	63	238	78	101
7	ENSG00000000938.11	FGR	28	52	29	74	243	70	31	16	44
8	ENSG00000000971.14	CFH	104	48	56	244	514	35	198	222	292
9	ENSG00000001036.12	FUCA2	3151	3859	3435	6287	10039	2738	2358	2846	3103
10	ENSG00000001084.9	GCLC	793	427	1119	2830	2342	331	2250	305	878
11	ENSG00000001167.13	NFYA	2239	885	1630	5285	2497	636	4345	824	1597
12	ENSG00000001460.16	STPG1	1383	256	348	625	594	225	784	396	1418
13	ENSG00000001461.15	NIPAL3	4812	521	758	3736	3310	1048	1889	835	1475
14	ENSG00000001497.15	LAS1L	1841	2912	3849	4038	3865	2583	3526	1228	2348
15	ENSG00000001561.6	ENPP4	34	15	42	79	120	4	14	4	5
16	ENSG00000001617.10	SEMA3F	211	498	223	527	411	370	258	134	358
17	ENSG00000001626.13	CFTR	1	0	0	2	2	0	0	0	0
18	ENSG00000001629.8	ANKIB1	578	271	2041	2638	2546	97	1617	273	461
19	ENSG00000001630.14	CYP51A1	119	137	192	660	672	113	312	55	55
20	ENSG00000001631.13	KRIT1	442	260	980	1779	1489	162	1279	302	266
21	ENSG00000002016.15	RAD52	192	152	645	687	437	105	501	153	144
22	ENSG00000002079.11	MYH16	3	4	5	5	4	4	0	4	6
23	ENSG00000002330.12	BAD	2414	4696	3088	2503	3620	3135	1608	2197	2766
24	ENSG00000002549.11	LAP3	1145	978	2249	11002	6844	567	2709	668	1820
25	ENSG00000002586.16	CD99	8663	10337	14686	6470	14531	6049	5921	10453	17924

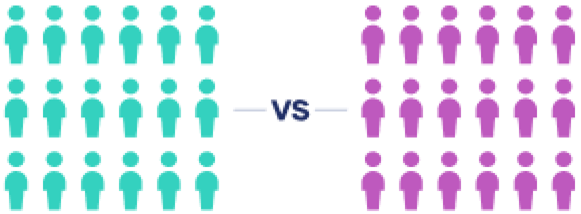
Differentially expressed genes (DEGs)

Paired-samples t test



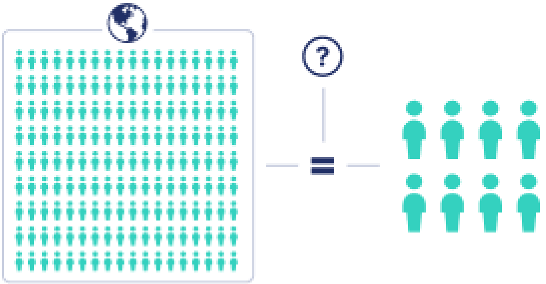
Investigate whether there's a difference within a group between two points in time (within-subjects).

Independent-samples t test

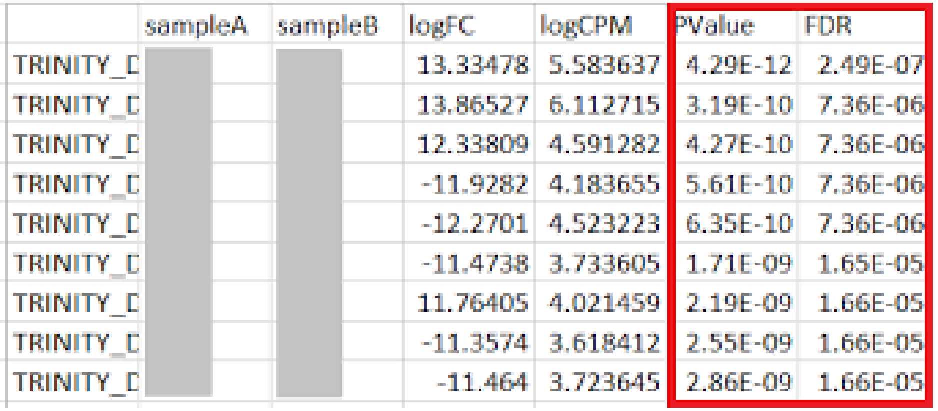


Investigate whether there's a difference between two groups (between-subjects).

One-sample t test

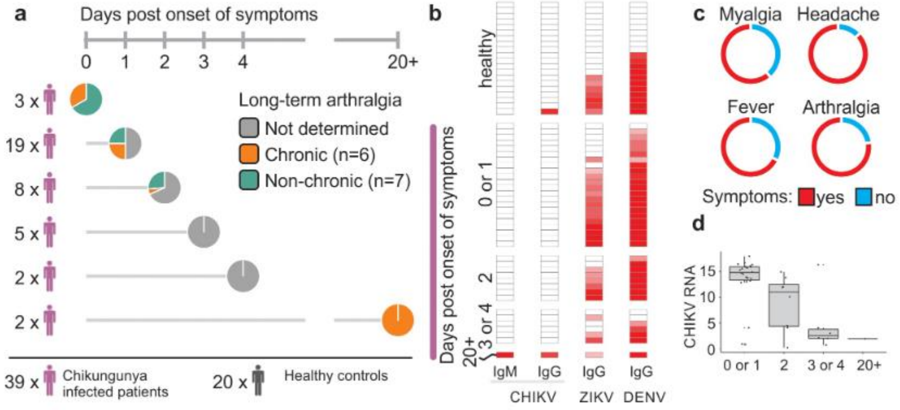


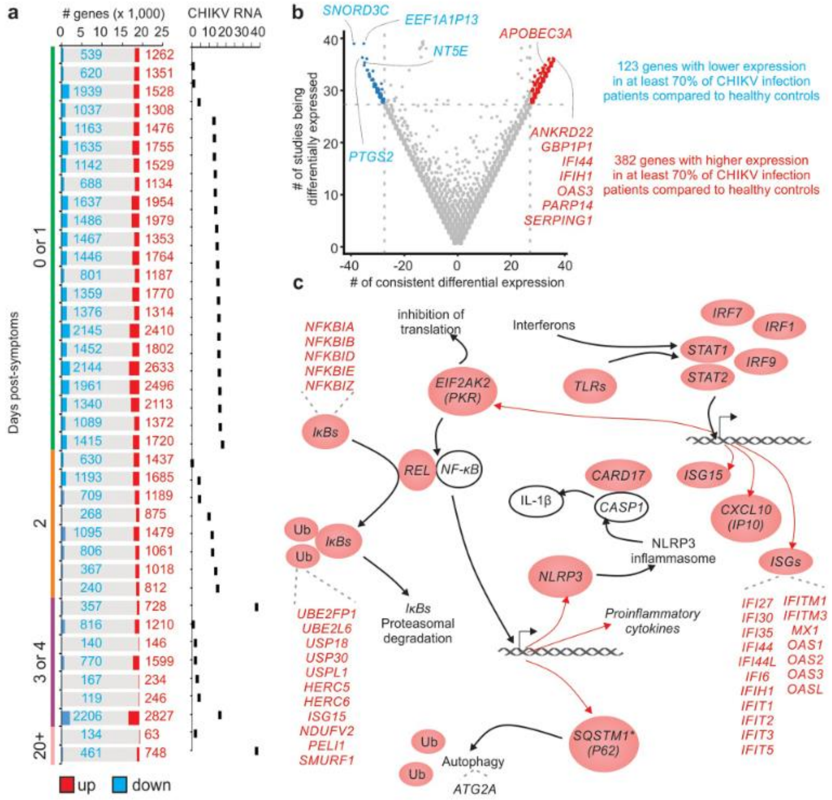
Investigate whether there's a difference between a group and a standard value or whether a subgroup belongs to a population.

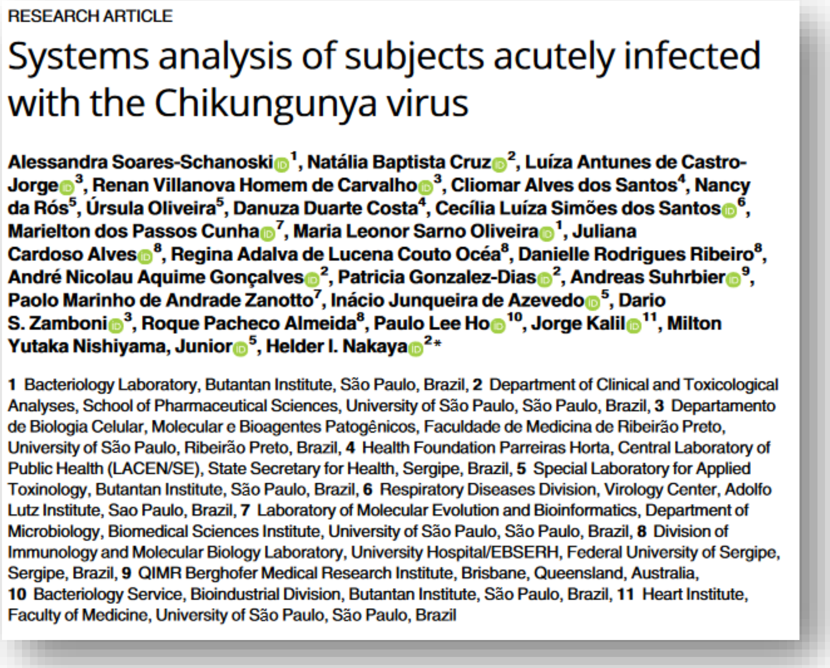


	sampleA	sampleB	logFC	logCPM	PValue	FDR
TRINITY_D	[redacted]	[redacted]	13.33478	5.583637	4.29E-12	2.49E-07
TRINITY_D	[redacted]	[redacted]	13.86527	6.112715	3.19E-10	7.36E-06
TRINITY_D	[redacted]	[redacted]	12.33809	4.591282	4.27E-10	7.36E-06
TRINITY_D	[redacted]	[redacted]	-11.9282	4.183655	5.61E-10	7.36E-06
TRINITY_D	[redacted]	[redacted]	-12.2701	4.523223	6.35E-10	7.36E-06
TRINITY_D	[redacted]	[redacted]	-11.4738	3.733605	1.71E-09	1.65E-05
TRINITY_D	[redacted]	[redacted]	11.76405	4.021459	2.19E-09	1.66E-05
TRINITY_D	[redacted]	[redacted]	-11.3574	3.618412	2.55E-09	1.66E-05
TRINITY_D	[redacted]	[redacted]	-11.464	3.723645	2.86E-09	1.66E-05

Host-CHIKV interaction potential therapeutic strategies of NLRP3 inflammasome



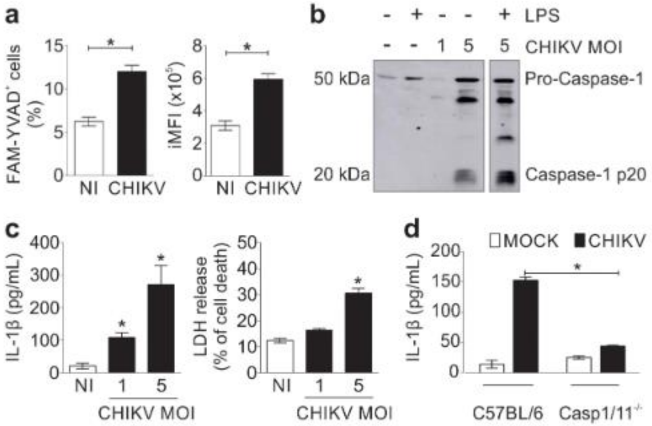



RESEARCH ARTICLE

Systems analysis of subjects acutely infected with the Chikungunya virus

Alessandra Soares-Schanoski[1], Natália Baptista Cruz[2], Luíza Antunes de Castro-Jorge[3], Renan Villanova Homem de Carvalho[3], Cliomar Alves dos Santos[4], Nancy da Rós[5], Úrsula Oliveira[5], Danuza Duarte Costa[4], Cecília Luíza Simões dos Santos[6], Marielton dos Passos Cunha[7], Maria Leonor Sarno Oliveira[1], Juliana Cardoso Alves[8], Regina Adalva de Lucena Couto Océa[8], Danielle Rodrigues Ribeiro[8], André Nicolau Aquime Gonçalves[2], Patricia Gonzalez-Dias[2], Andreas Suhrbier[9], Paolo Marinho de Andrade Zanotto[7], Inácio Junqueira de Azevedo[5], Dario S. Zamboni[3], Roque Pacheco Almeida[8], Paulo Lee Ho[10], Jorge Kalil[11], Milton Yutaka Nishiyama, Junior[5], Helder I. Nakaya[2]*

1 Bacteriology Laboratory, Butantan Institute, São Paulo, Brazil, 2 Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil, 3 Departamento de Biologia Celular, Molecular e Bioagentes Patogênicos, Faculdade de Medicina de Ribeirão Preto, University of São Paulo, Ribeirão Preto, Brazil, 4 Health Foundation Parreiras Horta, Central Laboratory of Public Health (LACEN/SE), State Secretary for Health, Sergipe, Brazil, 5 Special Laboratory for Applied Toxinology, Butantan Institute, São Paulo, Brazil, 6 Respiratory Diseases Division, Virology Center, Adolfo Lutz Institute, Sao Paulo, Brazil, 7 Laboratory of Molecular Evolution and Bioinformatics, Department of Microbiology, Biomedical Sciences Institute, University of São Paulo, São Paulo, Brazil, 8 Division of Immunology and Molecular Biology Laboratory, University Hospital/EBSERH, Federal University of Sergipe, Sergipe, Brazil, 9 QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia, 10 Bacteriology Service, Bioindustrial Division, Butantan Institute, São Paulo, Brazil, 11 Heart Institute, Faculty of Medicine, University of São Paulo, São Paulo, Brazil

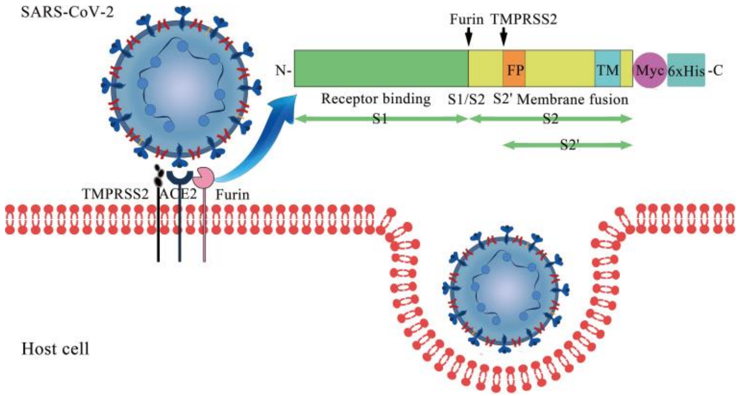


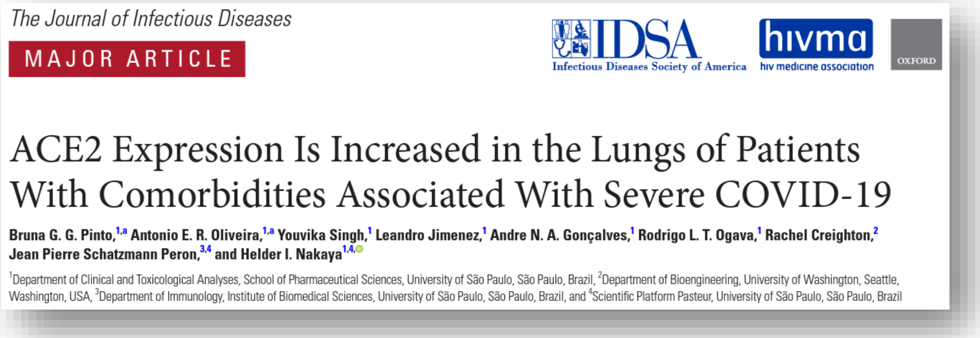
Hypothesis identified by RNA-analysis and validated in vitro or in vivo.

Is that possible to do the reverse path?

