



# Differential gene expression analysis using GEO2R

Sina Majidian

Postdoctoral Fellow at the University of Lausanne

SIB Member, Swiss Institute of Bioinformatics.

25 Nov. 2021



UNIL | Université de Lausanne



Swiss Institute  
of  
Bioinformatics

# Index

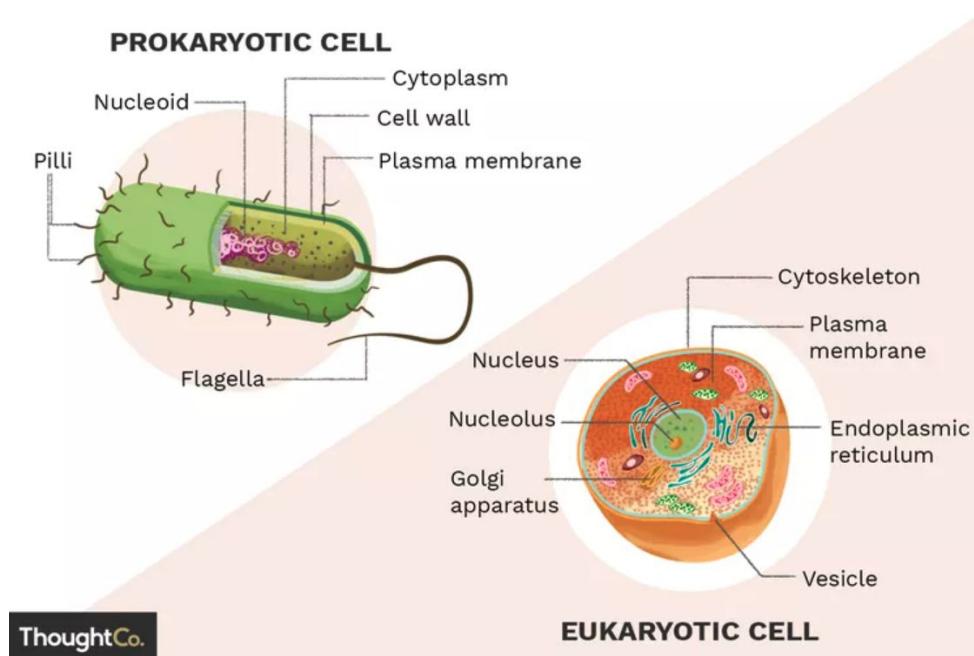
1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

# Index

1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

# Cell

The basic unit from which a living organism is made.



# Genome

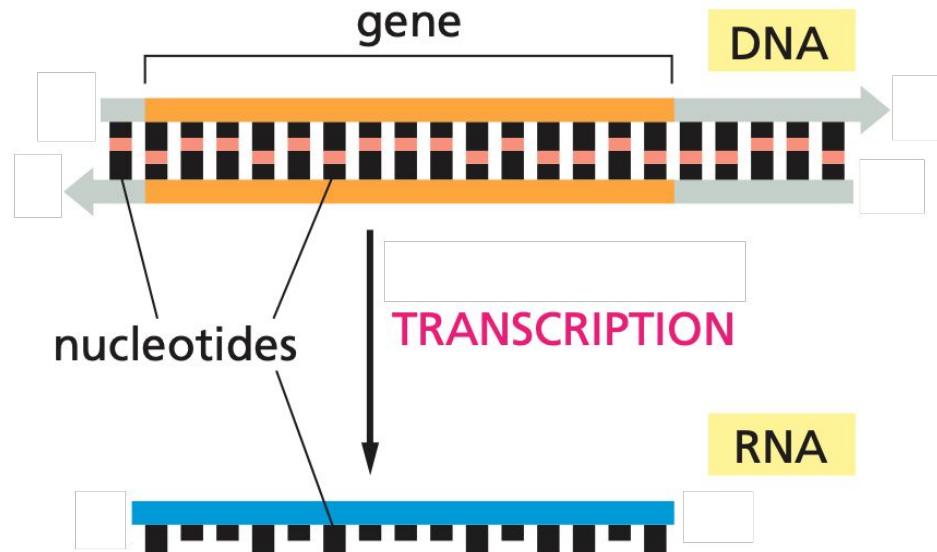
“The complete set of genes or genetic material present in a cell or organism.”

Oxford dictionaries

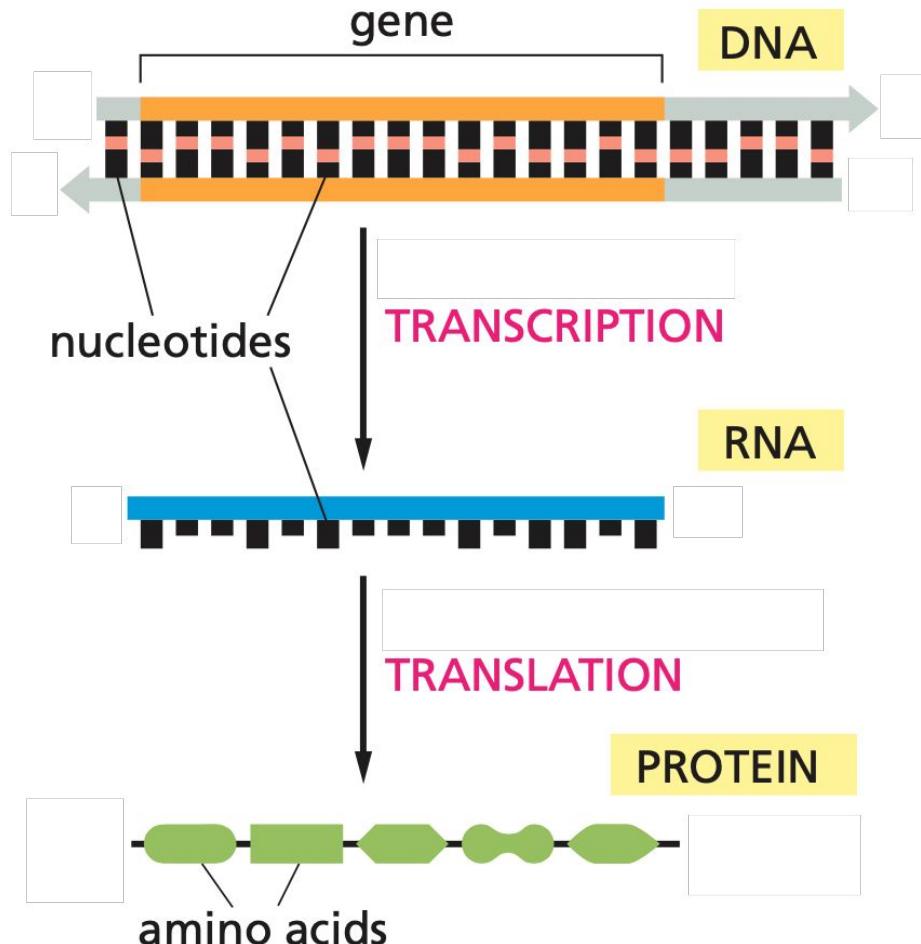
- containing the information needed to maintain organism's living.
- is made of DNA (or RNA in some viruses).

# RNA

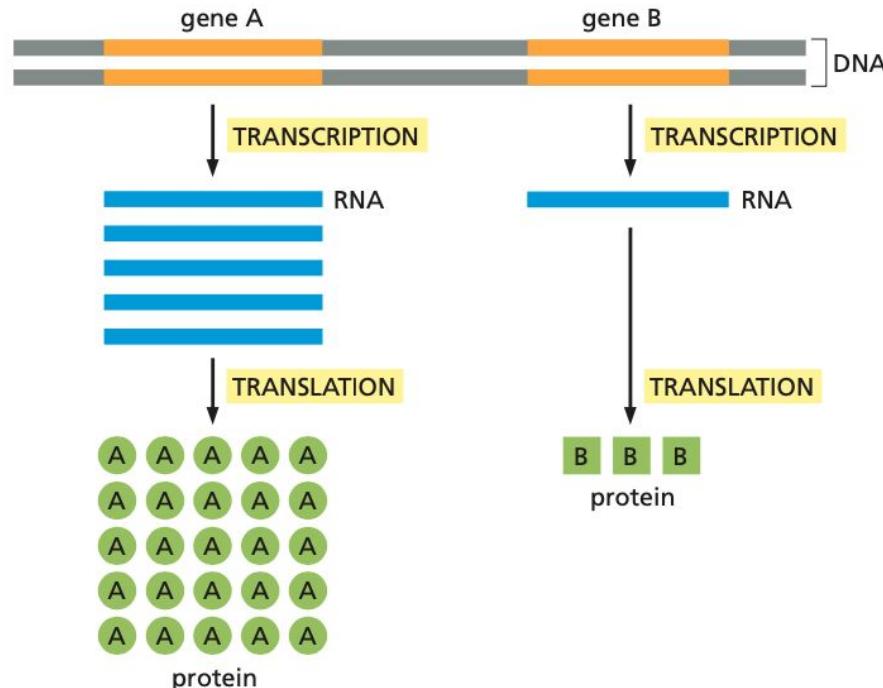
A,T,C,G



A,U,C,G



# A cell can express different genes at different rates.



# Transcriptomics

- Qualitative

Which genes are expressed?

- Quantitative

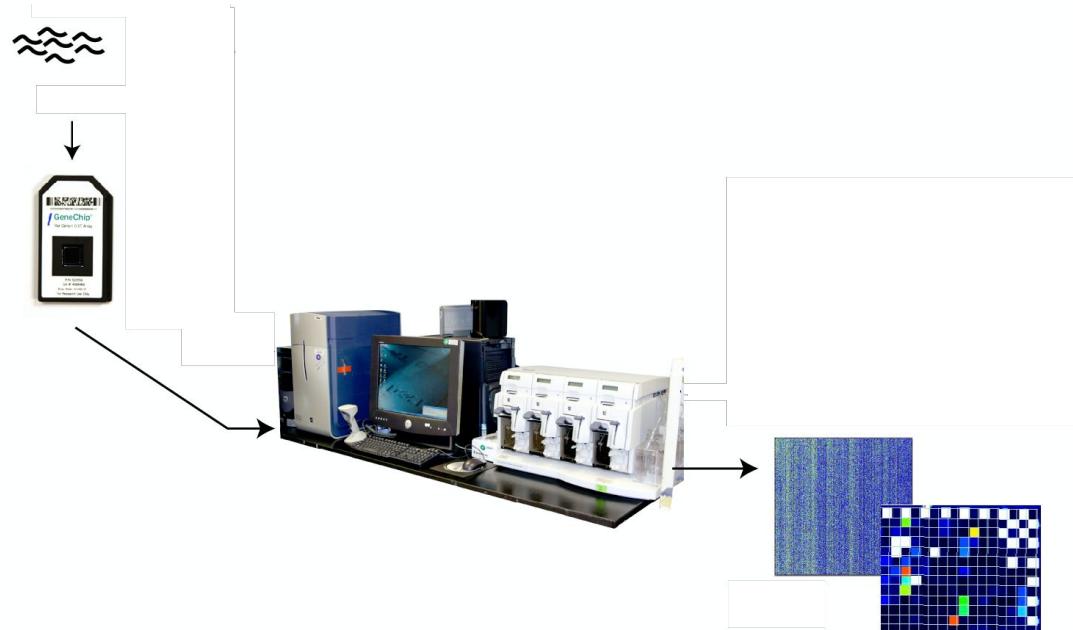
What are the expression levels of the genes?