UNIVERSITY OF OXFORD

Comorbidities have increased expression of ACE2 in the lungs





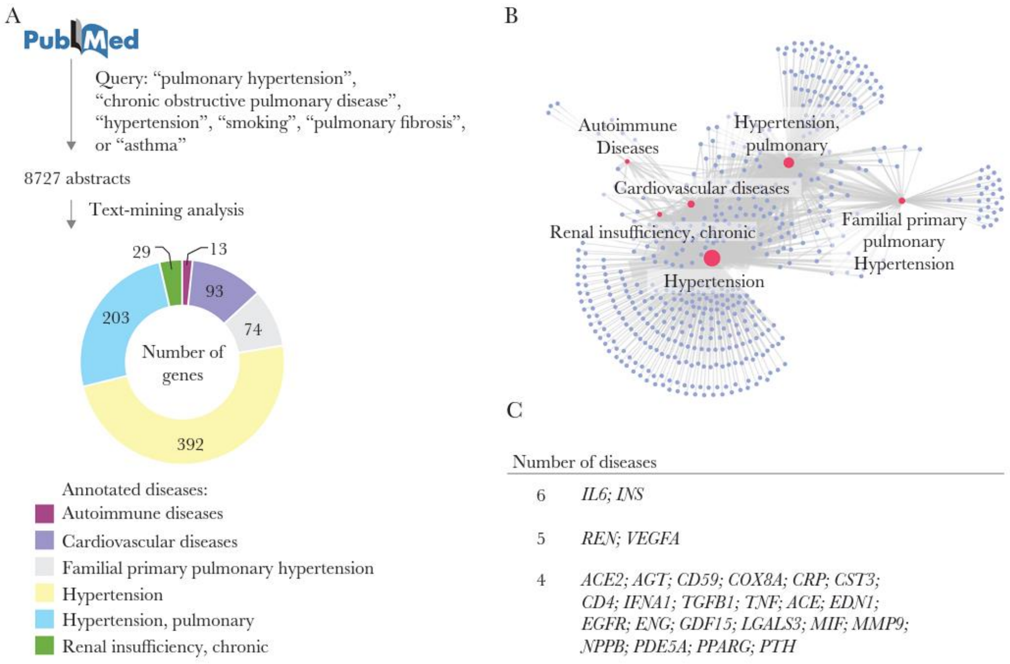
The Journal of Infectious Diseases

MAJOR ARTICLE

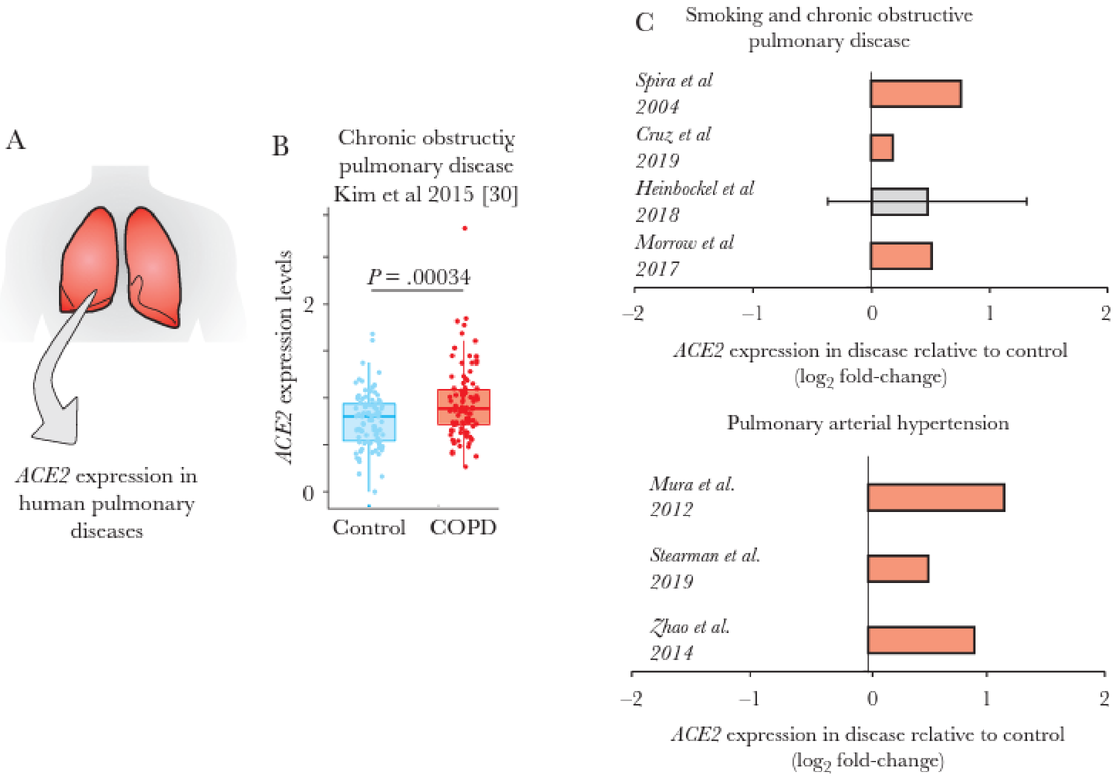
ACE2 Expression Is Increased in the Lungs of Patients With Comorbidities Associated With Severe COVID-19

Bruna G. G. Pinto,[1,a] Antonio E. R. Oliveira,[1,a] Youvika Singh,[1] Leandro Jimenez,[1] Andre N. A. Gonçalves,[1] Rodrigo L. T. Ogava,[1] Rachel Creighton,[2] Jean Pierre Schatzmann Peron,[3,4] and Helder I. Nakaya[1,4]

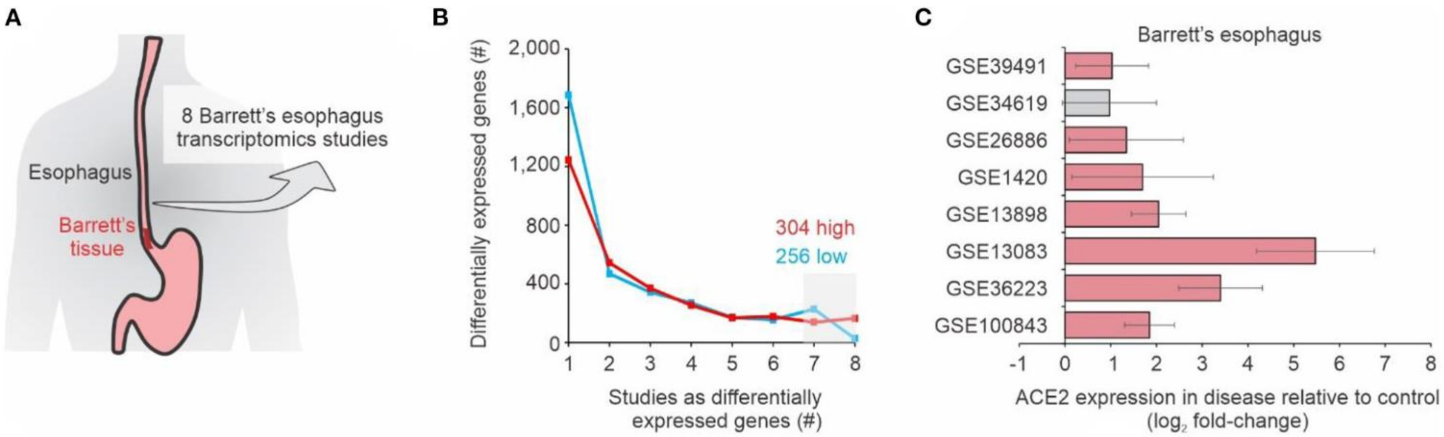
[1]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, São Paulo, Brazil, [2]Department of Bioengineering, University of Washington, Seattle, Washington, USA, [3]Department of Immunology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, Brazil, and [4]Scientific Platform Pasteur, University of São Paulo, São Paulo, Brazil



Number of diseases	
6	*IL6; INS*
5	*REN; VEGFA*
4	*ACE2; AGT; CD59; COX8A; CRP; CST3; CD4; IFNA1; TGFB1; TNF; ACE; EDN1; EGFR; ENG; GDF15; LGALS3; MIF; MMP9; NPPB; PDE5A; PPARG; PTH*

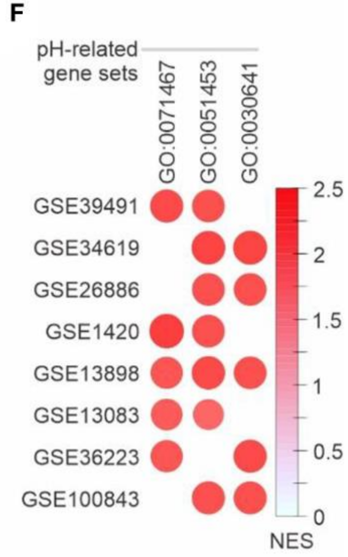


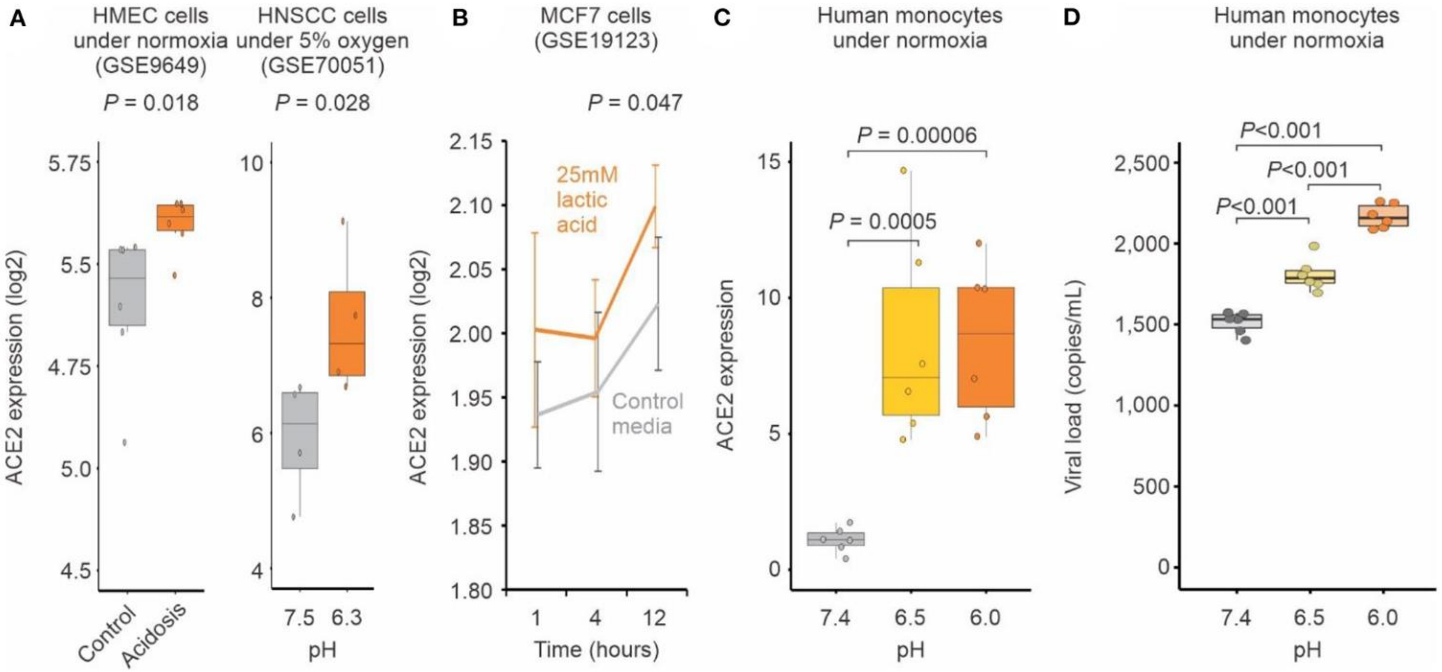
Risk of death by COVID-19 given by acid pH



Acid pH Increases SARS-CoV-2 Infection and the Risk of Death by COVID-19

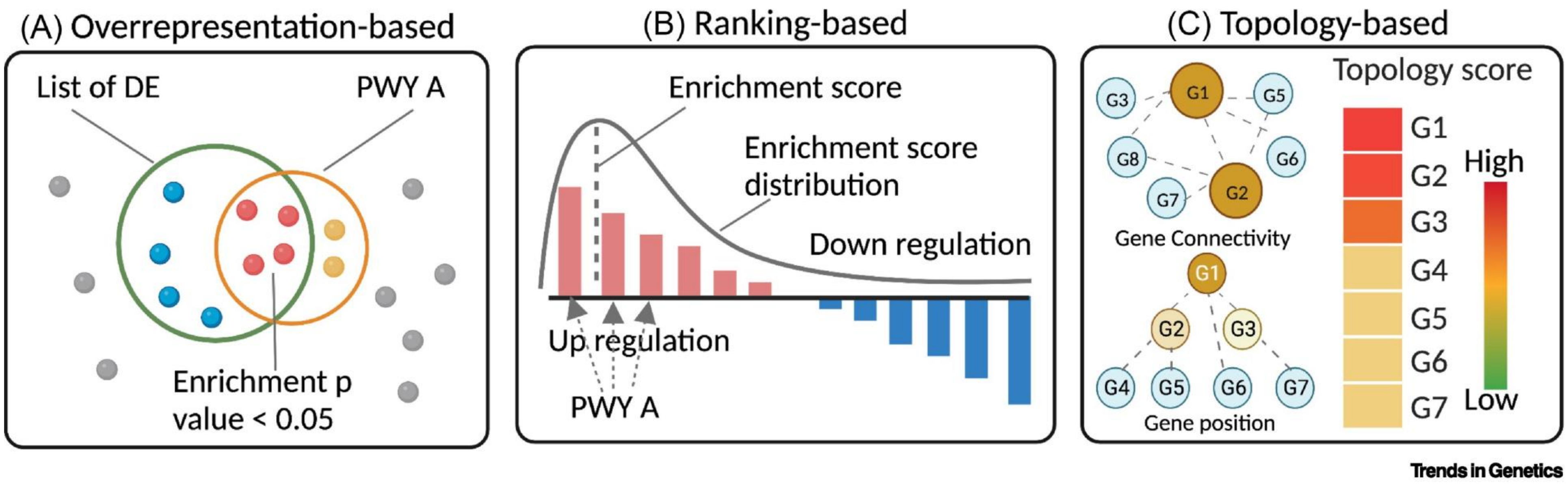
Leandro Jimenez[1,2], Ana Campos Codo[3], Vanderson de Souza Sampaio[4,5,6,7], Antonio E. R. Oliveira[1], Lucas Kaoru Kobo Ferreira[1], Gustavo Gastão Davanzo[3], Lauar de Brito Monteiro[3], João Victor Virgilio-da-Silva[3], Mayla Gabriela Silva Borba[4], Gabriela Fabiano de Souza[3], Nathalia Zini[8], Flora de Andrade Gandolfi[8], Stéfanie Primon Muraro[3], José Luiz Proença-Modena[3], Fernando Almeida Val[4,5,7], Gisely Cardoso Melo[4,5], Wuelton Marcelo Monteiro[4,5], Maurício Lacerda Nogueira[8], Marcus Vinícius Guimarães Lacerda[4,5,9], Pedro M. Moraes-Vieira[3,10,11] and Helder I. Nakaya[1,2,12]*





What means enrichment pathway analysis?

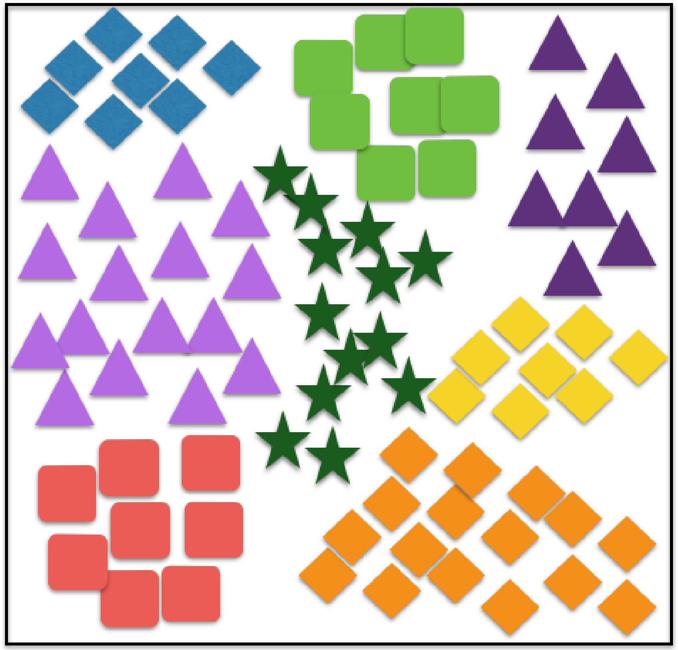
Pathway/Gene set enrichment analysis methods

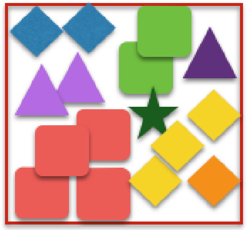


Biological pathway = Gene sets = Up-regulated genes at the specific condition

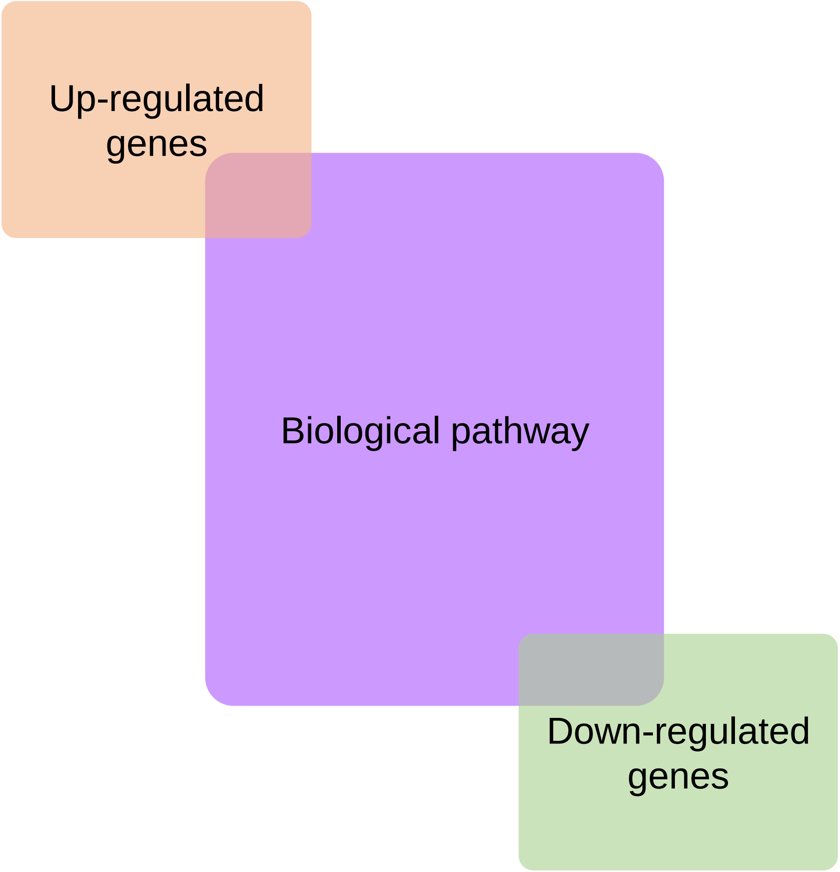
Overrepresentation-based enrichment

All known genes in a species (categorized into groups)





DEGs

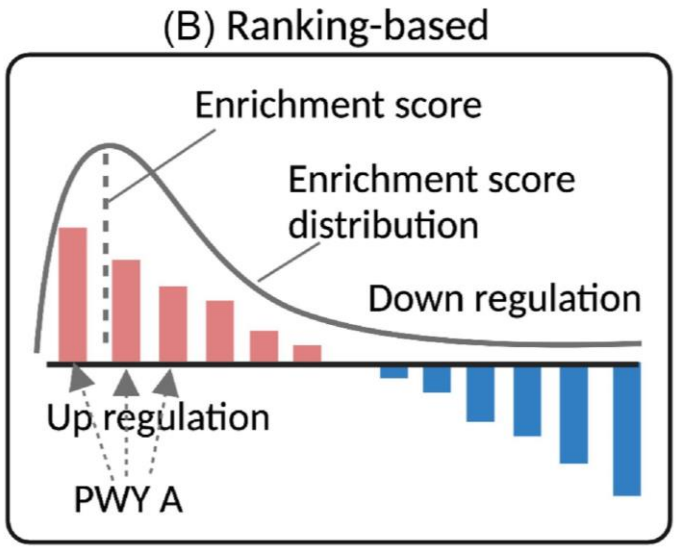


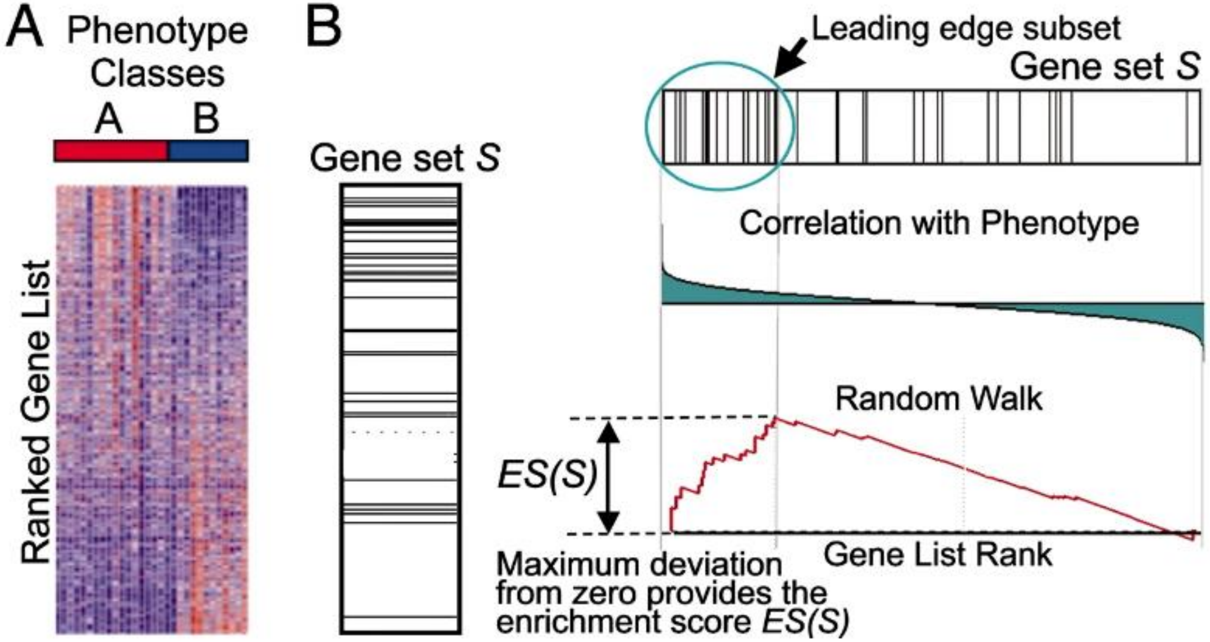
Up-regulated genes

Biological pathway

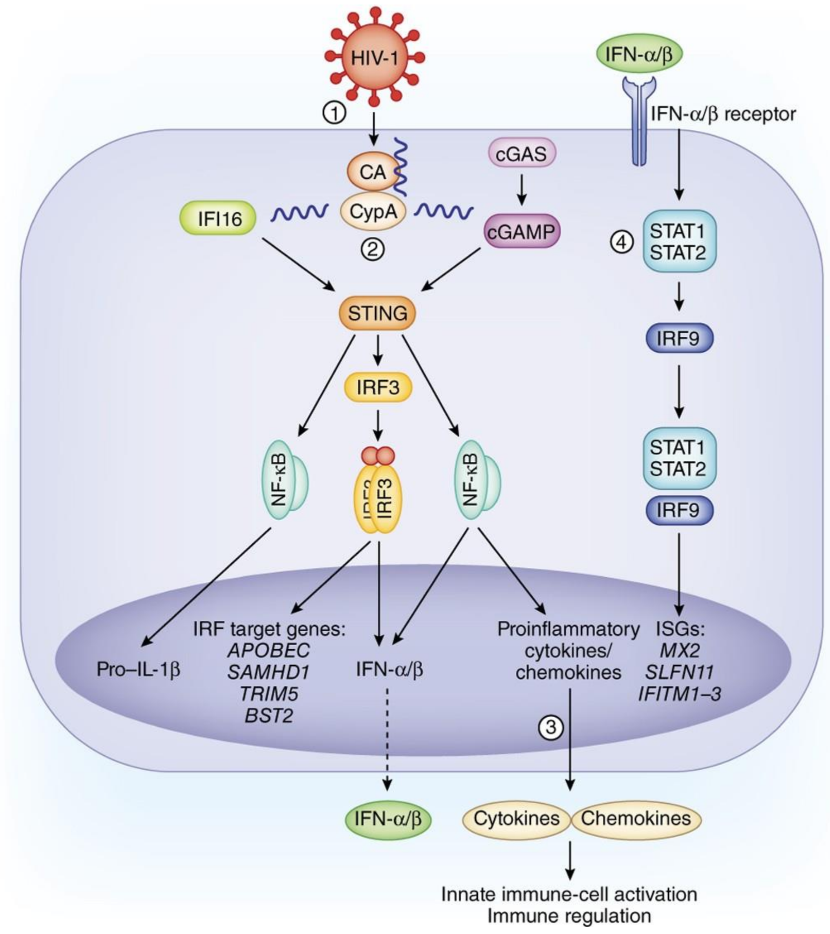
Down-regulated genes

Ranking-based enrichment

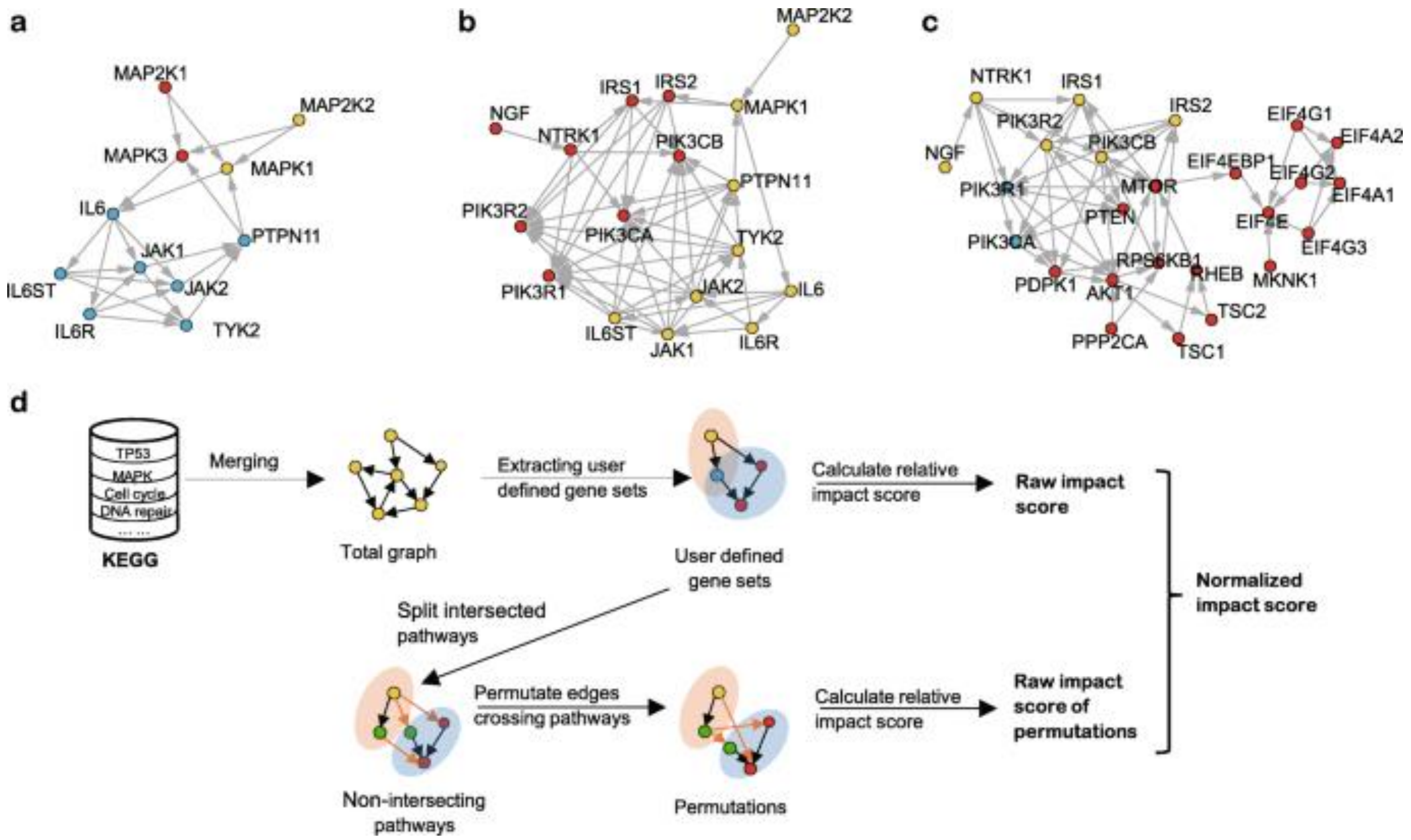




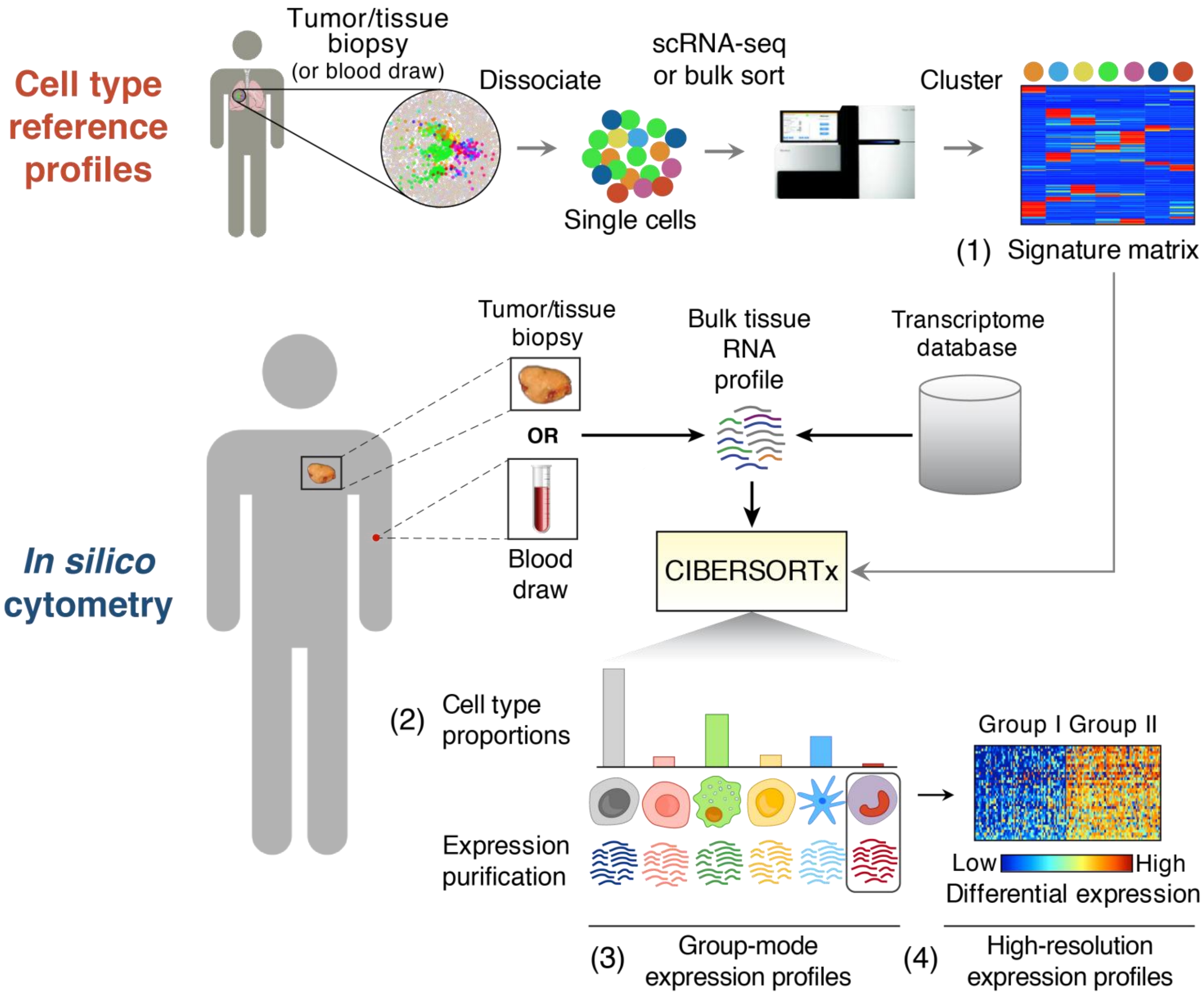
Topology-based enrichment analysis



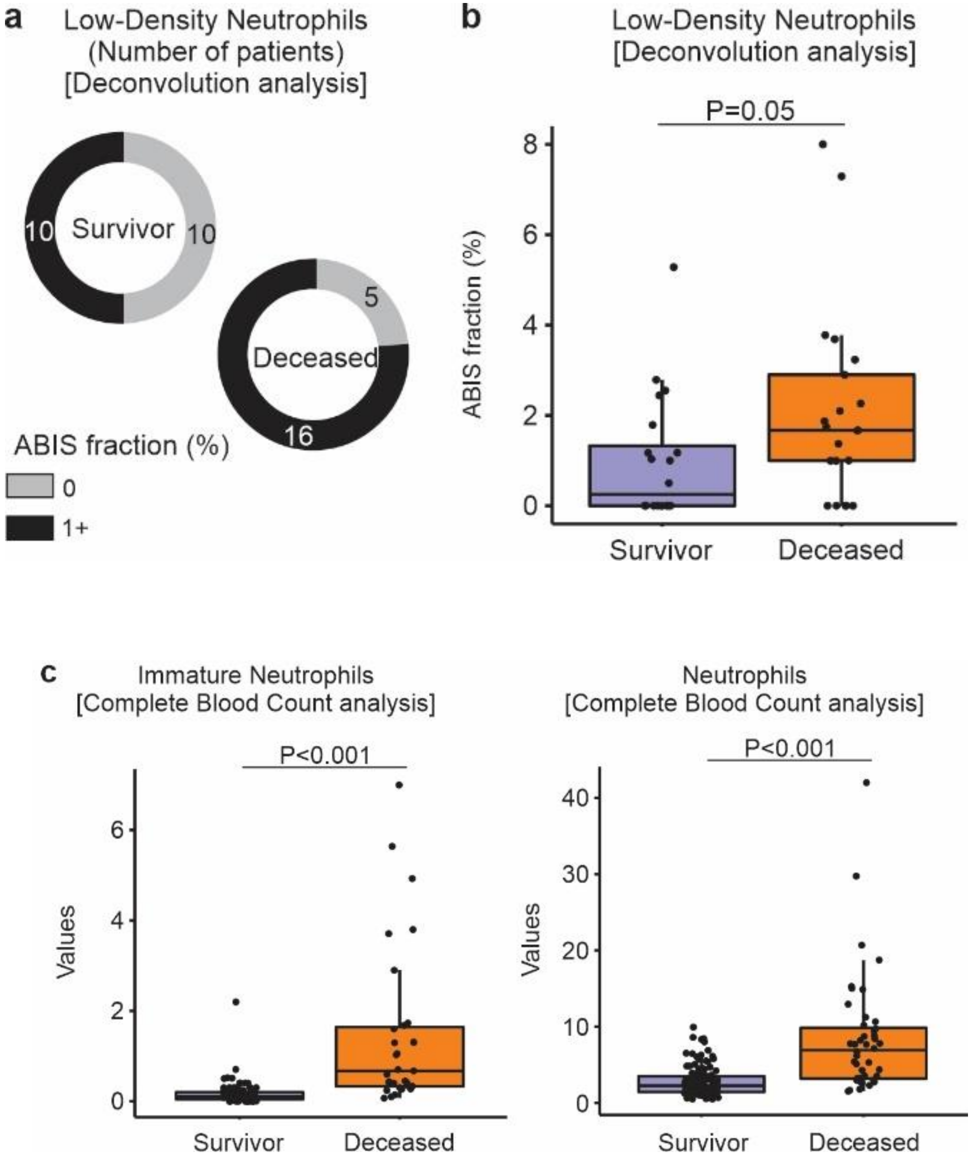
Kim Caesar/Nature Publishing Group

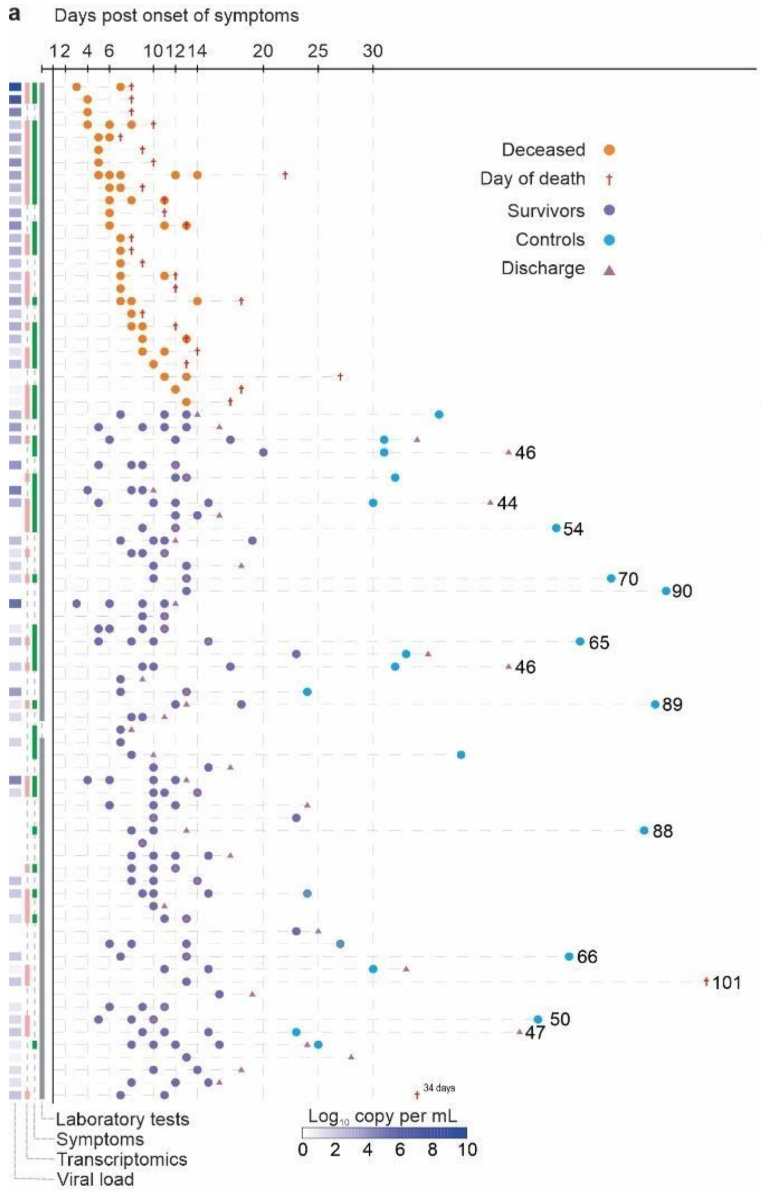


Deconvolution



Deconvolution

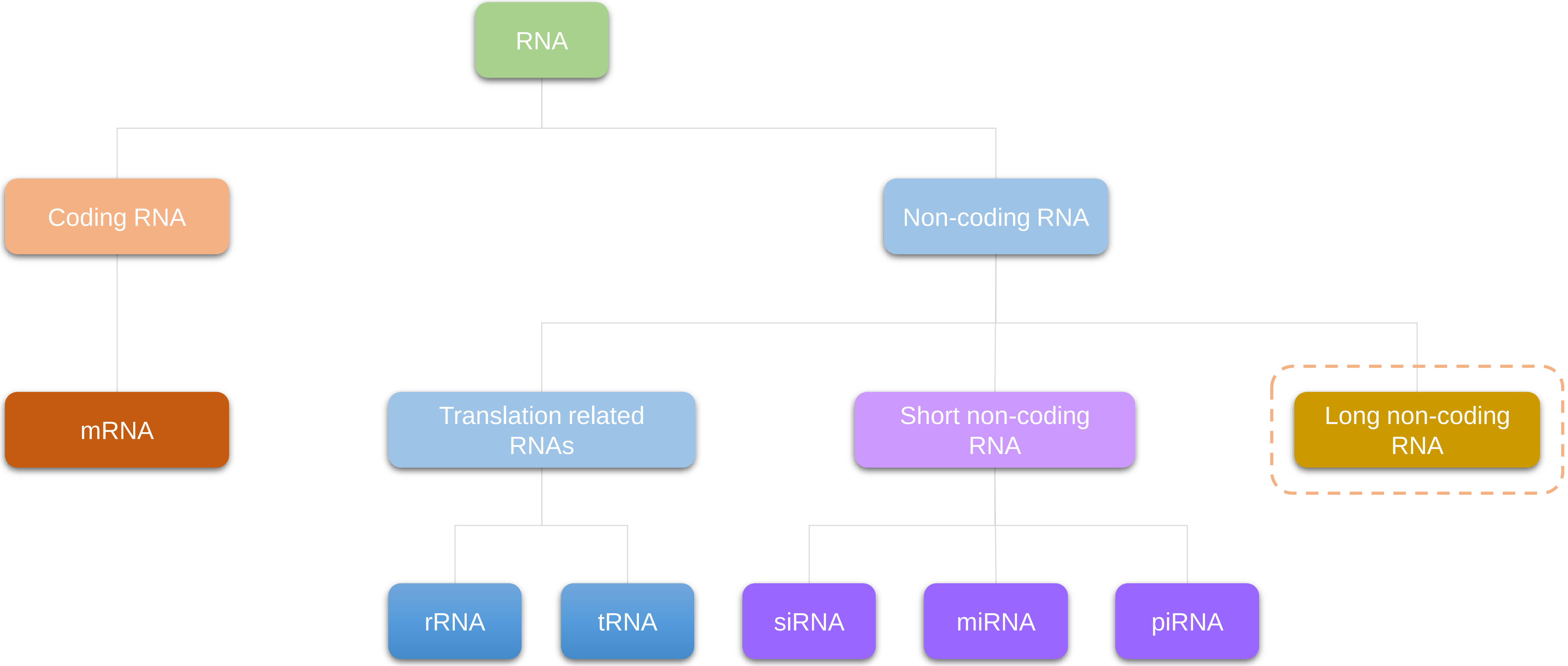




What we have seen until now…

- mRNA sequencing – transcriptome – library preparation
- Transcriptome pipeline
- Principal component analysis showing variability ingroup and between groups