# Index

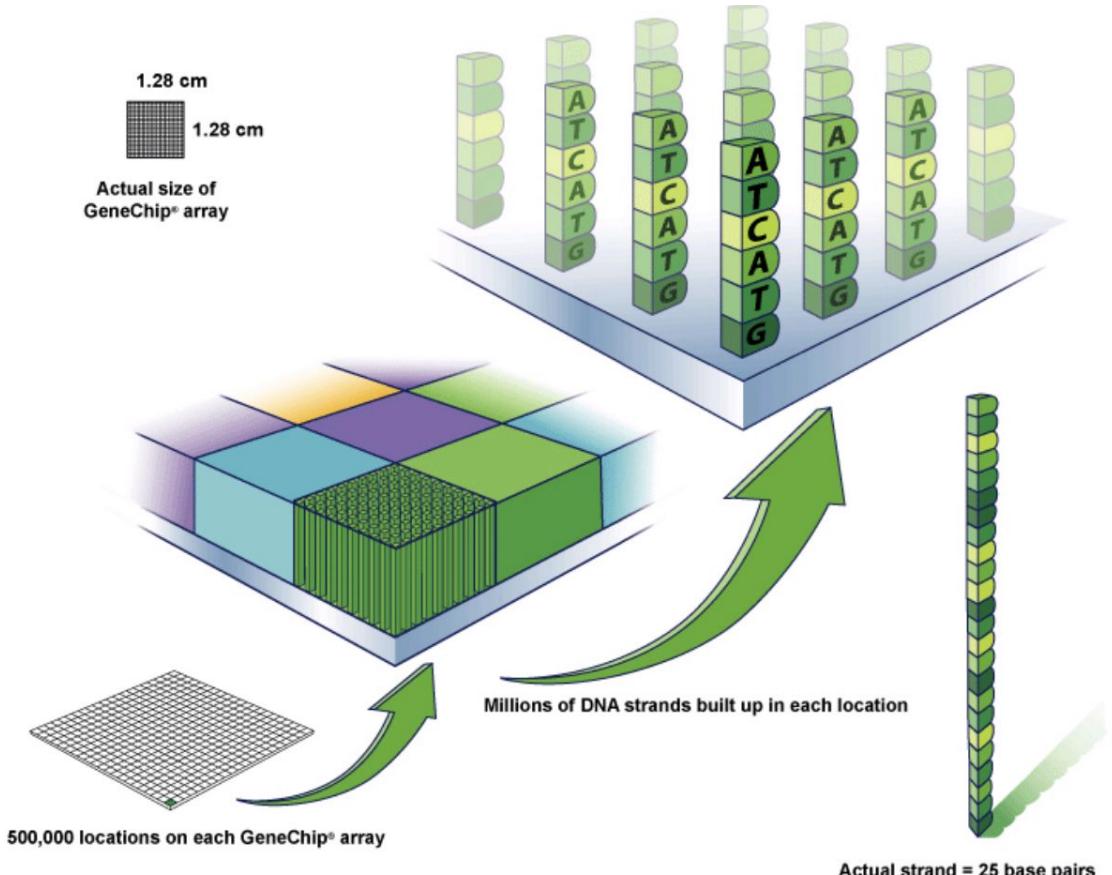
1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

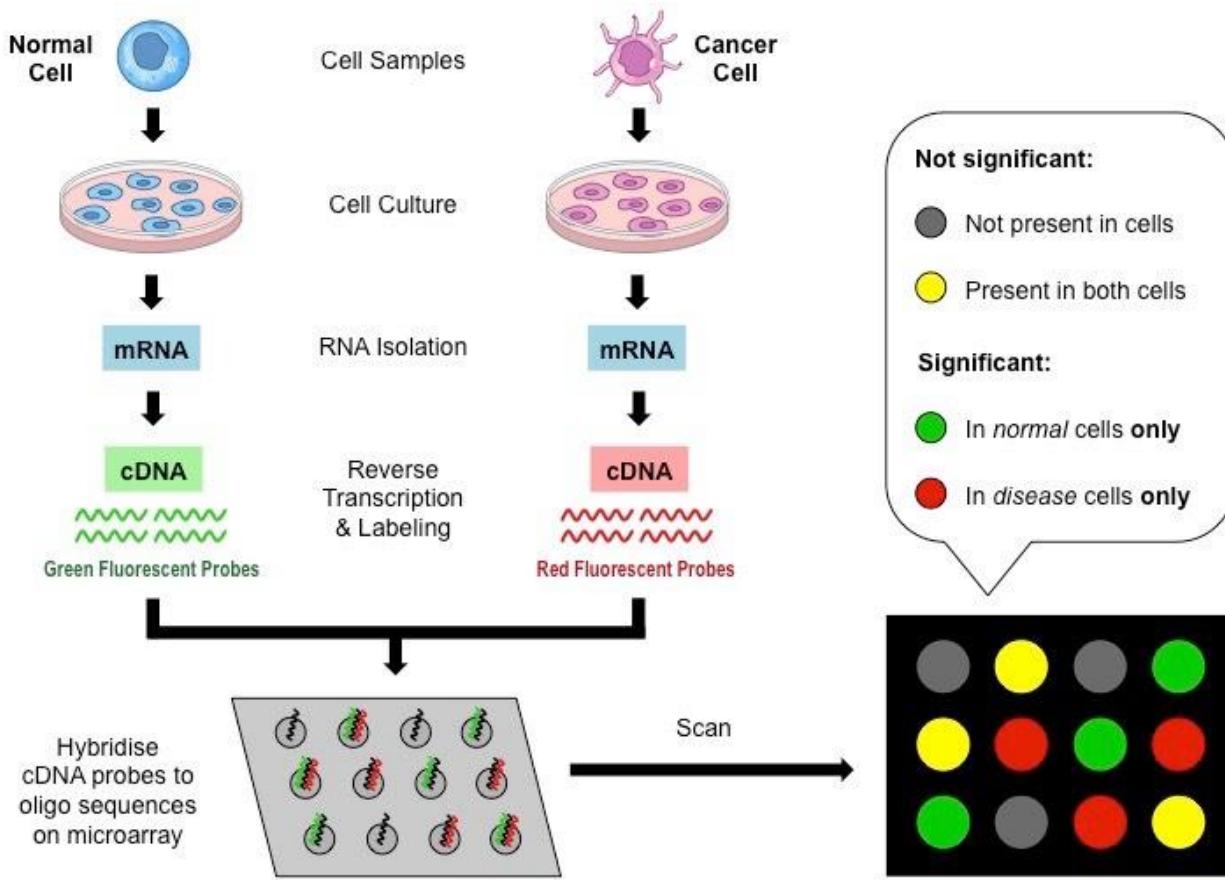
# Microarray

- to study the expression of many genes at once.
- Collect RNA from sample
- Generate complementary DNA from RNA using reverse transcriptase.
- label with a fluorescent dye



- placing thousands of gene sequences in known locations (probes) on a glass slide (gene chip).
- Measure light produced by complementary base pairing between the sample and the chip sequences.

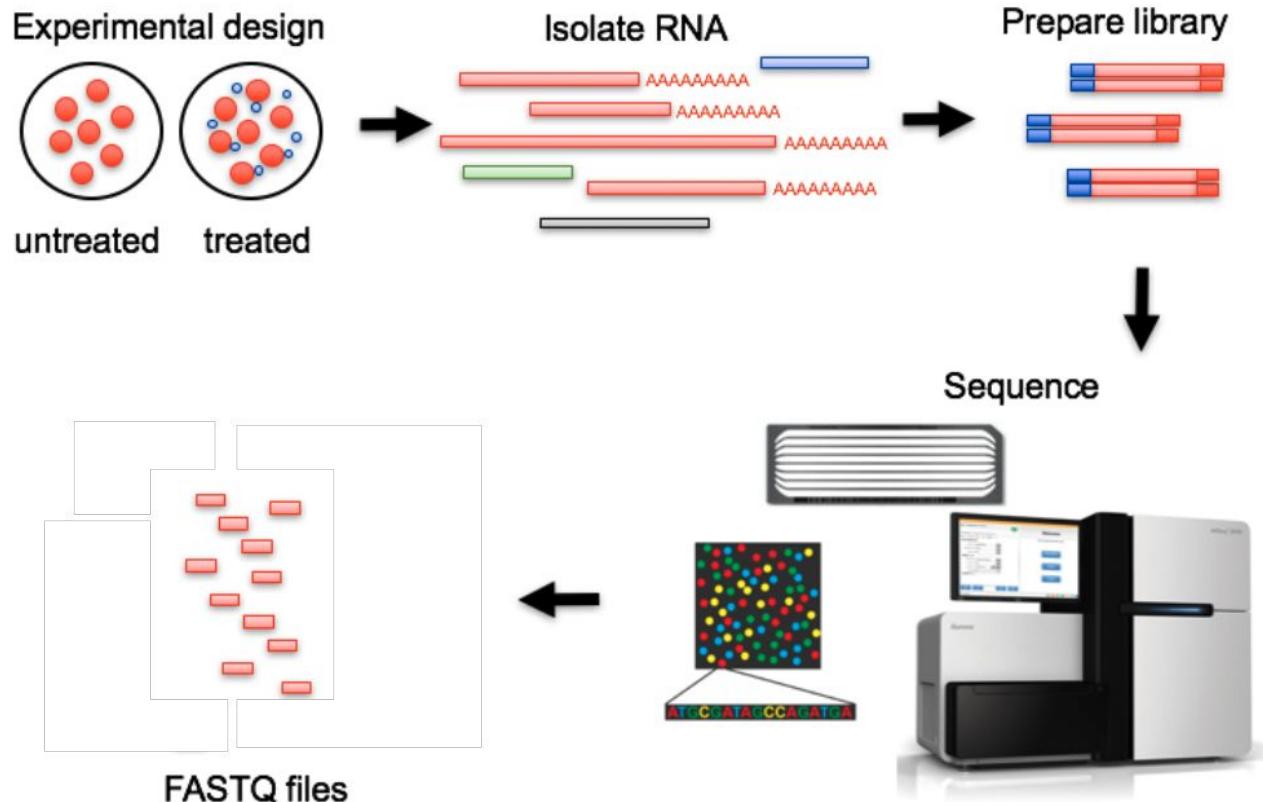




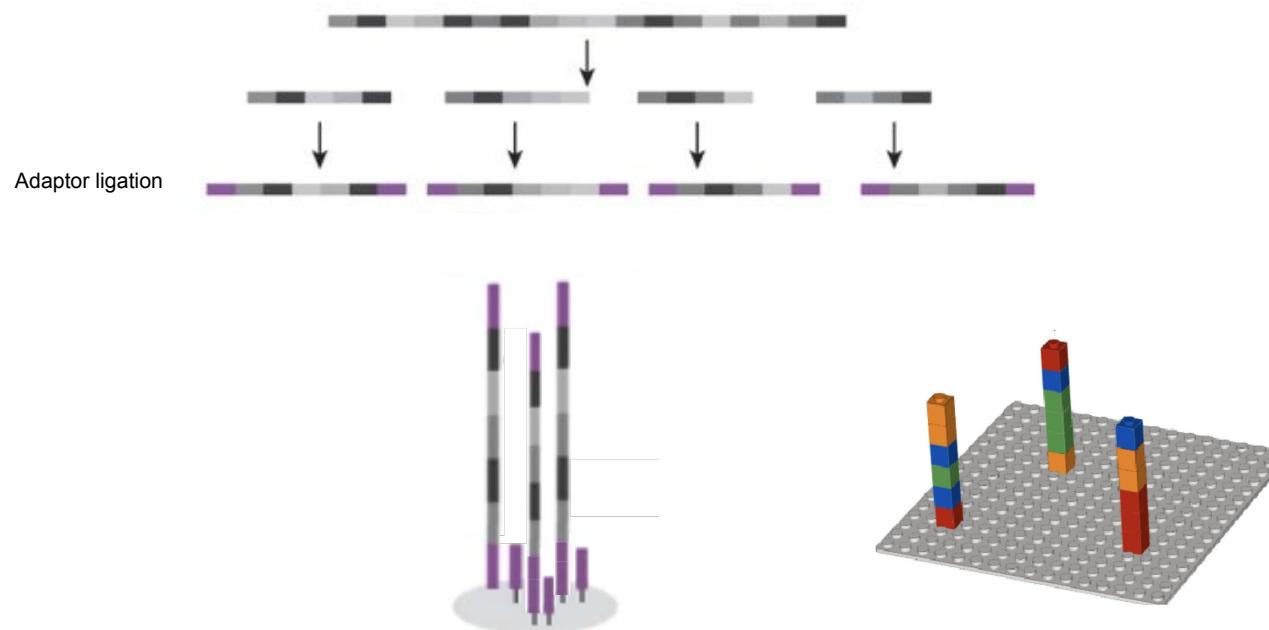
# Index

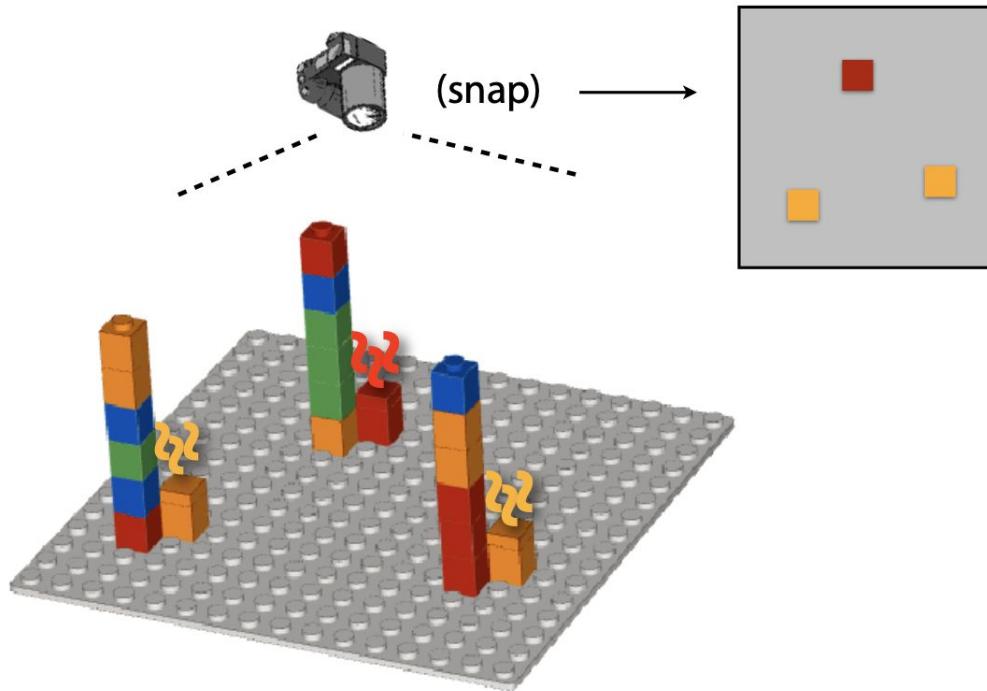
1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