- Assessment of potential outliers using molecular degree of perturbation
- Differentially expressed genes analysis
- Pathway enrichment analysis
- Deconvolution analysis

Coding and non-coding RNAs



RNA

Coding RNA

Non-coding RNA

mRNA

Translation related RNAs

Short non-coding RNA

Long non-coding RNA

rRNA

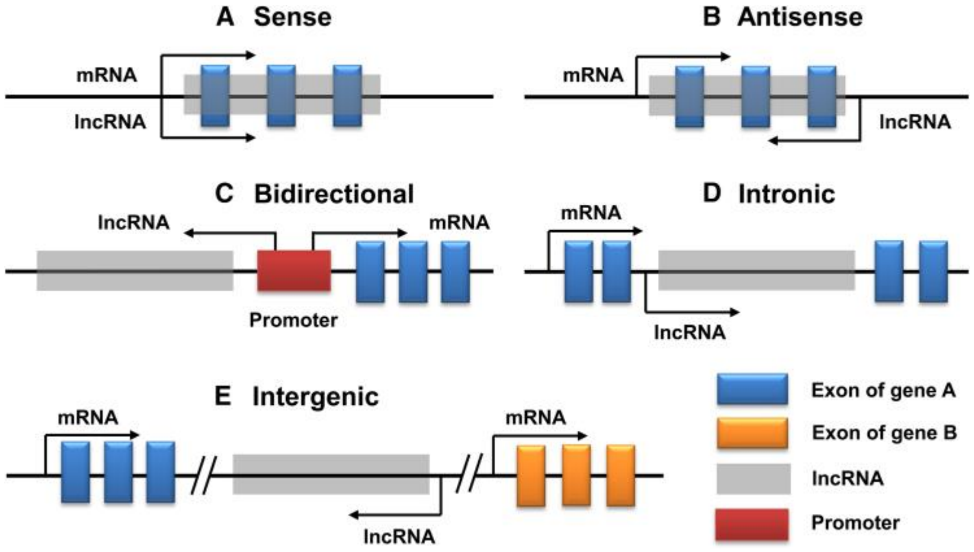
tRNA

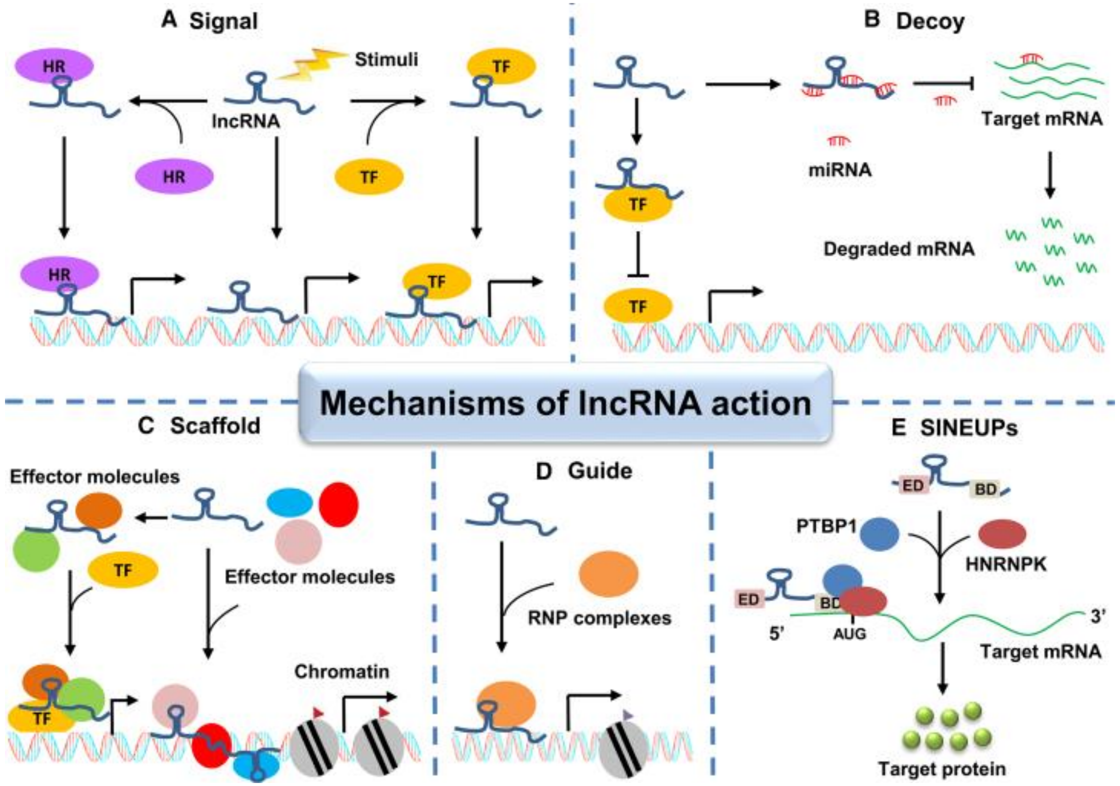
siRNA

miRNA

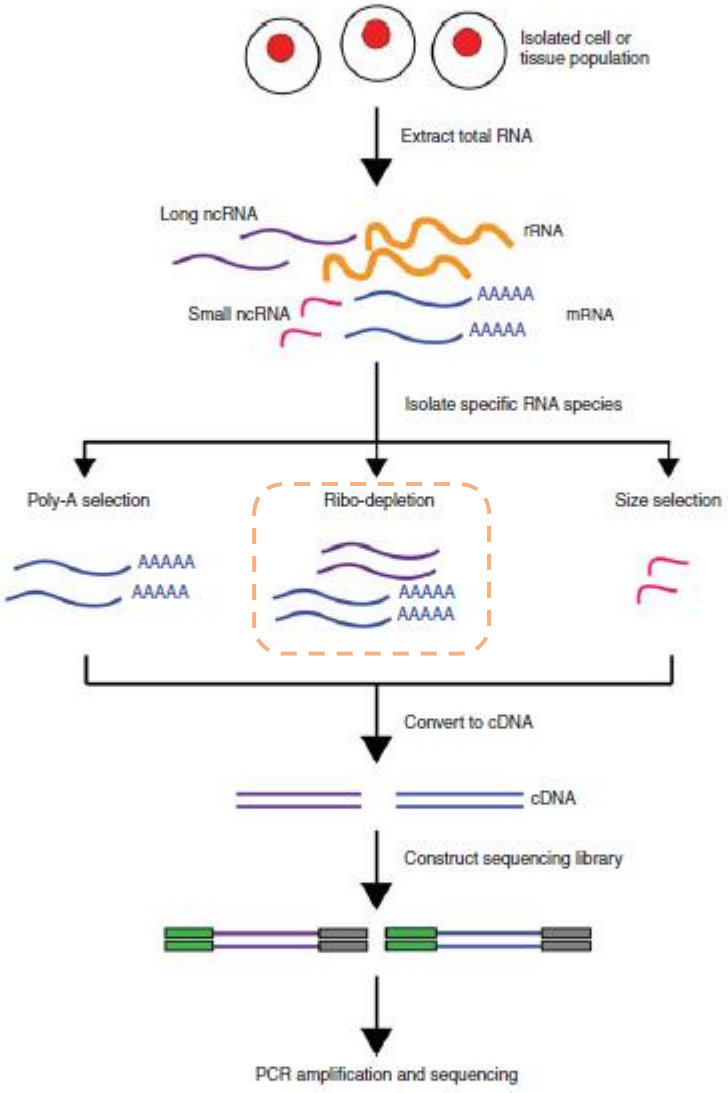
piRNA

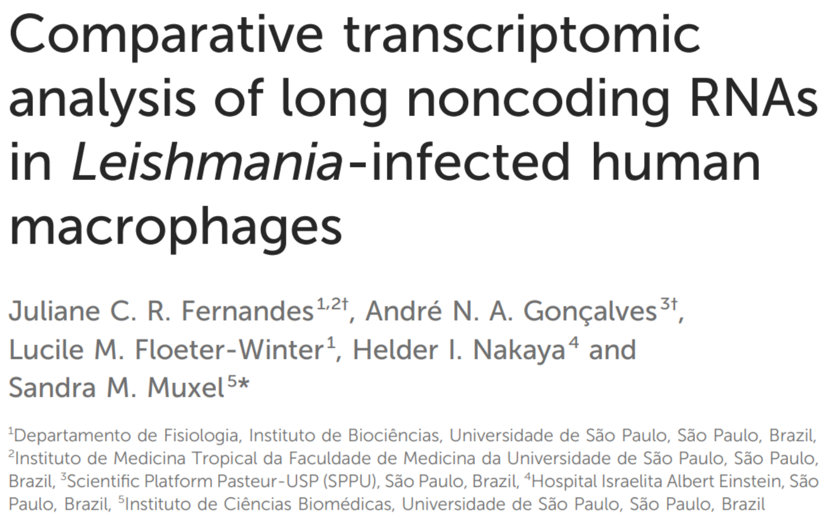
Long non-coding RNA biogenesis and mechanism of action





Long non-coding RNA sequencing strategy

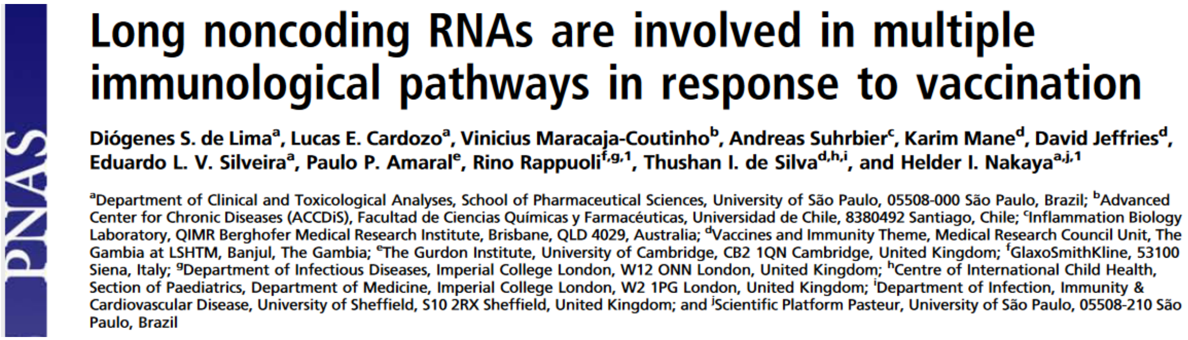




Comparative transcriptomic analysis of long noncoding RNAs in *Leishmania*-infected human macrophages