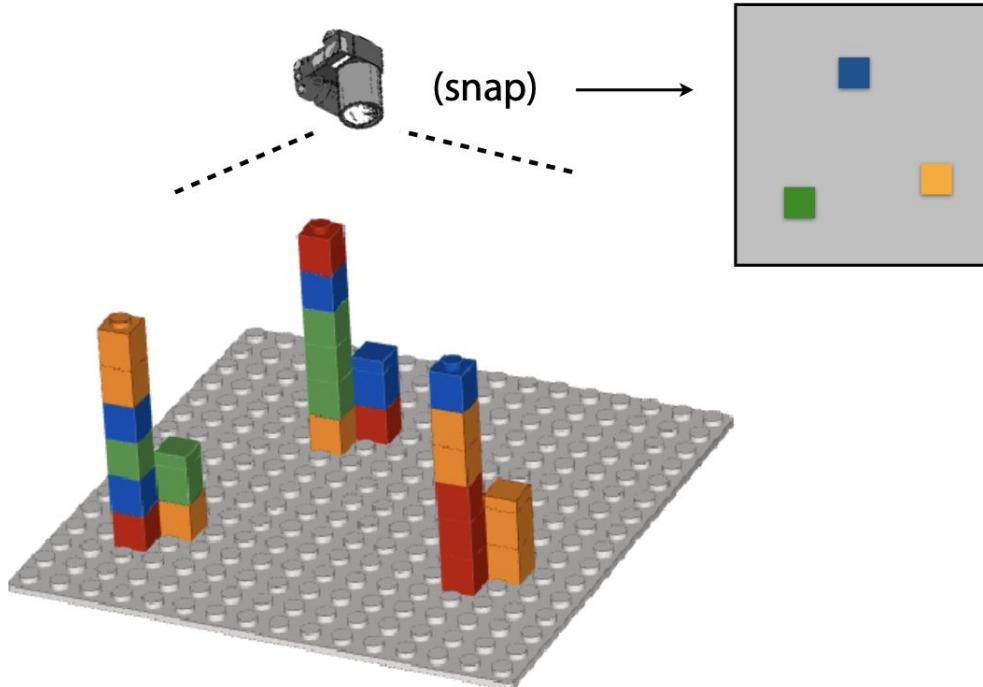
# RNA-seq

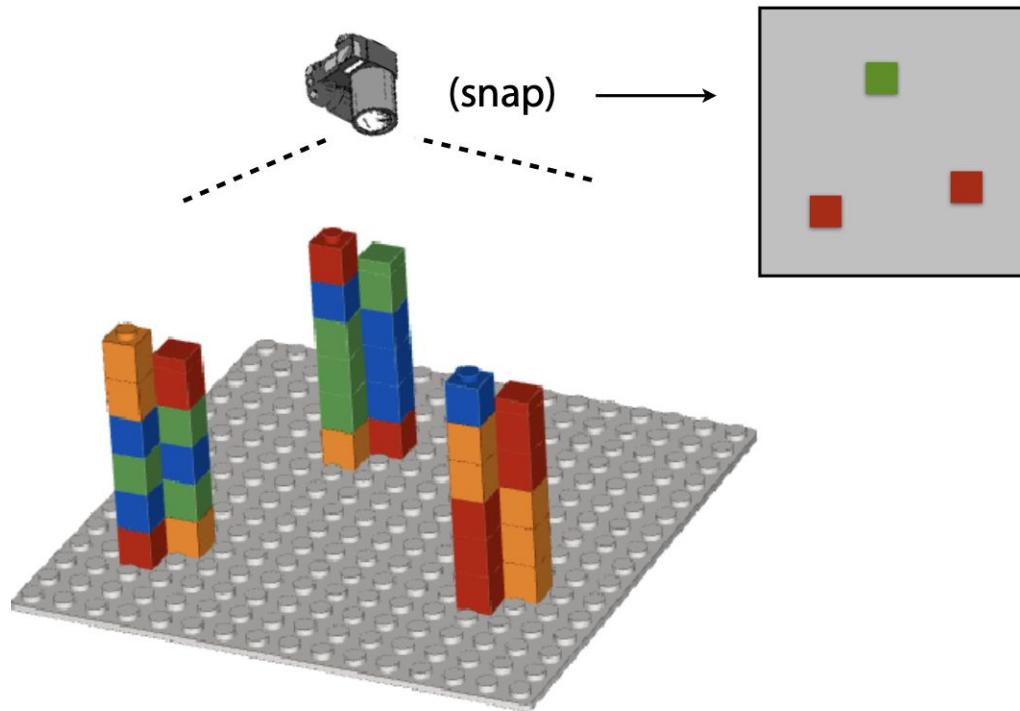


# Illumina sequencing

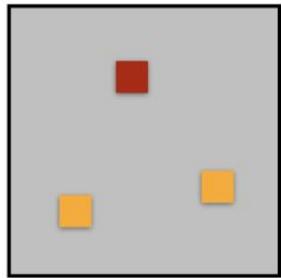




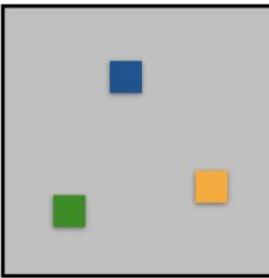




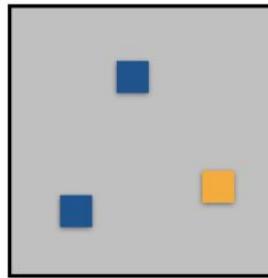
Cycle 1



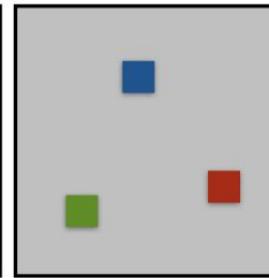
Cycle 2



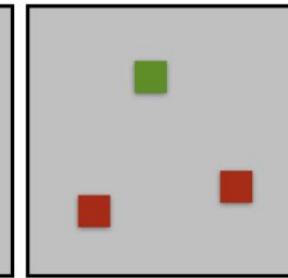
Cycle 3



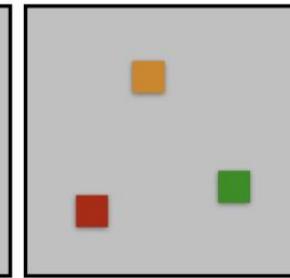
Cycle 4

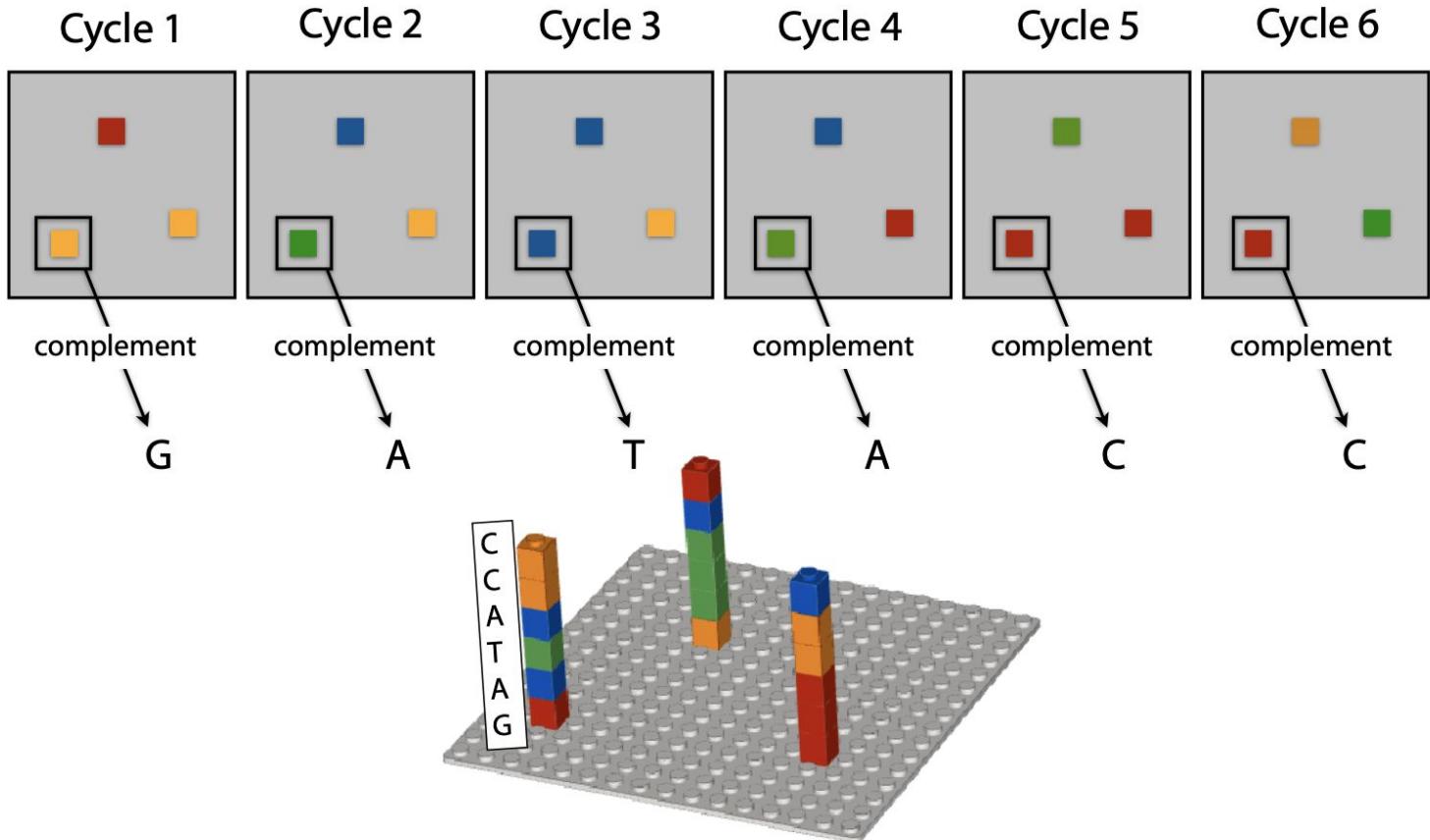


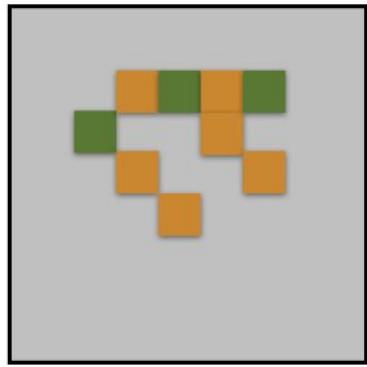
Cycle 5



Cycle 6





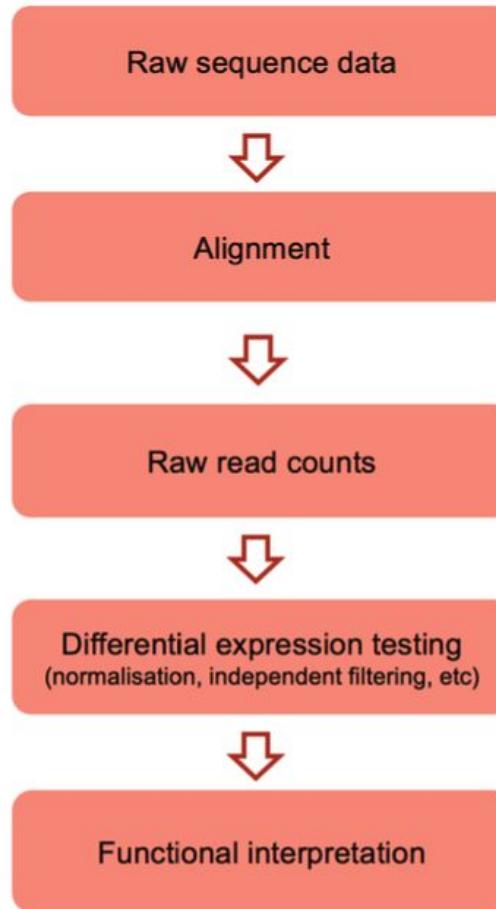


Call: orange (C)

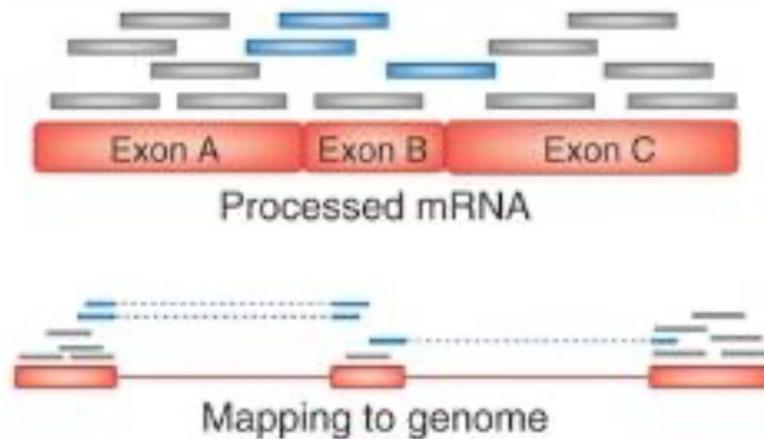
Estimate  $p$ , probability incorrect:  
non-orange light / total light

$$p = 3 \text{ green} / 9 \text{ total} = 1/3$$

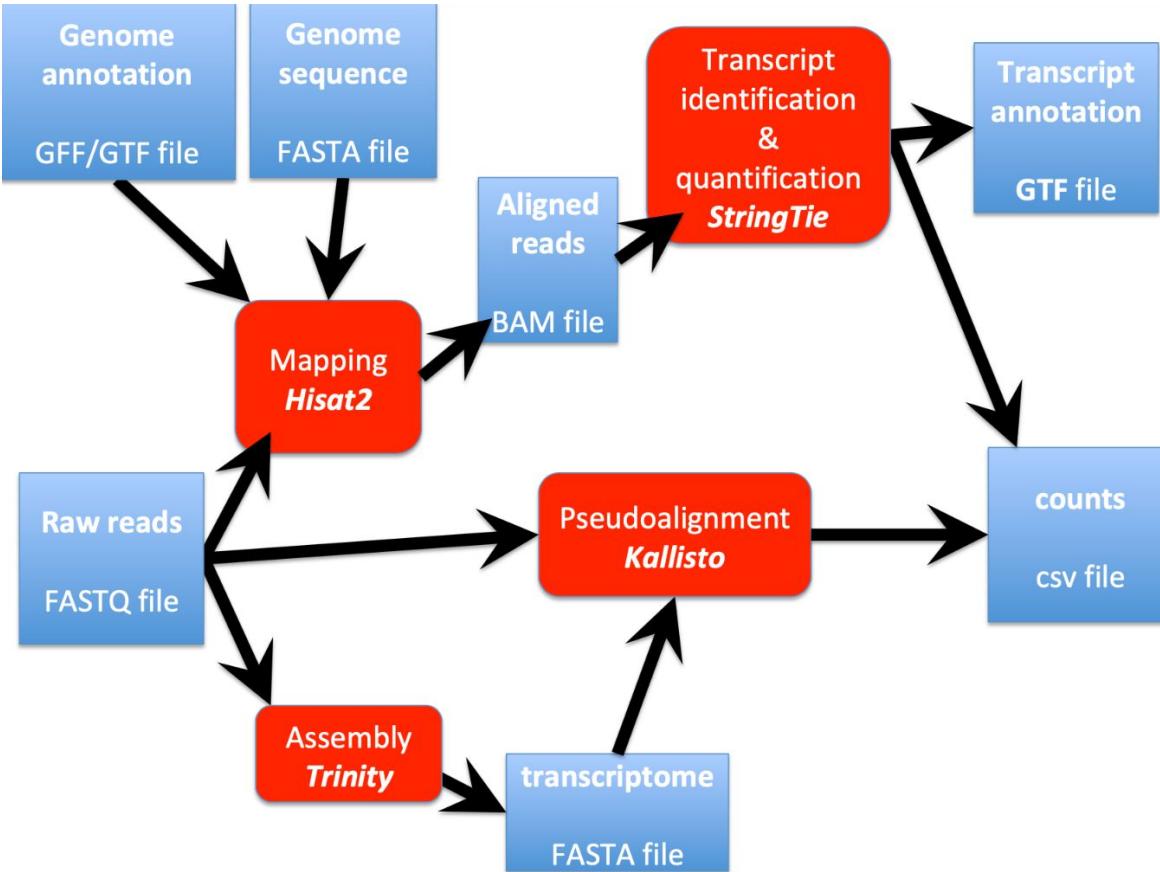
Illumina error rate 0.01%



Mapping RNA-Seq reads is a challenge because of introns, polyA...



HISAT2: <http://ccb.jhu.edu/software/hisat/index.shtml>



Quantification

Map reads to the genome

Exon/intron structure

Alternative splicing: what were the transcripts?

Map reads to the transcriptome Different isoforms

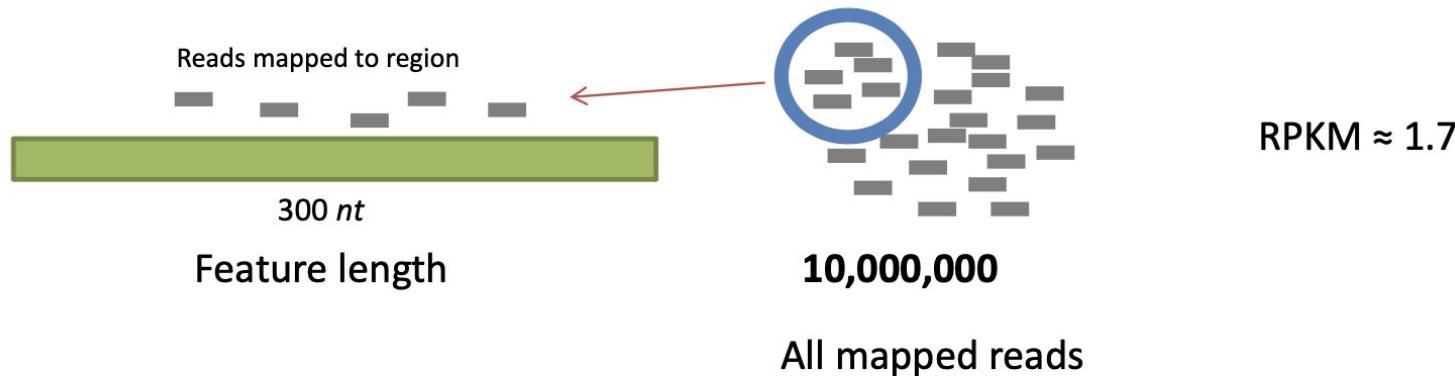
RPKM, FPKM, TPM, CPM Count the reads mapping to a transcript/gene

Compare the same gene between samples: normalize for the total number of mapped reads  
-CPM (counts per million)

Compare different genes/isoforms:

normalize for the length of the genes/isoforms -RPKM, FPKM, TPM (transcripts per million)

# Correcting for transcript length and total number of reads

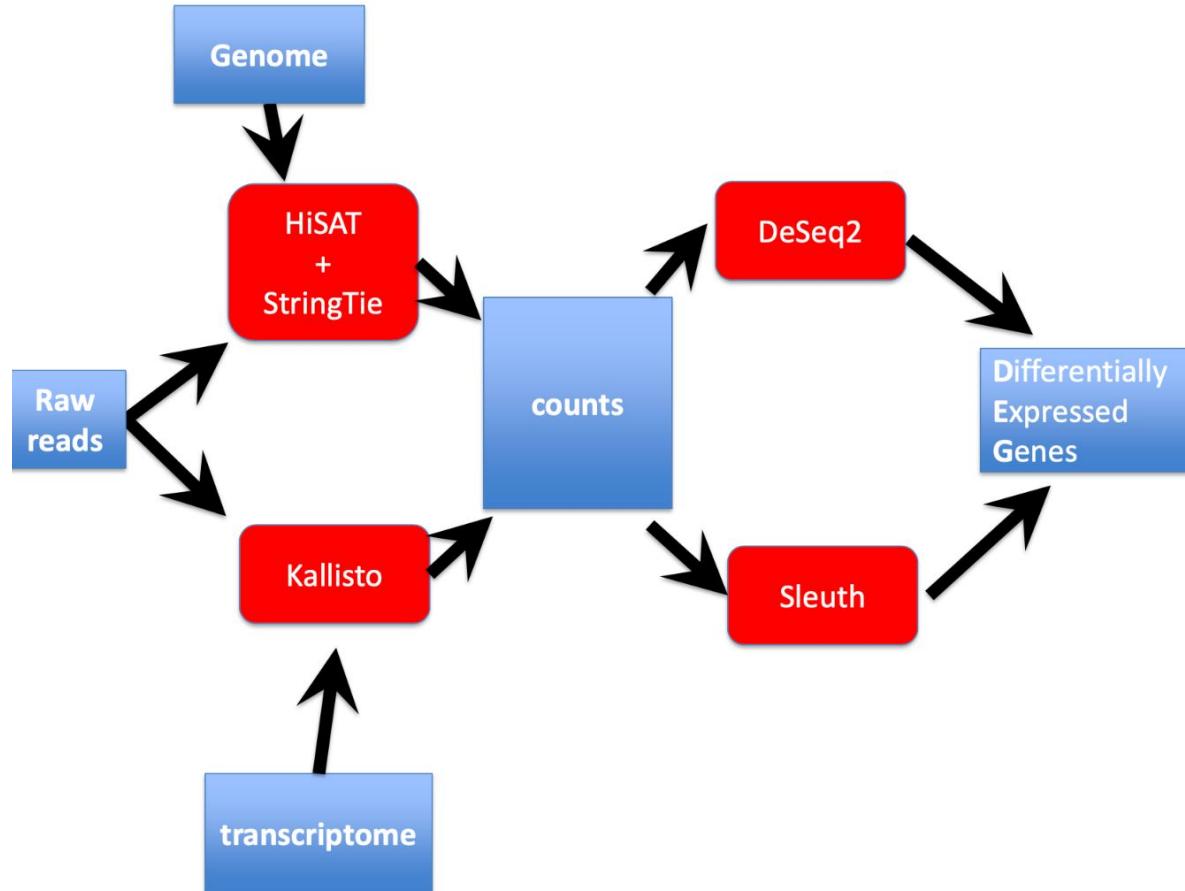


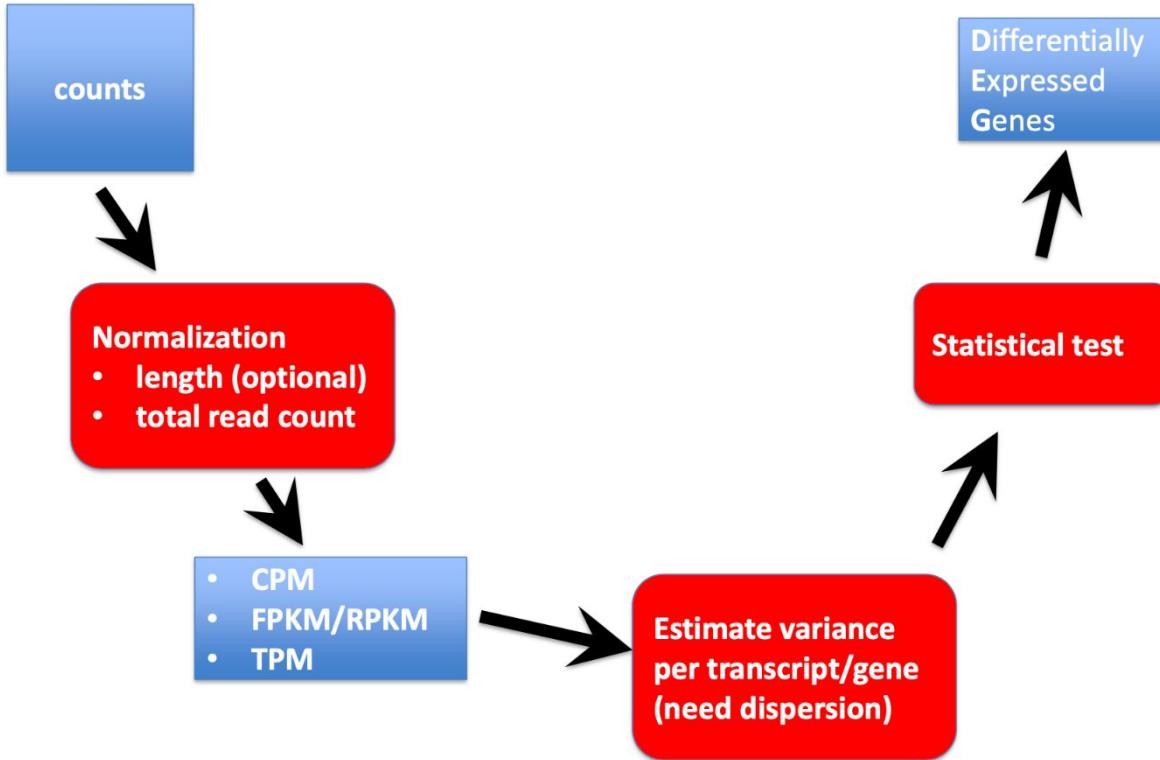
$$RPKM = 10^9 \times \frac{\text{Number of reads mapped to a region}}{\text{Total reads} \times \text{region length}}$$

RPKM: Reads Per Kilo base of transcript per Million reads

# Differential expression

- Which genes are higher/lower expressed
  - between tissues, after treatment ...





# RNA-seq read counting

The number of reads that align to each gene is an estimate of how many RNA transcripts were present in the sample for that gene.

Problems:

Paralogs (reads don't



Reads that span 2 or

Sampling error

# Index

1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. **GEO website**
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

# NCBI's GEO dataset



- Gene Expression Omnibus
- a public functional genomics data repository
- Array- and sequence-based data
- Tools are provided to
  - query and download
  - gene expression profiles

All Images Videos Books News More Tools

About 99,100,000 results (0.67 seconds)

<https://www.ncbi.nlm.nih.gov> › geo

⋮

## Gene Expression Omnibus - NCBI

GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted.

### Submitting data

Gene Expression Omnibus (GEO) is a database repository of high ...

### GEO Profiles

This database stores individual gene expression profiles from ...

### GEO Browser

Gene Expression Omnibus (GEO) is a database repository of high ...

### Login to Submit

Gene Expression Omnibus (GEO) is a database repository of high ...

### About GEO DataSets

Gene Expression Omnibus (GEO) is a database repository of ...

### Analyze with GEO2R

Gene Expression Omnibus (GEO) is a database repository of high ...



## Gene Expression Omnibus



Gene Expression Omnibus is a database for gene expression profiling and RNA methylation profiling managed by the National Center for Biotechnology Information. These high-throughput screening genomics data are derived from microarray or RNA-Seq experimental data. [Wikipedia](#)

**Primary citation:** Edgar R & al. (2002)

**Research center:** National Center for Biotechnology Information



## COVID-19 Information



[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

# Gene Expression Omnibus

GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.