Juliane C. R. Fernandes[1,2†], André N. A. Gonçalves[3†],
Lucile M. Floeter-Winter[1], Helder I. Nakaya[4] and
Sandra M. Muxel[5*]

[1]Departamento de Fisiologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil, [2]Instituto de Medicina Tropical da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, [3]Scientific Platform Pasteur-USP (SPPU), São Paulo, Brazil, [4]Hospital Israelita Albert Einstein, São Paulo, Brazil, [5]Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, Brazil



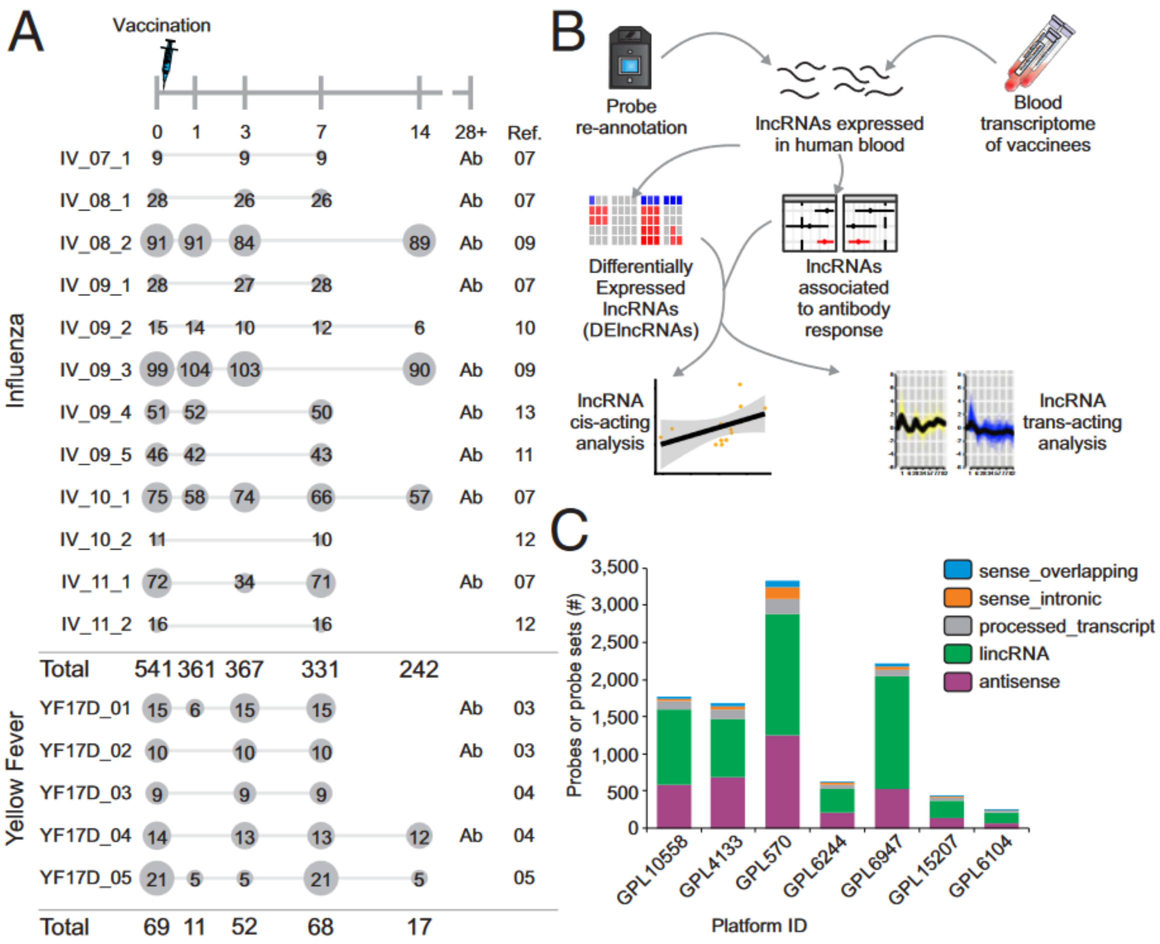
PNAS

Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination

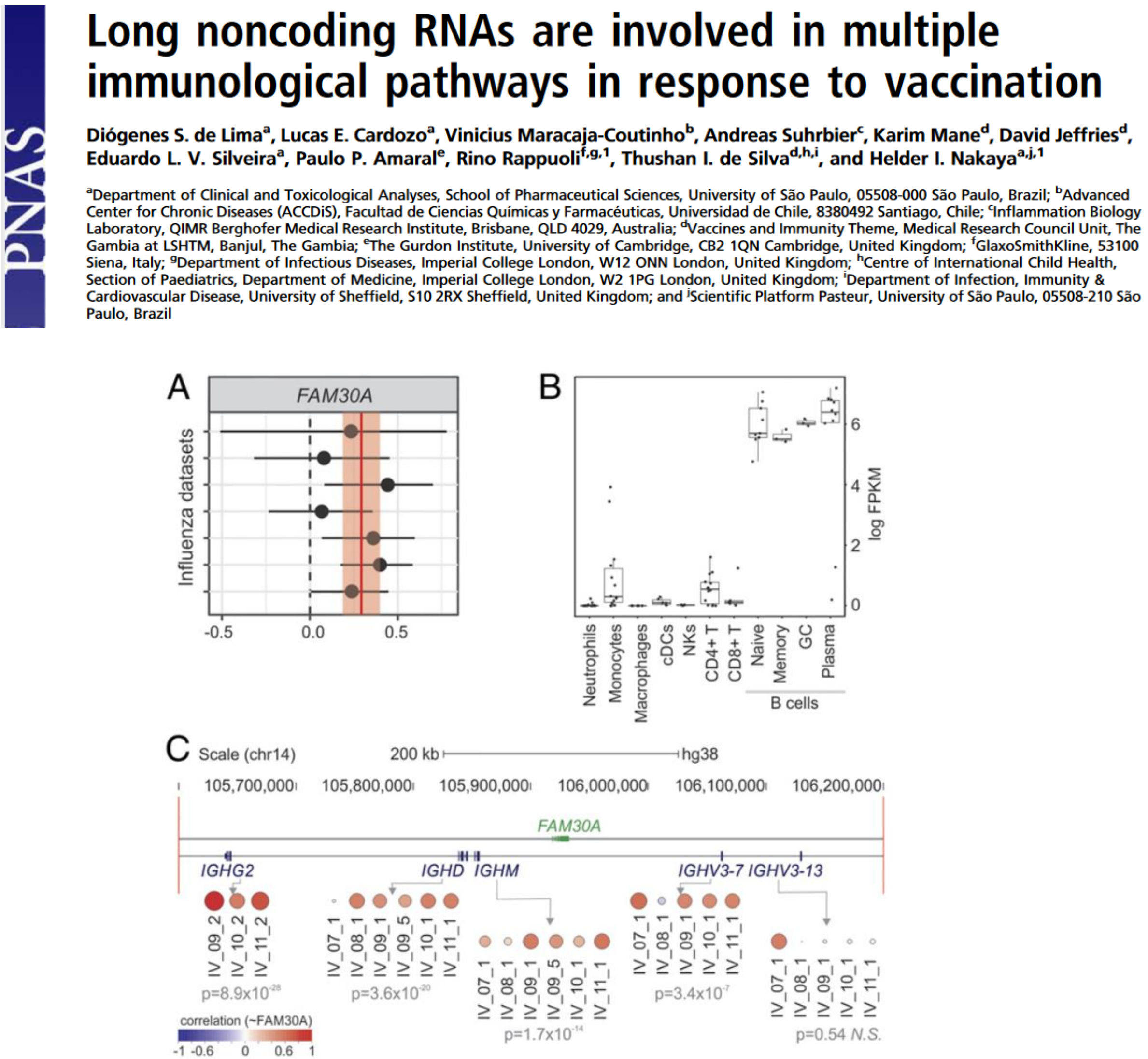
Diógenes S. de Lima[a], Lucas E. Cardozo[a], Vinicius Maracaja-Coutinho[b], Andreas Suhrbier[c], Karim Mane[d], David Jeffries[d], Eduardo L. V. Silveira[a], Paulo P. Amaral[e], Rino Rappuoli[f,g,1], Thushan I. de Silva[d,h,i], and Helder I. Nakaya[a,j,1]

[a]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, 05508-000 São Paulo, Brazil; [b]Advanced Center for Chronic Diseases (ACCDiS), Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, 8380492 Santiago, Chile; [c]Inflammation Biology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia; [d]Vaccines and Immunity Theme, Medical Research Council Unit, The Gambia at LSHTM, Banjul, The Gambia; [e]The Gurdon Institute, University of Cambridge, CB2 1QN Cambridge, United Kingdom; [f]GlaxoSmithKline, 53100 Siena, Italy; [g]Department of Infectious Diseases, Imperial College London, W12 0NN London, United Kingdom; [h]Centre of International Child Health, Section of Paediatrics, Department of Medicine, Imperial College London, W2 1PG London, United Kingdom; [i]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, S10 2RX Sheffield, United Kingdom; and [j]Scientific Platform Pasteur, University of São Paulo, 05508-210 São Paulo, Brazil

Evolution of Gene Expression Analysis: A Century of Discovery



	0	1	3	7	14	28+	Ref.
IV_07_1	9		9	9		Ab	07
IV_08_1	28		26	26		Ab	07
IV_08_2	91	91	84		89	Ab	09
IV_09_1	28		27	28		Ab	07
IV_09_2	15	14	10	12	6		10
IV_09_3	99	104	103		90	Ab	09
IV_09_4	51	52		50		Ab	13
IV_09_5	46	42		43		Ab	11
IV_10_1	75	58	74	66	57	Ab	07
IV_10_2	11			10			12
IV_11_1	72		34	71		Ab	07
IV_11_2	16			16			12
Total	541	361	367	331	242		
YF17D_01	15	6	15	15		Ab	03
YF17D_02	10		10	10		Ab	03
YF17D_03	9		9	9			04
YF17D_04	14		13	13	12	Ab	04
YF17D_05	21	5	5	21	5		05
Total	69	11	52	68	17		



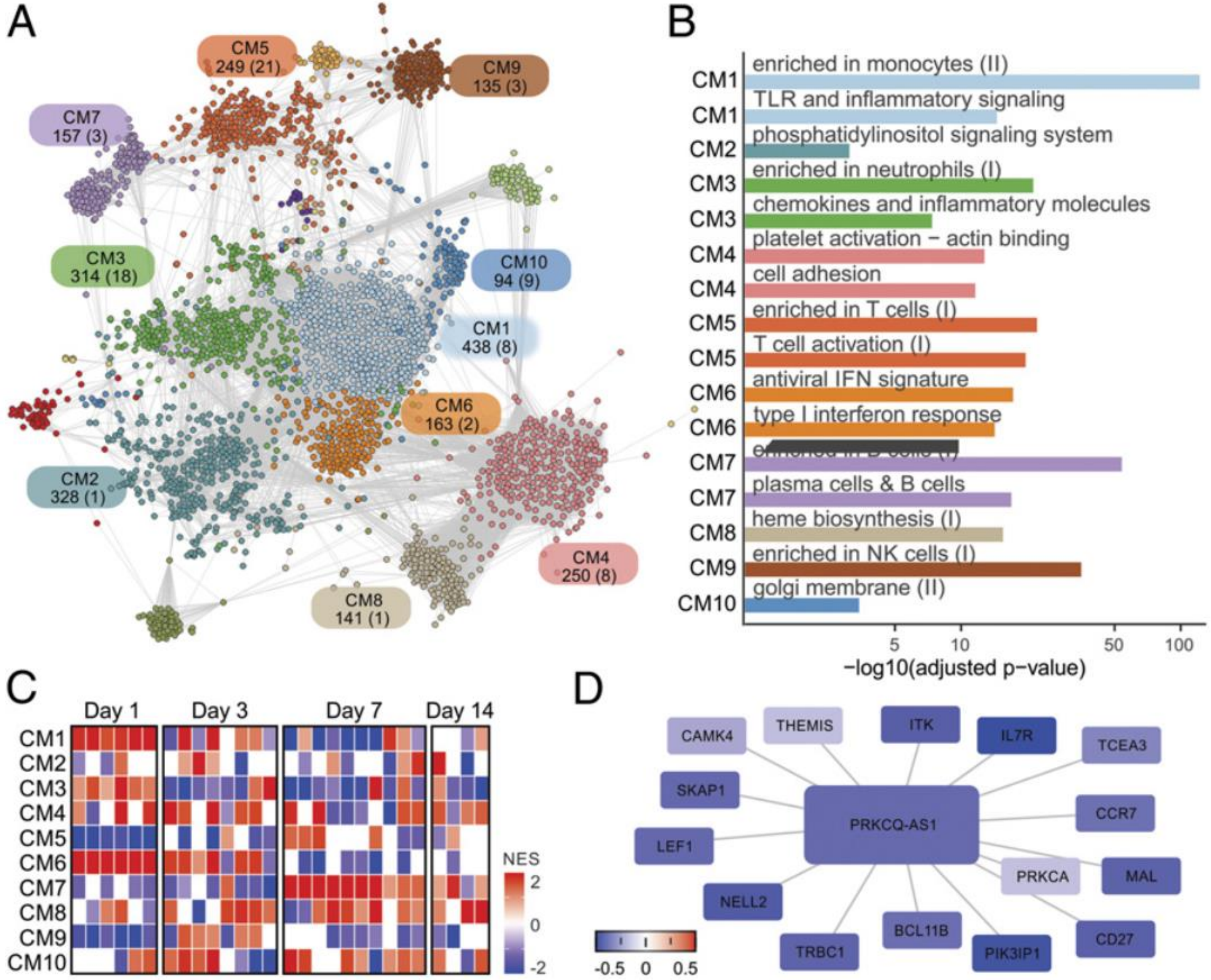
PNAS

Long noncoding RNAs are involved in multiple immunological pathways in response to vaccination

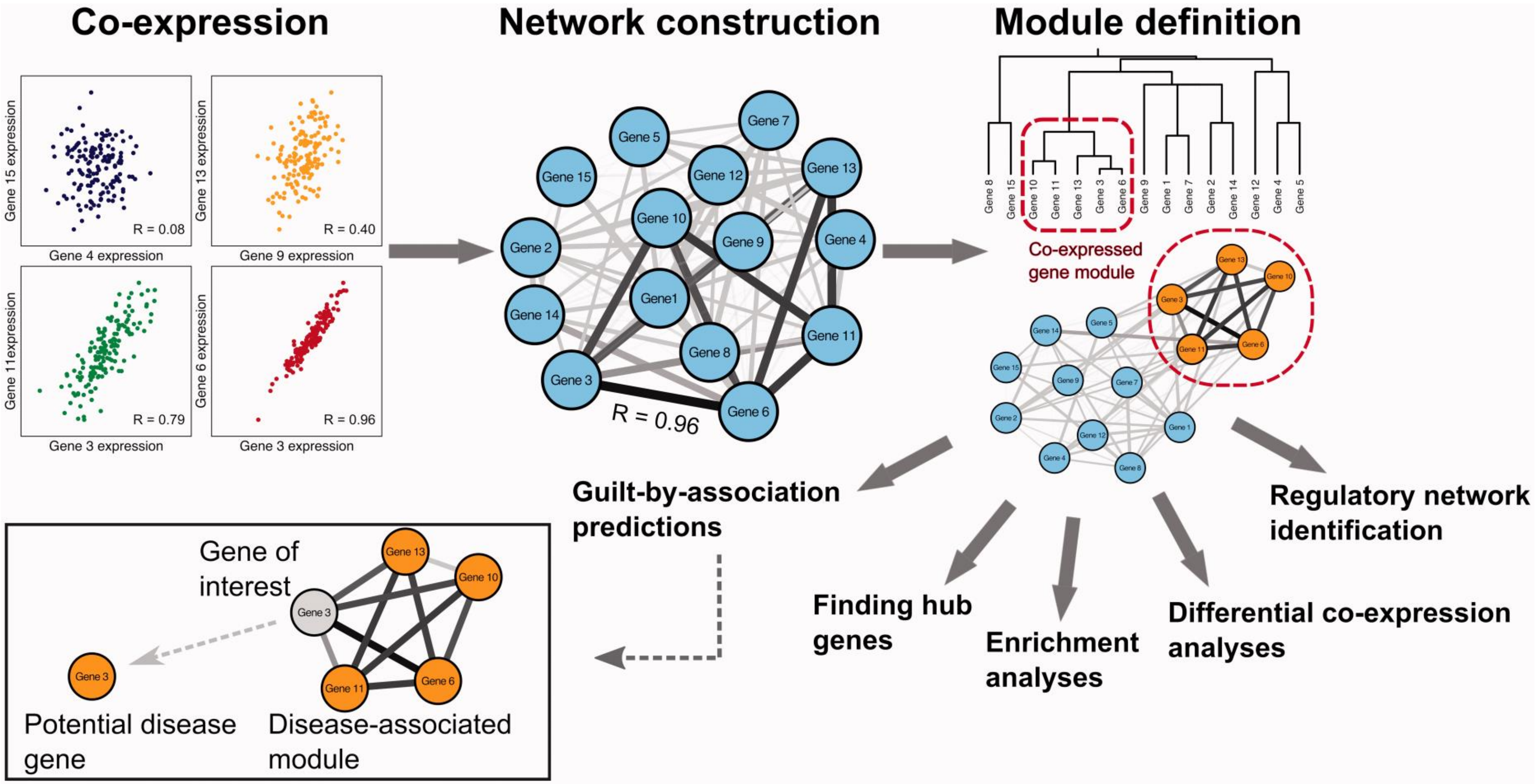
Diógenes S. de Lima[a], Lucas E. Cardozo[a], Vinicius Maracaja-Coutinho[b], Andreas Suhrbier[c], Karim Mane[d], David Jeffries[d], Eduardo L. V. Silveira[a], Paulo P. Amaral[e], Rino Rappuoli[f,g,1], Thushan I. de Silva[d,h,i], and Helder I. Nakaya[a,j,1]