## Getting Started

[Overview](#)[FAQ](#)[About GEO DataSets](#)[About GEO Profiles](#)[About GEO2R Analysis](#)[How to Construct a Query](#)[How to Download Data](#)

## Tools

[Search for Studies at GEO DataSets](#)[Search for Gene Expression at GEO Profiles](#)[Search GEO Documentation](#)[Analyze a Study with GEO2R](#)[Studies with Genome Data Viewer Tracks](#)[Programmatic Access](#)[FTP Site](#)[ENCODE Data Listings and Tracks](#)

## Browse Content

[Repository Browser](#)

DataSets: 4348

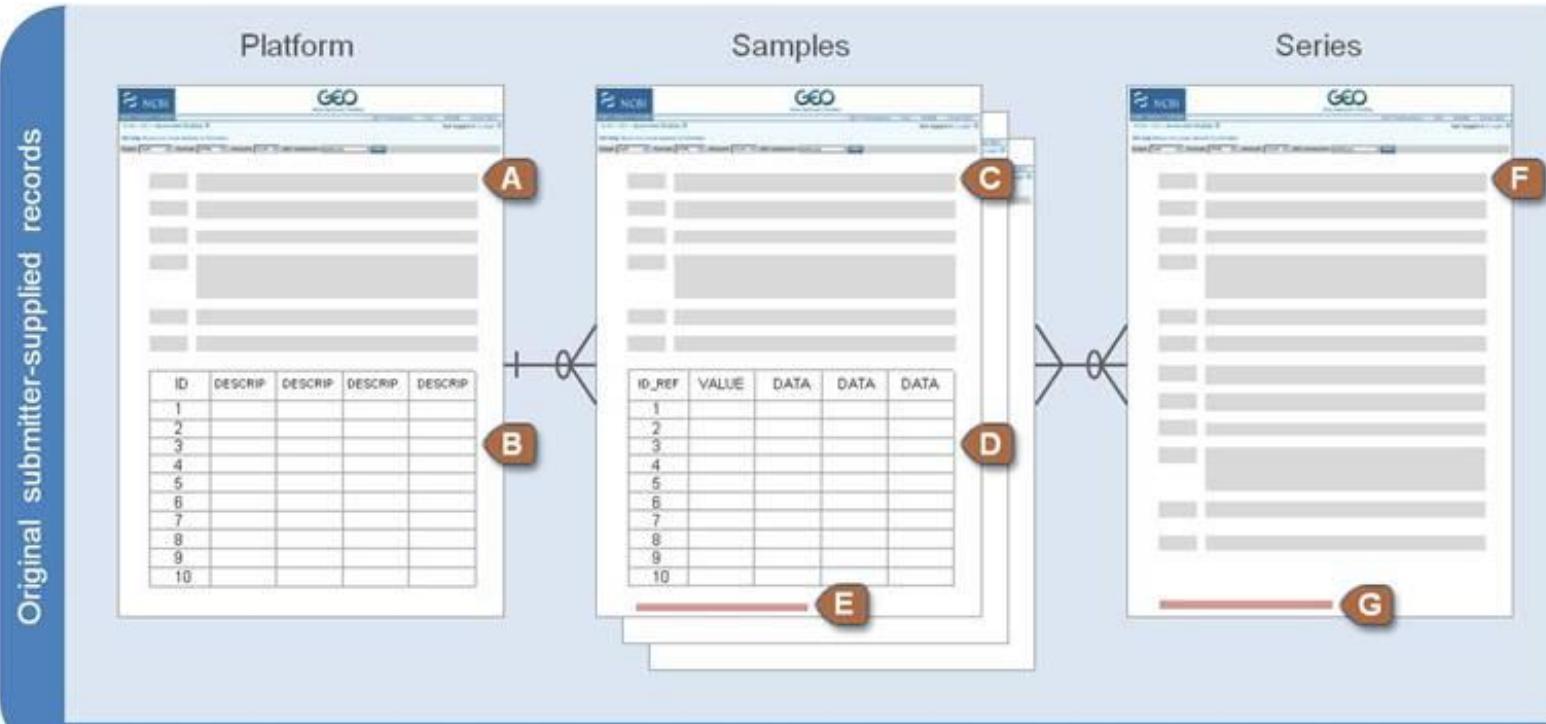
Series: 164748

Platforms: 22745

Samples: 4741813

## Information for Submitters

[Login to Submit](#)[Submission Guidelines](#)[MIAME Standards](#)



<https://www.ncbi.nlm.nih.gov/geo/info/overview.html>

# A sample record

- the conditions under which an individual Sample was handled,
  - the manipulations it underwent,
  - abundance measurement of each element derived from it.
- 
- with a unique and stable GEO accession number (GSMxxx).

COVID-19 is an emerging, rapidly evolving situation.

Get the latest public health information from CDC: <https://www.coronavirus.gov>.

Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

[HOME](#) | [SEARCH](#) | [SITE MAP](#)[GEO Publications](#) | [FAQ](#) | [MIAME](#) | [Email GEO](#)[NCBI](#) > [GEO](#) > [Accession Display](#) Not logged in | [Login](#) 

Scope: Self Format: HTML Amount: Quick GEO accession:

### Sample GSM81022

[Query DataSets for GSM81022](#)

Status	Public on Apr 29, 2006
Title	E19_epithelial cells_control
Sample type	RNA
Source name	E19 type II epithelial cells, 16h, control
Organism	<a href="#">Rattus norvegicus</a>
Characteristics	Strain:CD Gender: male and female Age: Embryonic day 19 Tissue: Fetal lung from time-pregnant rats Cells: Type II epithelial cells
Biomaterial provider	Juan Sanchez-Esteban
Treatment protocol	After isolation, E19 type II cells were cultured on silastic membranes coated with collagen 1. Next day monolayers were maintained on non-strained conditions for 16 h (control)
Growth protocol	Cells were maintained in serum-free DMEM medium.
Extracted molecule	total RNA
Extraction protocol	Total RNA was extracted from E19 type II cells using TRIzol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA) and purified further using the Rneasy Mini Kit (Invitrogen).
Label	biotin conversion of double stranded cDNA into biotin-labeled cRNA was

Submission date Oct 29, 2005  
 Last update date Nov 01, 2005  
 Contact name Juan Sanchez-Esteban  
 E-mail(s) [jsanchezesteban@wihri.org](mailto:jsanchezesteban@wihri.org)  
 Phone 401-2741122 ext1383  
 Fax 401-4537571  
 Organization name Women & Infants Hospital/Brown Medical School  
 Department Pediatrics  
 Street address 101 Dudley Street  
 City Providence  
 State/province RI  
 ZIP/Postal code 02905  
 Country USA

Platform ID [GPL341](#)

Series (1) [GSE3541](#) DNA microarray reveals novel genes induced by mechanical forces in fetal lung type II epithelial cells

#### Data table header descriptions

<b>ID_REF</b>	
<b>VALUE</b>	signal, MAS5 transformation
<b>ABS_CALL</b>	detection
<b>DETECTION P-VALUE</b>	MicroArray Suite 5.0

#### Data table

ID_REF	VALUE	ABS_CALL	DETECTION P-VALUE
1398836_s_at	28617.1	P	0.000244
AFFX_Rat_beta-actin_3_at	27325.3	P	0.000044
1375170_at	27204.5	P	0.000244
1398835_at	27068.4	P	0.000244
AFFX_Rat_beta-actin_M_at	26301.8	P	0.000044
1371297_at	25061.9	P	0.000244
1389956_a_at	24322.5	P	0.000244
1371307_at	24275.7	P	0.000244
1371344_at	24004.9	P	0.000244
1387890_at	23906	P	0.000244
AFFX-r2-P1-cre-3_at	23678.7	P	0.000244
1371318_at	23464.6	P	0.000244
1367579_a_at	23369.9	P	0.000244
1370212_at	23170	P	0.000244

# A Platform record

- composed of a summary description of the array or sequencer
- a data table defining the array template (for array-based Platforms)
- a unique and stable GEO accession number (GPLxxx).

Status	Public on Jun 23, 2003
Title	[RAE230A] Affymetrix Rat Expression 230A Array
Technology type	in situ oligonucleotide
Distribution	commercial
Organism	<a href="#">Rattus norvegicus</a>
Manufacturer	Affymetrix
Manufacture protocol	see manufacturer's web site

Array A of GeneChip Rat Expression Set 230  
Has 15923 entries and was indexed 09-Apr-2003

Sequences used in the design of the array were selected from GenBank, dbEST, and RefSeq. Sequence clusters were created from Build 99 of UniGene (June 2002) and refined by analysis and comparison with a number of other publicly available databases including the Baylor College of Medicine Human Genome Sequencing Center's preliminary rat genome assembly (June 2002). In addition, sequences were analyzed for untrimmed low-quality sequence information, correct orientation, false clustering, alternative splicing and alternative polyadenylation.

Description	Affymetrix submissions are typically submitted to GEO using the GEOarchive method described at <a href="http://www.ncbi.nlm.nih.gov/projects/geo/info/geo_affy.html">http://www.ncbi.nlm.nih.gov/projects/geo/info/geo_affy.html</a>
-------------	--

June 03, 2009: annotation table updated with netaffx build 28  
June 21, 2012: annotation table updated with netaffx build 32  
July 01, 2016: annotation table updated with netaffx build 35

Submission date	Jun 19, 2003
Last update date	Mar 03, 2017
Organization	Affymetrix, Inc.
E-mail(s)	<a href="mailto:geo@ncbi.nlm.nih.gov">geo@ncbi.nlm.nih.gov</a> , <a href="mailto:support@affymetrix.com">support@affymetrix.com</a>
Phone	888-362-2447
URL	<a href="http://www.affymetrix.com/index.affx">http://www.affymetrix.com/index.affx</a>
Street address	
City	Santa Clara
State/province	CA
ZIP/Postal code	95051
Country	USA

Samples (3303) [GSM12793](#), [GSM12795](#), [GSM12797](#), [GSM12799](#), [GSM12801](#), [GSM12803](#)

[+ More...](#)

Series (149) [GSE487](#) PGA Rat Liver Methylprednisolone

[+ More...](#)

[GSE830](#) Rat germ cells

[GSE925](#) A Comparison of Neonatal Rat Ventricular Myocytes Treated With Phenylephrine or PAMH

### Relations

Alternative to [GPL8627](#) (Alternative CDF)

Alternative to [GPL14122](#) (Alternative CDF)

### Data table header descriptions

**ID** Affymetrix Probe Set ID

**GB\_ACC** GenBank Accession Number

**SPOT\_ID** identifies controls

**Species Scientific Name** The genus and species of the organism represented by the probe set.

**Annotation Date** The date that the annotations for this probe array were last updated. It will generally be earlier than the date when the annotations were posted on the Affymetrix web site.

**Sequence Type**

**Sequence Source** The database from which the sequence used to design this probe set was

Scope: **Self** Format: **HTML** Amount: **Quick** GEO accession: **GPL70** **GO**

## Platform GPL70

Query DataSets for GPL70

Status Public on Feb 22, 2002  
Title Affymetrix roDROMEGAa Drosophila full genome  
Technology type in situ oligonucleotide  
Distribution custom-commercial  
Organism [Drosophila melanogaster](#)

Annotation SOFT table...

**Download family**

SOFT formatted family file(s)  
MINIML formatted family file(s)

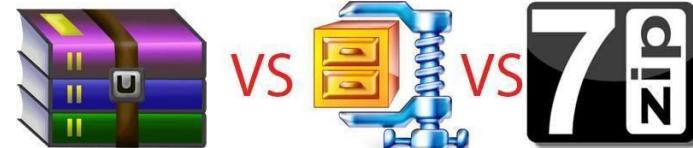
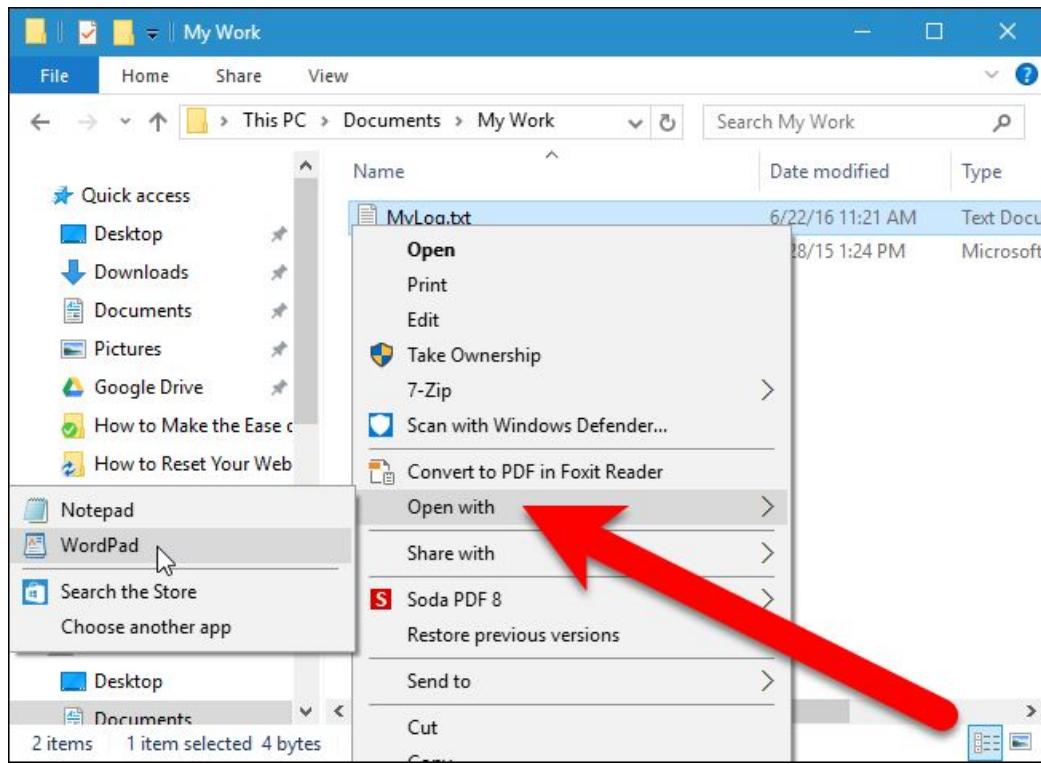
**Format**

SOFT   
MINIML 

**Supplementary data files not provided**

| [NLM](#) | [NIH](#) | [GEO Help](#) | [Disclaimer](#) | [Accessibility](#) |

# Open with Notepad/Excel



```
^Annotation
!Annotation_date = Aug 09 2016
!Annotation_platform = GPL70
!Annotation_platform_title = Affymetrix roDRomeGAa Drosophila full genome
!Annotation_platform_organism = Drosophila melanogaster
#ID = ID from Platform data table
#Gene title = Entrez Gene name
#Gene symbol = Entrez Gene symbol
#Gene ID = Entrez Gene identifier
#UniGene title = Entrez UniGene name
#UniGene symbol = Entrez UniGene symbol
#UniGene ID = Entrez UniGene identifier
#Nucleotide Title = Entrez Nucleotide title
#GI = GenBank identifier
#GenBank Accession = GenBank accession
#Platform_CLONEID = CLONE_ID from Platform data table
#Platform_ORF = ORF from Platform data table
#Platform_SPOTID = SPOT_ID from Platform data table
#Chromosome location = Entrez gene chromosome and location
#Chromosome annotation = Entrez gene chromosome annotation
#GO:Function = Gene Ontology Function term
#GO:Process = Gene Ontology Process term
#GO:Component = Gene Ontology Component term
```

G0:0005634					
CG4794_at	CG31213 gene product from transcript CG31213-RB	CG31213	42382	CG4794	
92C1-92C1	Chromosome 3R, NT_033777.3 (19972898..19979569)	ATP binding//ATPase activity, coupled to transmembrane movement of substances			
transport	integral component of membrane//membrane	G0:0005524///G0:0042626	G0:0006810	G0:0016021///G0:0016020	
CG4796_at	curled cu	41339	CG4796	3-50.0 cM	
NT_033777.3 (11197592..11208633)	locomotor rhythm//nuclear-transcribed mRNA poly(A) tail shortening//response to light stimulus//sensory perception of pain	CCR4-NOT complex//cytoplasm//lipid particle	G0:0045475///G0:000289///G0:0009416///G0:0019233		
G0:0030014///G0:0005737///G0:0005811					
CG4797_at	CG4797 gene product from transcript CG4797-RB	CG4797	37770	CG4797	
60A2-60A2	Chromosome 2R, NT_033778.4 (23812185..23818327, complement)	transport integral component of membrane	G0:0005355 G0:0055085	glucose transmembrane transporter activity	transmembrane
CG4798_s at	lethal (2) k01209	l(2)k01209	36953	CG4798	54B7-54B7

ID	Gene title	Gene symbol
CG11045_at	Equilibrative nucleoside transporter 2	Ent2
CG11046_at	Ecdysone-inducible gene	E3

CG11074_at	CG30385	ger CG30385///(246585//246584//3564	CG11074		43A1-43A1// Chromosome molecular function biological_process cellular_com	GO:0003674	GO:0008150		
CG11075_at	maternal ger mamo		32353	CG11075	12D1-12D1	Chromosome chromatin binding female meiotic cytoplasm//	GO:0003682	GO:0007143	
CG11076_at	CG11076	ger CG11076	43830	CG11076	102F6-102F6	Chromosome 4, NC_004353.4 (1029711..1031757, complement)			
CG11077_at	CG11077	ger CG11077	43831	CG11077	102F6-102F6	Chromosome molecular function homeostatic integral component	GO:0003674	GO:0042592	
CG11079_at	CG34423	ger CG34423///(5740313//37699	CG11079		59D8-59D8// Chromosome	ATPase inhibitory negative regulation	mitochondria	GO:0042030	GO:0032780
CG11082_at	maternal ger mamo		32353	CG11082	12D1-12D1	Chromosome chromatin binding female meiotic cytoplasm//	GO:0003682	GO:0007143	
CG11084_at	prickle	pk	45343	CG11084	2-55.3 cM	Chromosome protein binding establishment	meiotic cytoplasm//	GO:0005515	GO:0001737
CG11085_at	CG11085	ger CG11085	32213	CG11085	11B4-11B4	Chromosome sequence-specific regulation	of transcription	GO:0043565	GO:0006355
CG11086_at	CG11086	ger Gadd45	35646	CG11086	43A2-43A2	Chromosome 2R, NT_033	JNK cascade, nucleus		GO:0007254
CG11089_at	CG11089	ger CG11089	42973	CG11089	96B2-96B2	Chromosome IMP cyclohydrolase	purine nucleotide biosynthesis	GO:0003937	GO:0006164/
CG1109_at	CG1109	gene	CG1109	40698	CG1109	83B8-83B8	Chromosome 3R, NT_033	neuron projection morphogenesis	GO:0048812
CG11091_at				CG11091					
CG11093_at	fussel	fuss	43835	CG11093	102F4-102F4	Chromosome co-SMAD binding	imaginal disc cytoplasm	GO:0070410	GO:0007474
CG11095_at	CG11095	ger CG11095	32348	CG11095	12C5-12C5	Chromosome hydrolase activity		GO:0016787	
CG11096_at	inactivation	inaE	32343	CG11096	12C4-12C5	Chromosome lipoprotein binding	lipid metabolism	GO:0004465	GO:0006629
CG11097_s_	elf4E-Transp	4E-T	43836	CG11097	102F3-102F3	Chromosome 4, NC_004355	sensory perception cytoplasm//cytoplasmic	GO:0019233	
CG11098_at	Transport	an Tango1	33930	CG11098	26F3-26F5	Chromosome chaperone binding	Golgi organanelle	GO:0051087	GO:0007030
CG11099_at	CG11099	ger CG11099	37282	CG11099	56F10-56F11	Chromosome 2R, NT_033778.4 (20236235..20239200)			
CG11100_at	Mesoderm	-e Mes2	40514	CG11100	80C1-80C1	Chromosome molecular function	ecdysis, chitin nucleus//nucleus	GO:0003674	GO:0018990
CG11101_at	pawn	pwn	44011	CG11101	2-55.4 cM	Chromosome calcium ion binding	chaeta morphogenesis//	GO:0005509	GO:0008407/
CG11102_at	inactivation	inaE	32343	CG11102	12C4-12C5	Chromosome lipoprotein binding	lipid metabolism	GO:0004465	GO:0006629
CG11103_at	CG11103	ger CG11103	32342	CG11103	12C4-12C4	Chromosome molecular function	biological_process cellular_component	GO:0003674	GO:0008150
CG11104_at	inactivation	inaE	32343	CG11104	12C4-12C5	Chromosome lipoprotein binding	lipid metabolism	GO:0004465	GO:0006629
CG11105_at	CG44422	ger CG44422	32063	CG11105	10B2-10B2	Chromosome calcium ion binding		GO:0005509	

# Exercise

What is the probe ID for the gene **GAPDH** in the array **GPL570**?

Status	Public on Nov 07, 2003
Title	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
Technology type	in situ oligonucleotide
Distribution	commercial
Organism	<i>Homo sapiens</i>
Manufacturer	Affymetrix
Manufacture protocol	see manufacturer's web site

Complete coverage of the Human Genome U133 Set plus 6,500 additional genes for analysis of over 47,000 transcripts  
 All probe sets represented on the GeneChip Human Genome U133 Set are identically replicated on the GeneChip Human Genome U133 Plus 2.0 Array. The sequences from which these probe sets were derived were selected from GenBank®, dbEST, and RefSeq. The sequence clusters were created from the UniGene database (Build 133, April 20, 2001) and then refined by analysis and comparison with a number of other publicly available databases, including the Washington University EST trace repository and the University

1552258_at	NM_052871	Homo sapiens Oct 6, 2014	Con
1552261_at	NM_080735	Homo sapiens Oct 6, 2014	Con
1552263_at	NM_138957	Homo sapiens Oct 6, 2014	Con
1552264_a_at	NM_138957	Homo sapiens Oct 6, 2014	Con

Total number of rows: **54675**

Table truncated, full table size **77582 Kbytes**.

[Download full table...](#)

[Annotation SOFT table...](#)

2122891_at	/glyceraldehyde-3-phosphate dehydrogenase	GAPDH	2597	601346543F1 NIH_MCC_8 Homo sapiens cDNA clone IMAGE:3679498 5'-mRNA sequence	9805199 B561479	12p13 Chromosome 12, NC_000812.12 (6534405..6538375)	NAD binding//NADP binding//glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity//glyceraldehyde-3-phosphate dehydrogenase (NAD+) (phosphorylating) activity//identical protein binding//microtubule binding//peptidyl-cysteine S-nitrosylase activity//protein binding canonical glycolysis//cellular response to interferon-gamma//gluconeogenesis//microtubule cytoskeleton organization//negative regulation of translation//negative regulation of translation//neuron apoptosis process//peptidyl-cysteine S-trans-nitrosylation//protein stabilization//regulation of macroautophagy GAIT complex//cytoplasm//cytoplasm/cytosol//cytosol//extracellular exosome//extracellular matrix//intracellular membranebounded organelle//intracellular ribosomes//secretion process//vesicle transport//nucleus/nuclear envelope//nucleus/nuclear membrane//nucleoplasm//plasma membrane//vesicle G0:0051287//G0:0050651//G0:0004365//G0:0004365//G0:0042802//G0:0008017//G0:0035605//G0:0085515 G0:0061621//G0:0071346//G0:0006904//G0:000226//G0:0017148//G0:0017148//G0:0051402//G0:0035606//G0:0050821//G0:0016241 G0:0097452//G0:0005737//G0:0005829//G0:0005829//G0:0070062//G0:0031012//G0:0042321//G0:0030529//G0:0005811//G0:0016020//G0:0015630//G0:0031965//G0:0005634//G0:0005634//G0:0048471//G0:0008868//G0:0031982
212582_a	/oxysterol binding protein like 8	OSBPPL	114882	Homo sapiens mRNA; cDNA DKFZp564E2282 (from clone DKFZp564E2282)	4884169 AL049923	12q11 Chromosome 12, NC_000812.12 (76351797..76559899)	complement C8 cholesterolemia//cholesterol ester binding//phosphatidylinositol binding//phosphatidylinositol 4-phosphate binding//positive regulation of protein kinase B activity//cell differentiation//negative regulation of cell migration//negative regulation of sequestering of triglyceride//phospholipid transport//positive regulation of glucose import in response to insulin stimulus//positive regulation of insulin receptor signaling pathway//insensitive regulation of protein kinase R epsilon//protein localization to nuclear envelope endoplasmic reticulum membrane//internal

# A Series record

- links together a group of related Samples
- provides description of the whole study
- contain tables describing extracted data, summary conclusions, or analyses.
- unique and stable GEO accession number (GSExxx).