[a]Department of Clinical and Toxicological Analyses, School of Pharmaceutical Sciences, University of São Paulo, 05508-000 São Paulo, Brazil; [b]Advanced Center for Chronic Diseases (ACCDiS), Facultad de Ciencias Químicas y Farmacéuticas, Universidad de Chile, 8380492 Santiago, Chile; [c]Inflammation Biology Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4029, Australia; [d]Vaccines and Immunity Theme, Medical Research Council Unit, The Gambia at LSHTM, Banjul, The Gambia; [e]The Gurdon Institute, University of Cambridge, CB2 1QN Cambridge, United Kingdom; [f]GlaxoSmithKline, 53100 Siena, Italy; [g]Department of Infectious Diseases, Imperial College London, W12 0NN London, United Kingdom; [h]Centre of International Child Health, Section of Paediatrics, Department of Medicine, Imperial College London, W2 1PG London, United Kingdom; [i]Department of Infection, Immunity & Cardiovascular Disease, University of Sheffield, S10 2RX Sheffield, United Kingdom; and [j]Scientific Platform Pasteur, University of São Paulo, 05508-210 São Paulo, Brazil

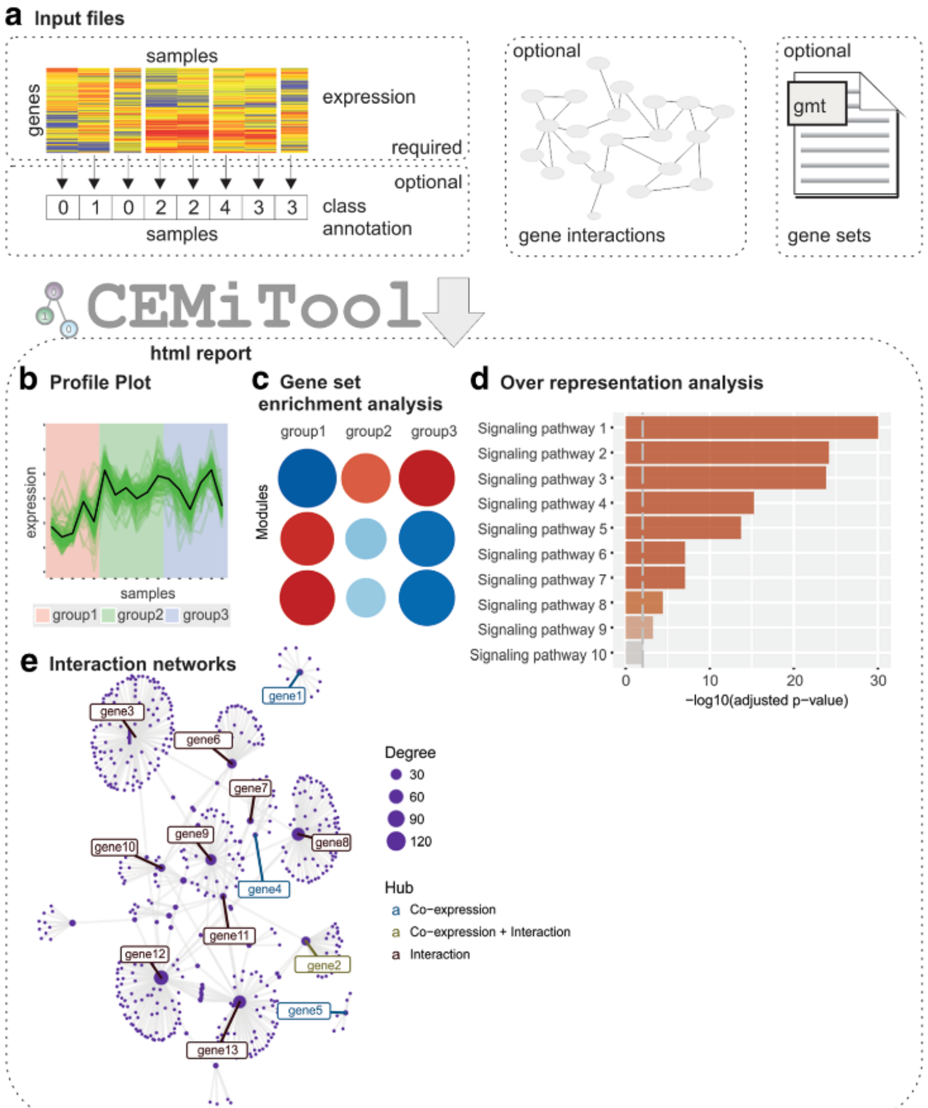
Long non-coding and mRNA expression relationship

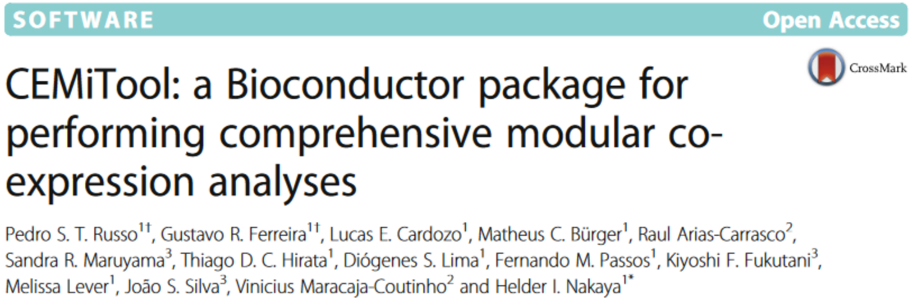


Co-expression module analysis



Co-Expression Module identification Tool (CEMiTool)

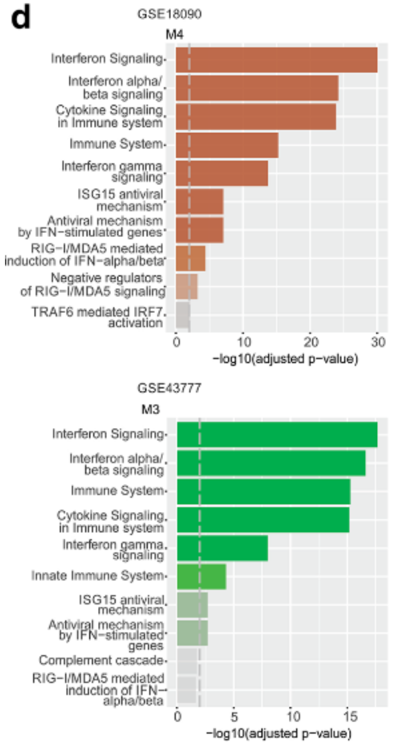


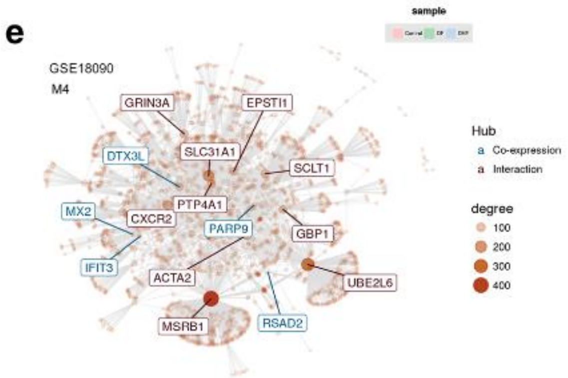

SOFTWARE Open Access

CrossMark

CEMiTool: a Bioconductor package for performing comprehensive modular co-expression analyses

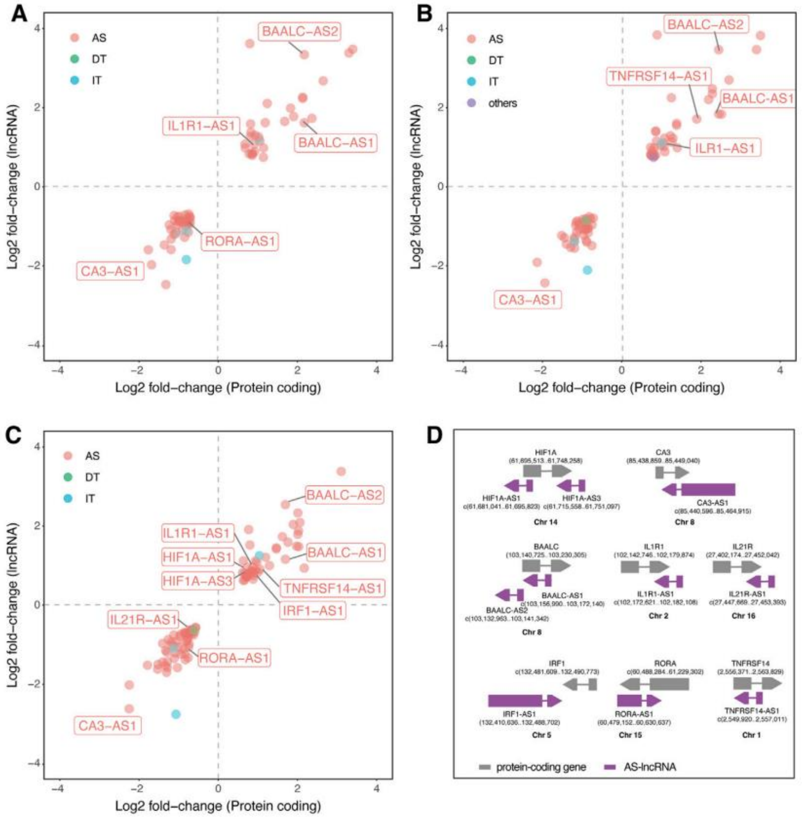
Pedro S. T. Russo[1†], Gustavo R. Ferreira[1†], Lucas E. Cardozo[1], Matheus C. Bürger[1], Raul Arias-Carrasco[2], Sandra R. Maruyama[3], Thiago D. C. Hirata[1], Diógenes S. Lima[1], Fernando M. Passos[1], Kiyoshi F. Fukutani[3], Melissa Lever[1], João S. Silva[3], Vinicius Maracaja-Coutinho[2] and Helder I. Nakaya[1*]





Antisense lncRNAs are co-expressed with sense protein coding may be involved in immune response

Category	# from total transcripts	% from total	# from DE	% total DEGs
Protein coding	13892	72.95	2,926	67.87
LncRNA	4,627	24.30	1,267	29.39
miRNA	139	0.73	29	0.67
snoRNA	121	0.64	22	0.51
snRNA	33	0.17	20	0.46
rRNA	4	0.02	4	0.09
Others	227	1.19	43	1
Total	19043		4,311	

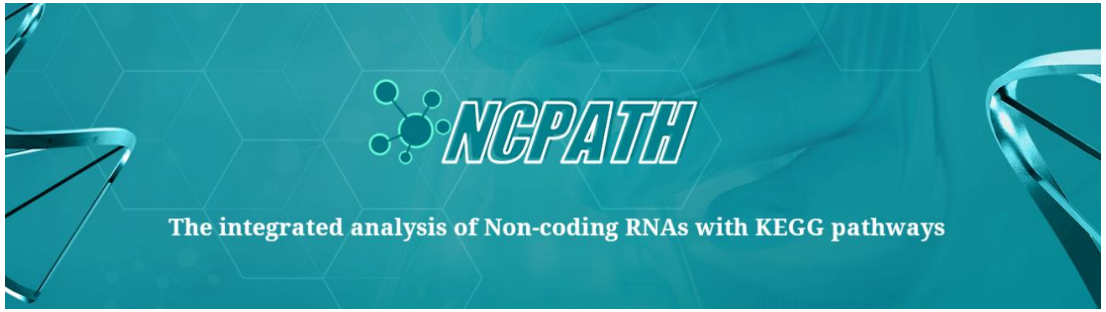


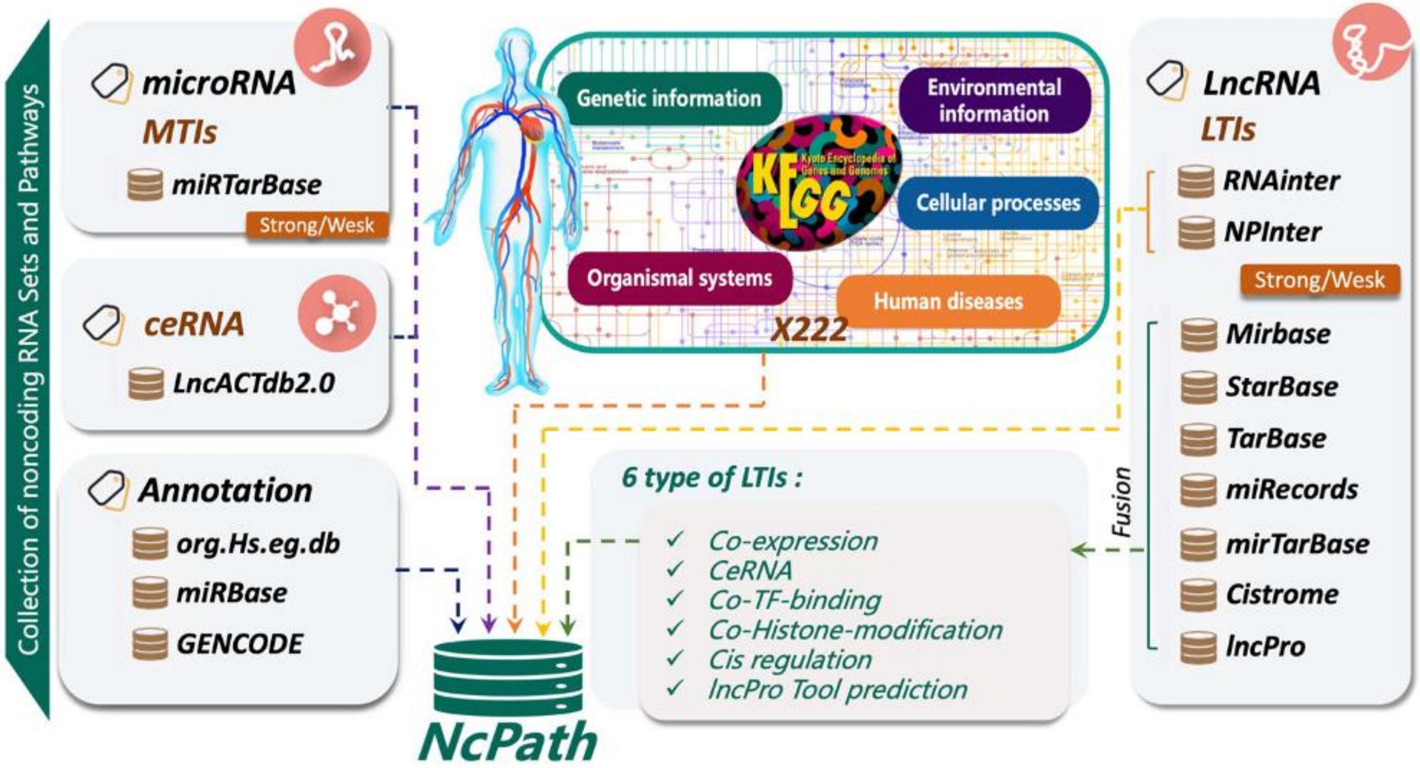
Comparative transcriptomic analysis of long noncoding RNAs in *Leishmania*-infected human macrophages

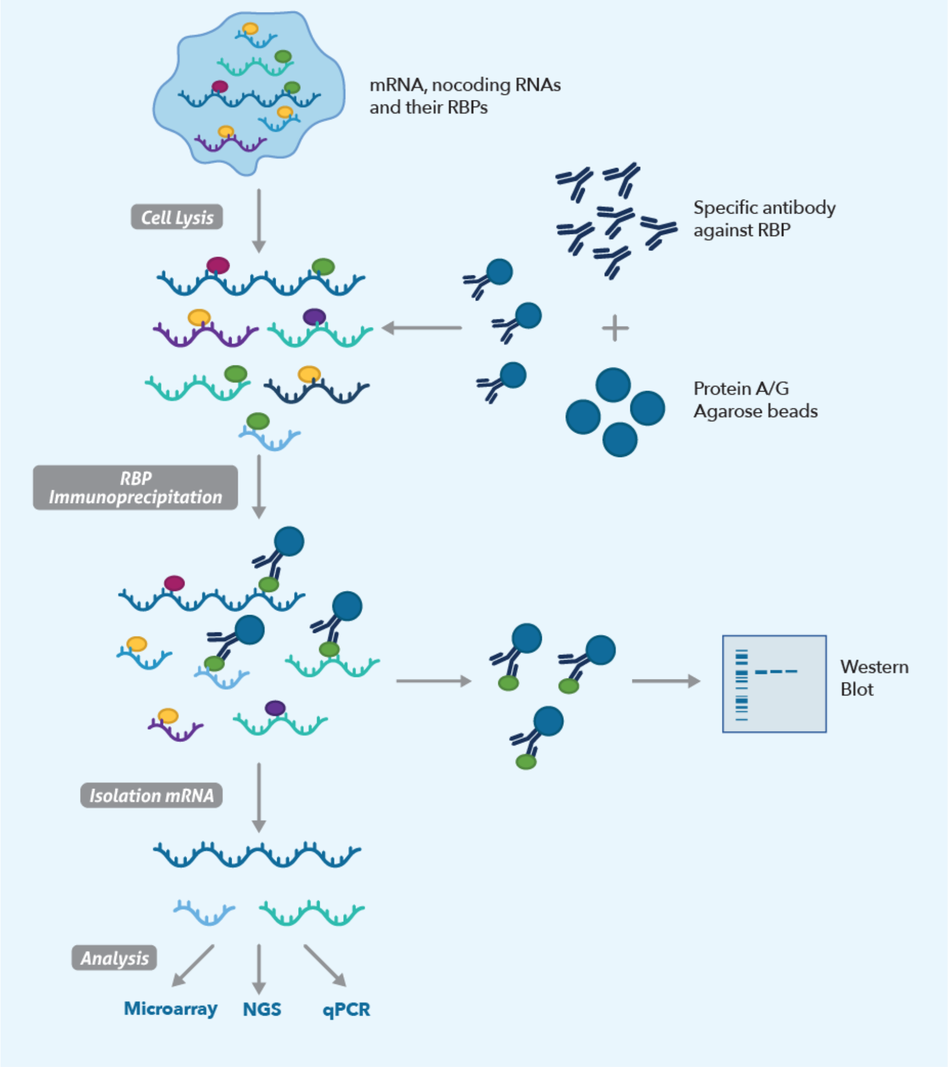
Juliane C. R. Fernandes[1,2†], André N. A. Gonçalves[3†], Lucile M. Floeter-Winter[1], Helder I. Nakaya[4] and Sandra M. Muxel[5*]

[1]Departamento de Fisiologia, Instituto de Biociências, Universidade de São Paulo, São Paulo, Brazil, [2]Instituto de Medicina Tropical da Faculdade de Medicina da Universidade de São Paulo, São Paulo, Brazil, [3]Scientific Platform Pasteur-USP (SPPU), São Paulo, Brazil, [4]Hospital Israelita Albert Einstein, São Paulo, Brazil, [5]Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo, Brazil

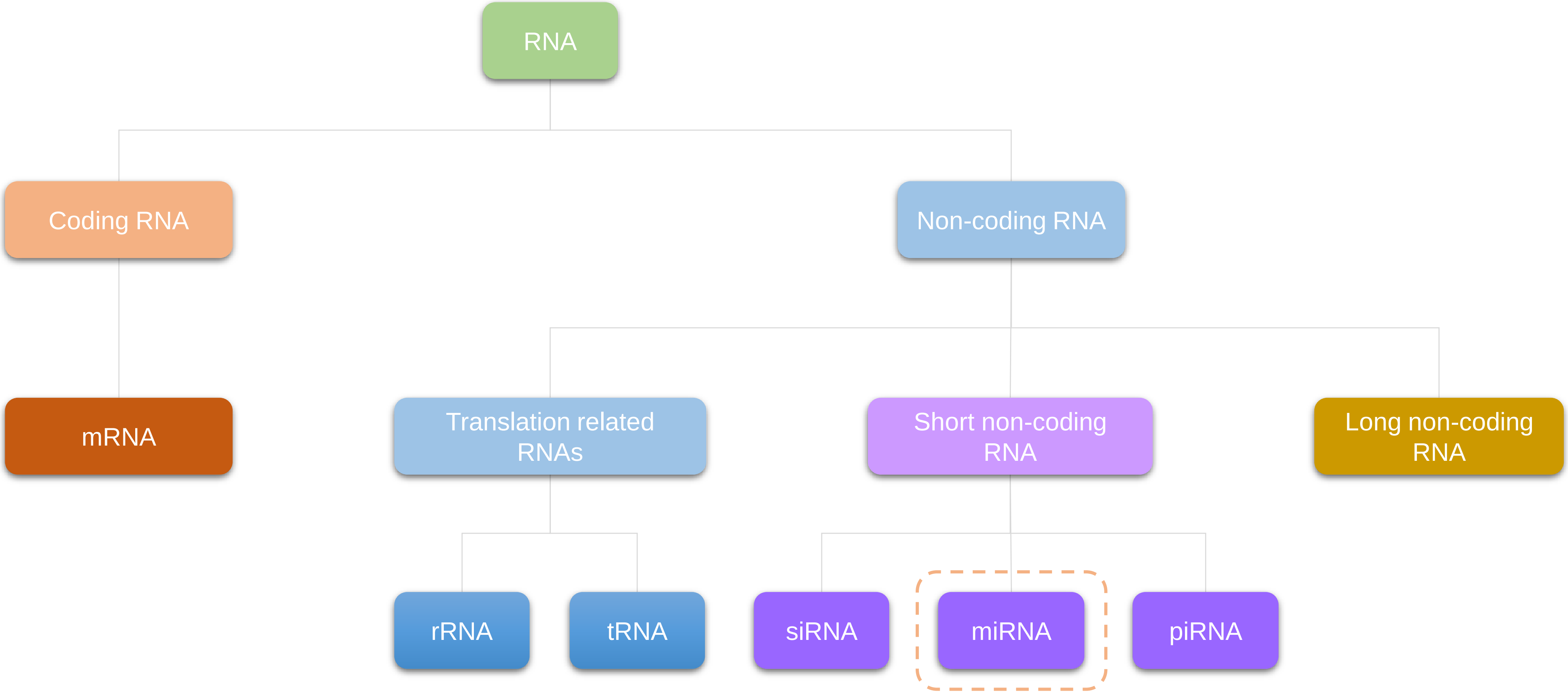
Long non-coding RNA function







Coding and non-coding RNAs



RNA

Coding RNA

Non-coding RNA

mRNA

Translation related RNAs

Short non-coding RNA

Long non-coding RNA

rRNA

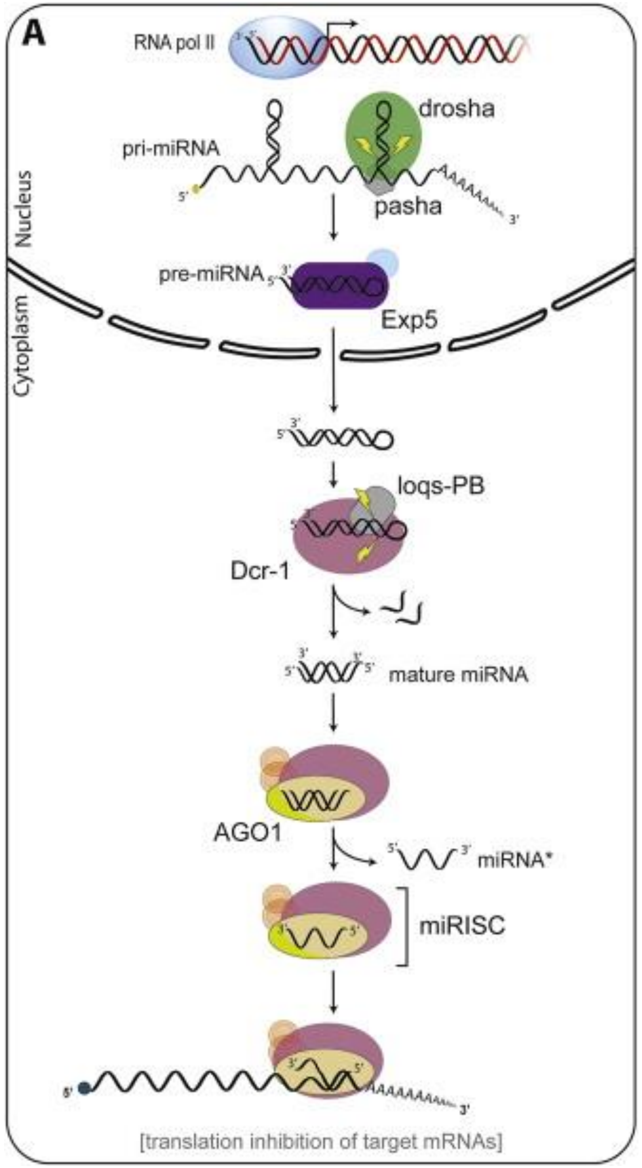
tRNA

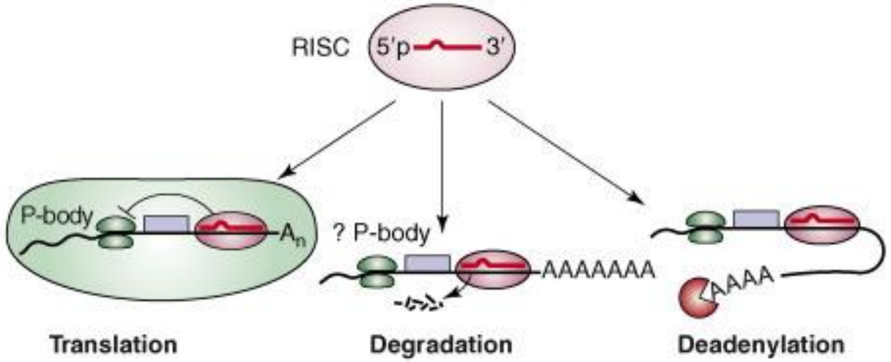
siRNA

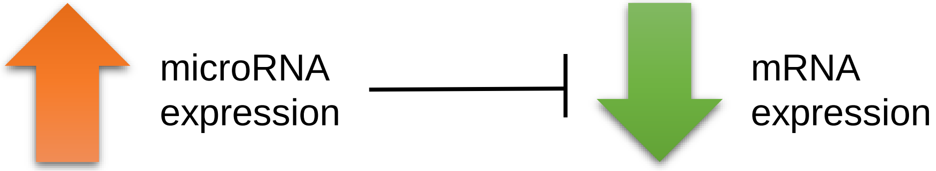
miRNA

piRNA

MicroRNA biogenesis and function



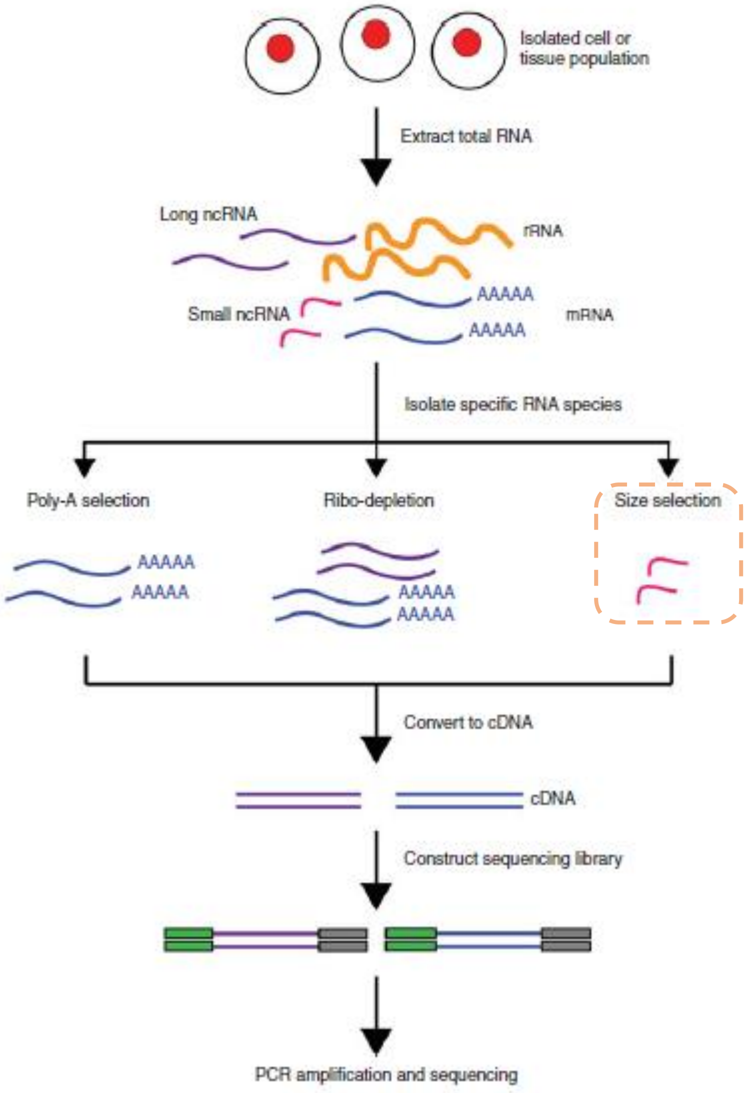


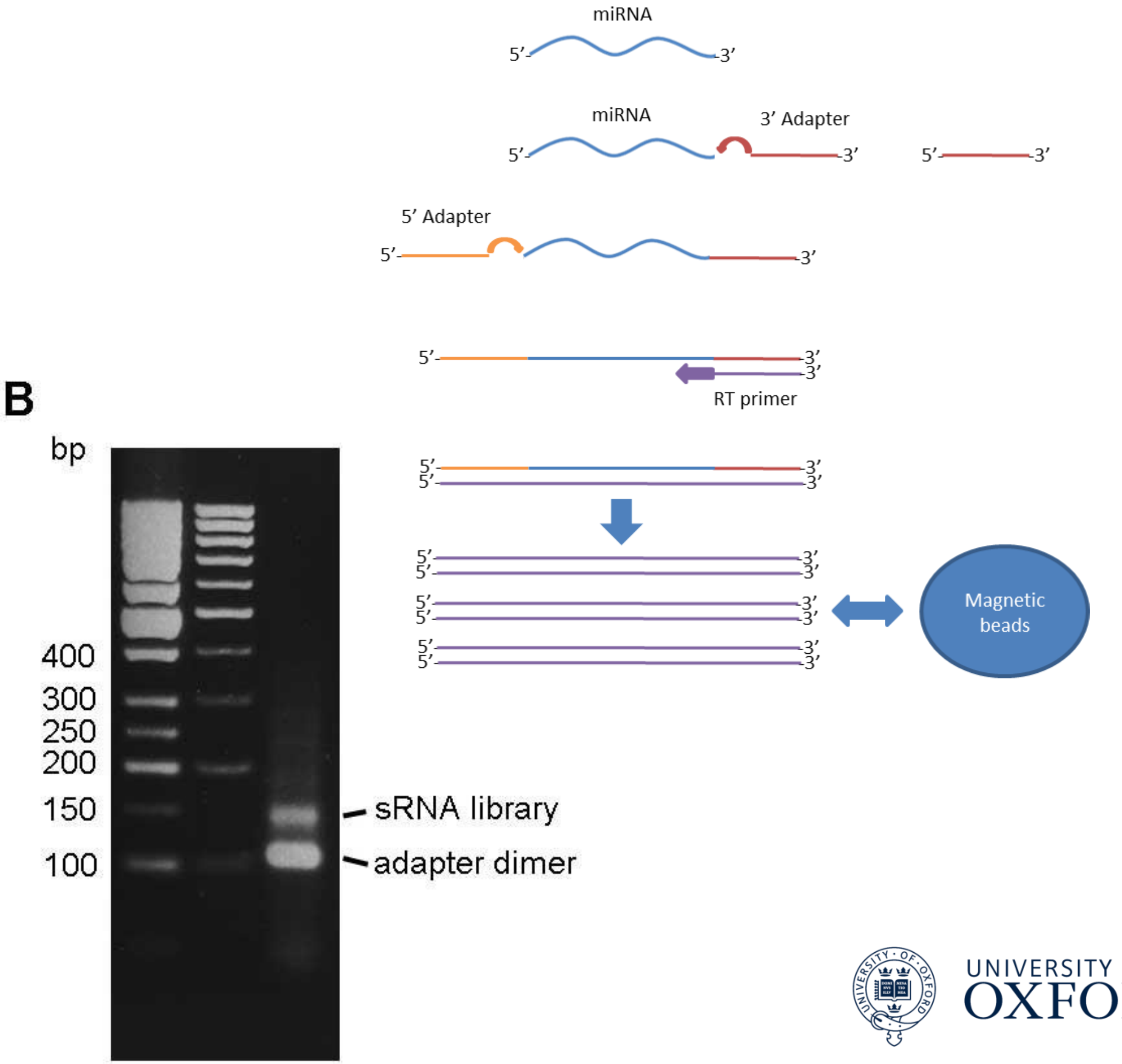


microRNA expression

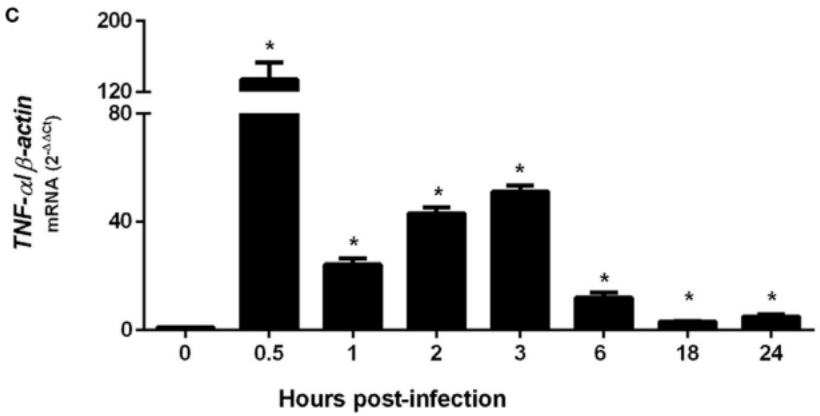
mRNA expression

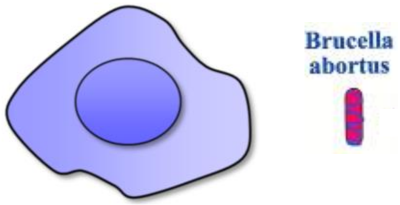
Small RNA sequencing protocol

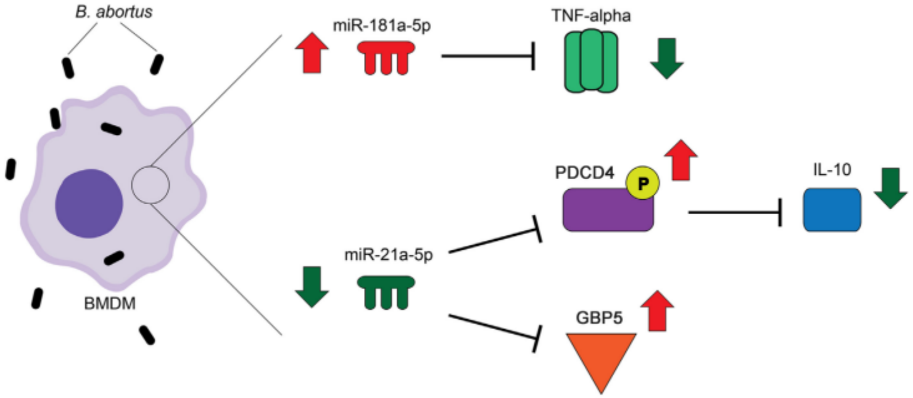


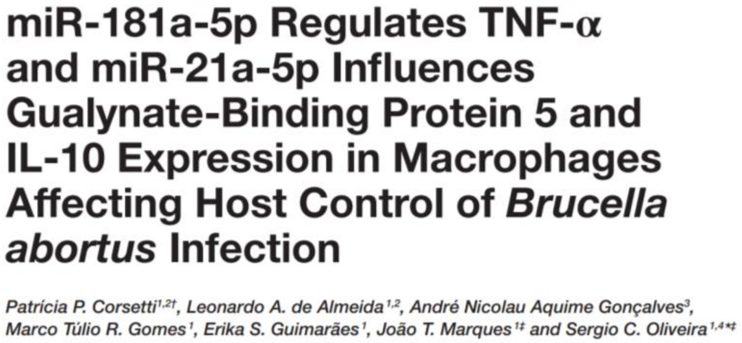


miR-21a-5p Regulates GBP5 Expression and Partially Influences Intracellular B. abortus Growth



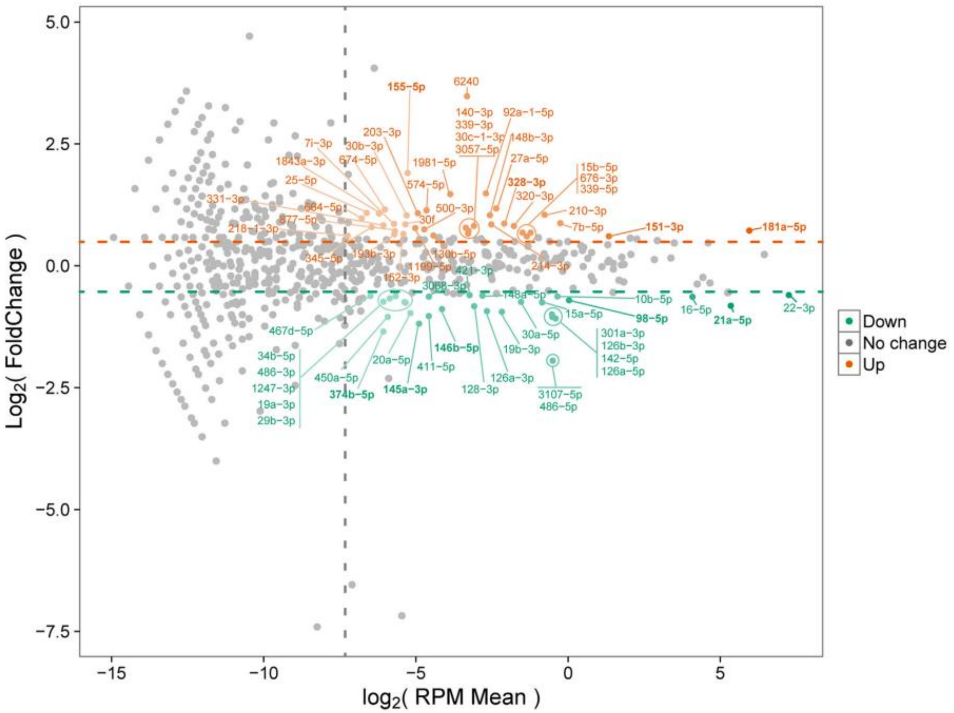




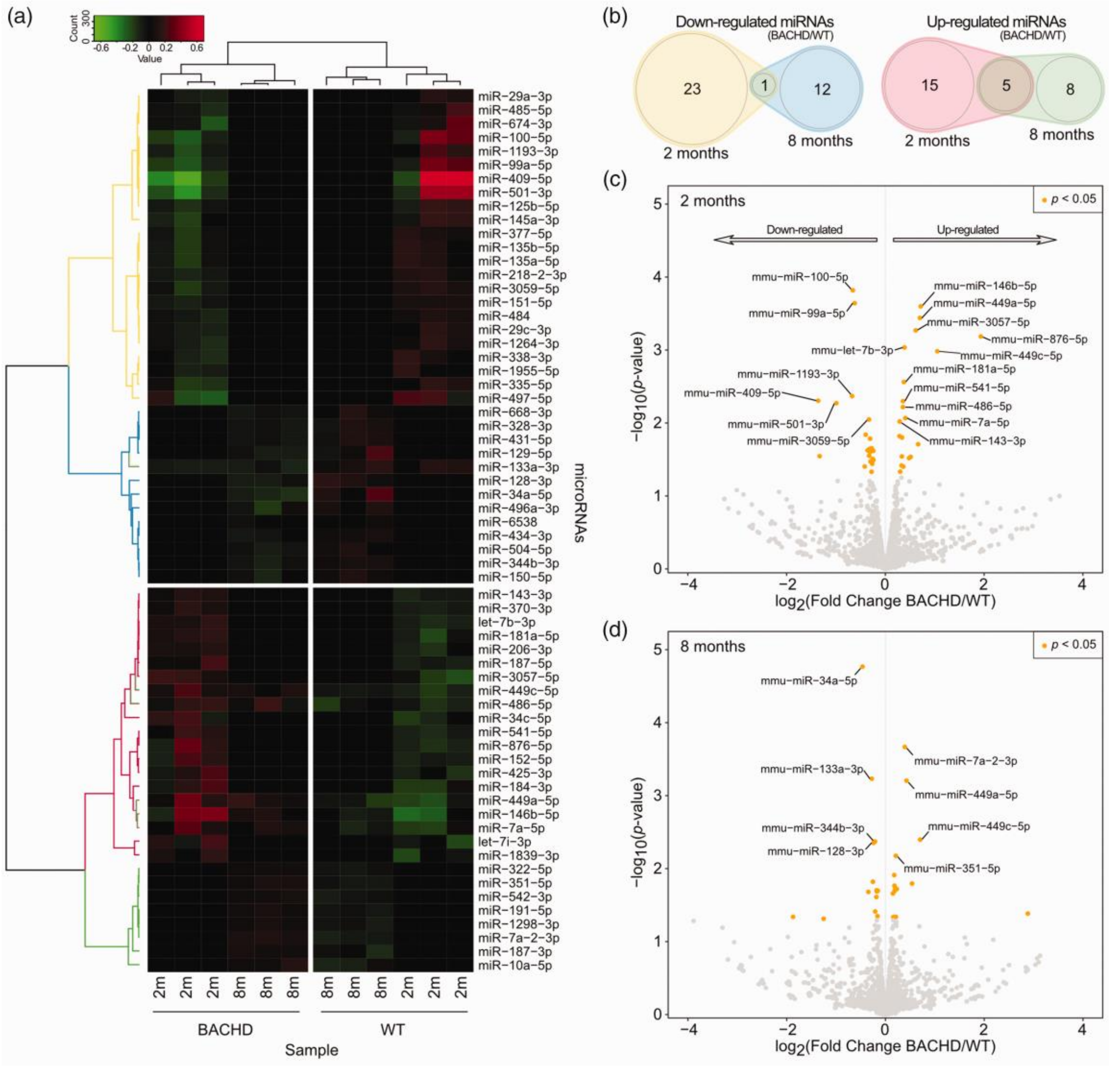


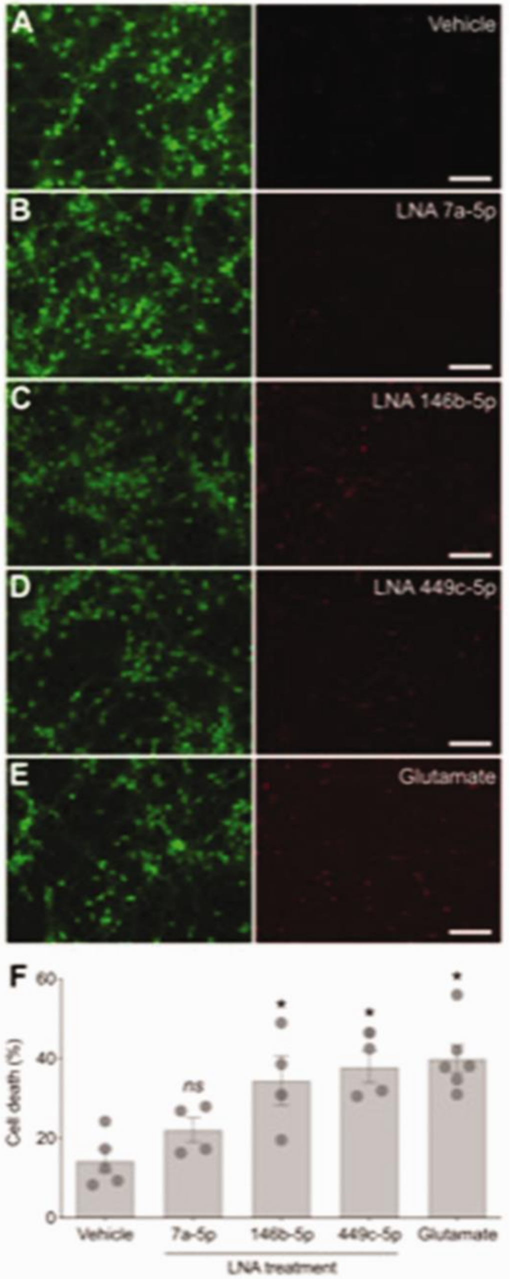
miR-181a-5p Regulates TNF-α and miR-21a-5p Influences Gualynate-Binding Protein 5 and IL-10 Expression in Macrophages Affecting Host Control of *Brucella abortus* Infection

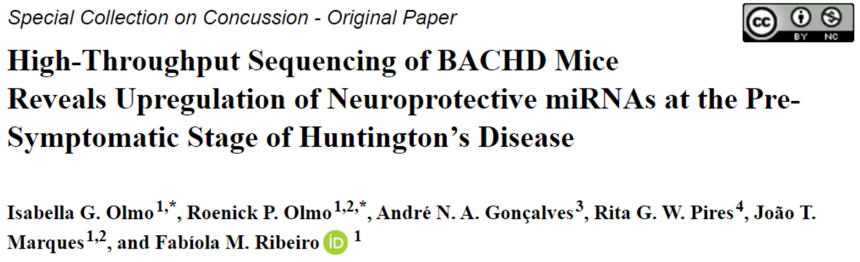
Patrícia P. Corsetti[1,2†], Leonardo A. de Almeida[1,2], André Nicolau Aquime Gonçalves[3], Marco Túlio R. Gomes[1], Erika S. Guimarães[1], João T. Marques[1‡] and Sergio C. Oliveira[1,4‡]*



miR-146b-5p and miR-449c-5p Blockade Promote Neurotoxicity in Striatal Neurons

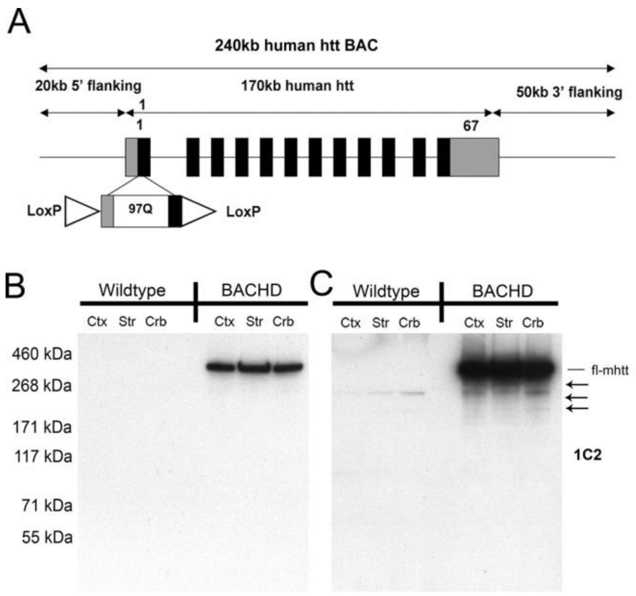





Special Collection on Concussion - Original Paper

High-Throughput Sequencing of BACHD Mice Reveals Upregulation of Neuroprotective miRNAs at the Pre-Symptomatic Stage of Huntington's Disease

Isabella G. Olmo[1,*], Roenick P. Olmo[1,2,*], André N. A. Gonçalves[3], Rita G. W. Pires[4], João T. Marques[1,2], and Fabíola M. Ribeiro[1]



Animals

2 months

8 months

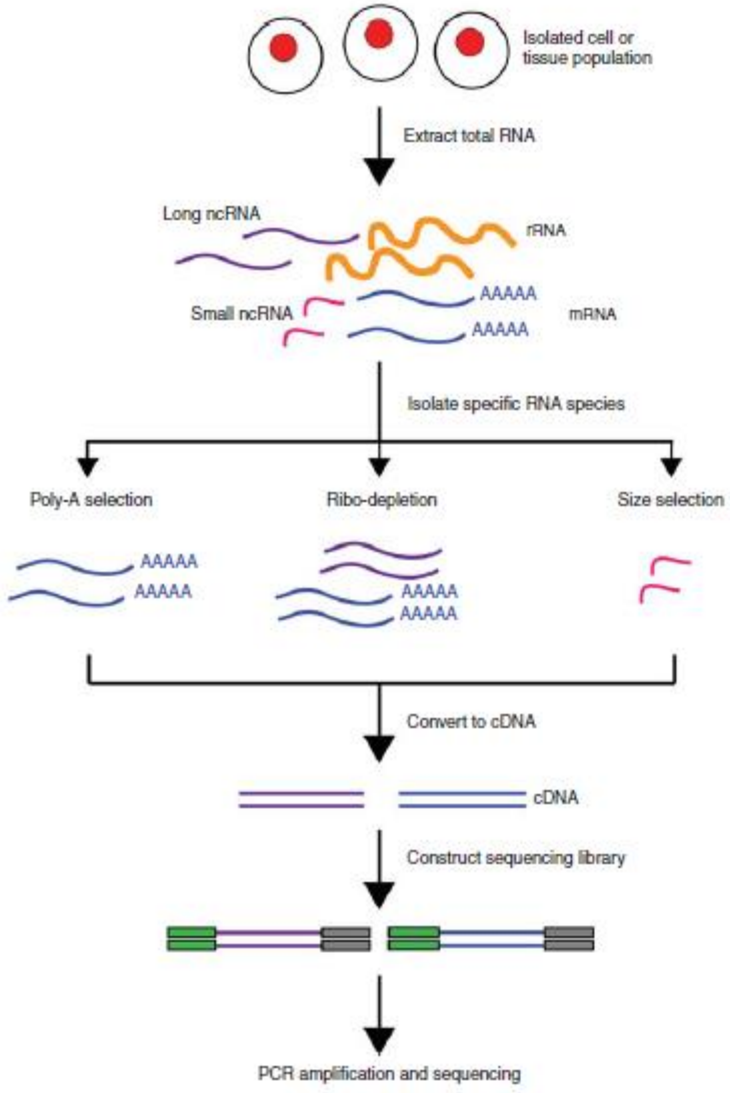
In summary

Differentially expressed genes

Pathway enrichment

Deconvolution

Co-expression analysis



Long non-RNA expression

Target mRNAs

miRNA expression

Any questions?

Andre Goncalves
Postdoctoral Researcher
30th October 2023