COVID-19 is an emerging, rapidly evolving situation.

Get the latest public health information from CDC: <https://www.coronavirus.gov>.

Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

[HOME](#) | [SEARCH](#) | [SITE MAP](#)[GEO Publications](#)[FAQ](#)[MIAME](#)[Email GEO](#)NCBI > GEO > **Accession Display** Not logged in | [Login](#) Scope: SelfFormat: HTMLAmount: QuickGEO accession: **Series GSE3541**[Query DataSets for GSE3541](#)

Status Public on Apr 29, 2006

Title DNA microarray reveals novel genes induced by mechanical forces in fetal lung type II epithelial cells

Organism **Rattus norvegicus**

Experiment type Expression profiling by array

Summary Mechanical forces are essential for normal fetal lung development. However, the cellular and molecular mechanisms regulating this process remain largely unknown. In the present study, we used oligonucleotide microarray technology to investigate gene expression profile in cultured E19 rat fetal lung type II epithelial cells exposed to a level of mechanical strain similar to that observed *in utero*. Significance Analysis of Microarrays (SAM) identified 92 genes differentially expressed by strain. Interestingly, several members of the solute carrier family of amino acid transporters, genes involved in amino acid synthesis and development, and amiloride-sensitive epithelial sodium channel gene were induced by strain. These results were confirmed by quantitative real-time polymerase chain reaction (qRT-PCR). Thus, this study identifies genes induced by strain that may be important for amino acid signaling pathways, protein synthesis and development in fetal type II cells. In addition, these data suggest that mechanical forces may contribute to facilitate lung fluid reabsorption in preparation for birth. Taken together, the present investigation provides further insights into how mechanical forces may modulate fetal lung development.  
Keywords: lung development, fetal type II epithelial cells, strain response, microarray

Overall design

Freshly isolated E19 fetal type II epithelial cells were plated on silastic membranes coated with collagen 1 and exposed to mechanical strain for 16 h to simulate mechanical forces *in utero* (5% cyclic strain, 60 cycles/min for 15 min + 2.5% continuous distention for the remaining 45 min of each hour) using the Flexercell FX-4000 Strain Unit. Cells grown on non-strained substrates

Submission date Oct 31, 2005  
Last update date Mar 03, 2017  
Contact name Juan Sanchez-Esteban  
E-mail(s) [jsanchezesteban@wihri.org](mailto:jsanchezesteban@wihri.org)  
Phone 401-2741122 ext1383  
Fax 401-4537571  
Organization name Women & Infants Hospital/Brown Medical School  
Department Pediatrics  
Street address 101 Dudley Street  
City Providence  
State/province RI  
ZIP/Postal code 02905  
Country USA

Platforms (1) [GPL341 \[RAE230A\] Affymetrix Rat Expression 230A Array](#)

Samples (6) [GSM81022 E19\\_epithelial cells\\_control1](#)

[More...](#) [GSM81159 E19\\_epithelial cells\\_control2](#)

[GSM81160 E19\\_epithelial cells\\_control3](#)

### Relations

BioProject [PRJNA93603](#)

Analyze with GEO2R

### Download family

[SOFT formatted family file\(s\)](#)

### Format

SOFT [?](#)

[MINiML formatted family file\(s\)](#)

MINiML [?](#)

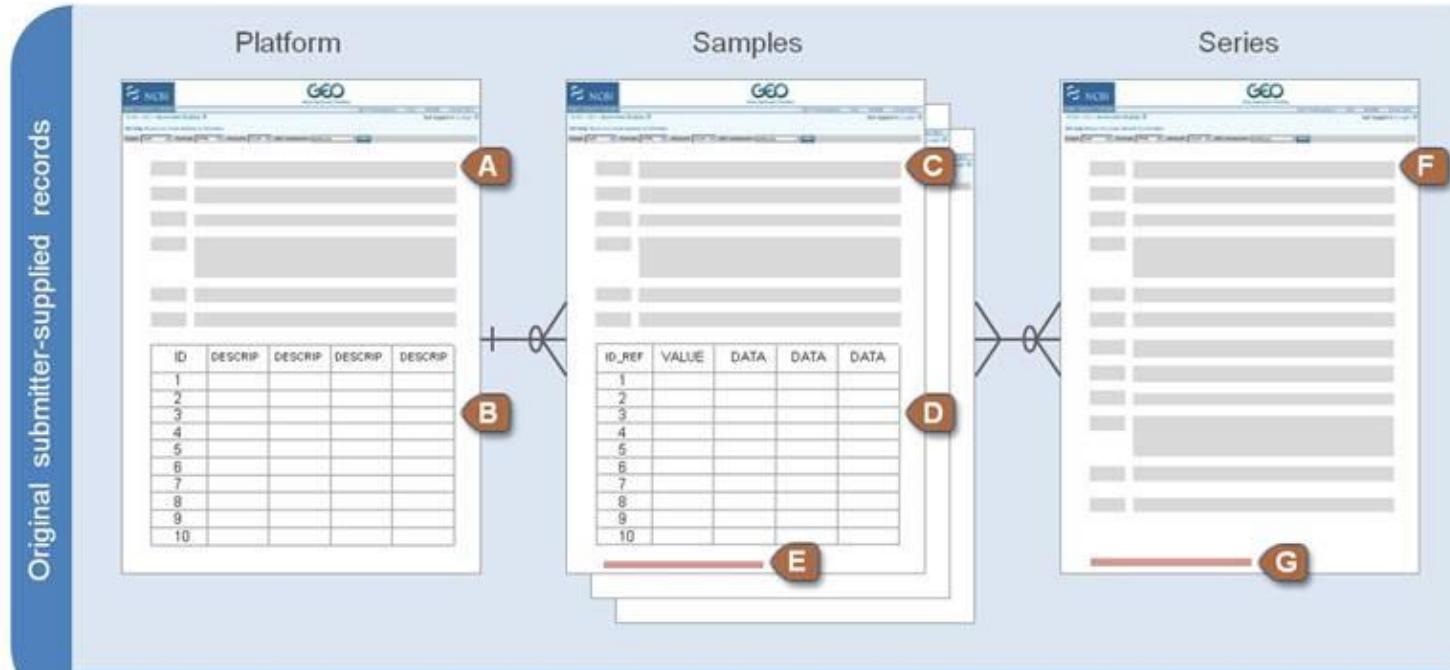
[Series Matrix File\(s\)](#)

TXT [?](#)

Supplementary file	Size	Download	File type/resource
GSE3541_RAW.tar	14.1 Mb	<a href="#">(http)(custom)</a>	TAR (of CEL, EXP)

Platforms (1)	<a href="#">GPL341</a> [RAE230A] Affymetrix Rat Expression 230A Array
Samples (6)	<a href="#">GSM81022</a> E19_epithelial cells_control1
<a href="#">Less...</a>	<a href="#">GSM81159</a> E19_epithelial cells_control2
	<a href="#">GSM81160</a> E19_epithelial cells_control3
	<a href="#">GSM81161</a> E19_epithelial cells_strain1
	<a href="#">GSM81162</a> E19_epithelial cells_strain2
	<a href="#">GSM81163</a> E19_epithelial cells_strain3

# Overview



# Index

1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis



## COVID-19 Information



[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)



# GEO DataSets

This database stores curated gene expression DataSets, as well as original Series and Platform records in the Gene Expression Omnibus (GEO) repository. Enter search terms to locate experiments of interest. DataSet records contain additional resources including cluster tools and differential expression queries.

## Getting Started

[GEO Documentation](#)

[GEO FAQ](#)

[About GEO DataSets](#)

[Construct a Query](#)

[Download Options](#)

## GEO Tools

[Submit to GEO](#)

[Advanced Search](#)

[DataSet Browser](#)

[Programmatic Access](#)

[GEO2R](#)

## More Resources

[GEO Home](#)

[GEO Profiles](#)

[SRA](#)

## Example Searches

Keywords and species

(smok\* OR diet) AND (mammals[organism] NOT human[organism])

Study type

"expression profiling by high throughput sequencing"[DataSet Type]

Studies with CEL files

cel[Supplementary Files]

DataSets that have 'age' as an experimental variable

age[Subset Variable Type]

Studies with between 100 and 500 samples

100:500[Number of Samples]



## COVID-19 Information

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

## Entry type

DataSets (59)

Series (2,296)

Samples (52,921)

Platforms (18)

## Organism

Customize ...

## Study type

Expression profiling by array

Methylation profiling by array

Customize ...

## Author

Customize ...

## Attribute name

tissue (21,004)

strain (9,848)

Customize ...

## Publication dates

30 days

1 year

Custom range...

[Clear all](#)[Show additional filters](#)

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾

Filters: [Manage Filters](#)

## Search results

Items: 1 to 20 of 55294

&lt;&lt; First &lt; Prev Page 1 of 2765 Next &gt; Last &gt;&gt;

- [RNA seq data of CD8+ T cells in peripheral blood lymphocytes from melanoma and lung cancer patients](#)

1. (Submitter supplied) CD8+ T cells are the primary target of immune checkpoint inhibitor (ICI) therapy in the treatment of cancers. ICI therapy only benefits a subset of patients and complicating this issue is a reliable prediction method that does not require invasive biopsies. We built an ICI response prediction model using scRNA seq data, and for validation of the model here we profiled transcriptome of CD8+ T cells in peripheral blood lymphocytes of 4 melanoma and 4 **lung cancer** patients that responded to ICI or not using RNA sequencing.

Organism: Homo sapiens  
Type: Expression profiling by high throughput sequencing  
Platform: GPL22790 8 Samples  
[Download data: XLSX](#)

Series Accession: GSE152590 ID: 200152590  
[SRA Run Selector](#)

- [Gene expression profiles of SCLC cells treated with KIX domain-blockade](#)

2. (Submitter supplied) This comparison is to determine a global change in gene expression influenced by blocking the CREBBP/EP300 KIX domain-protein interactions with MYB-MLL fusion peptides

Organism: Homo sapiens  
Type: Expression profiling by high throughput sequencing  
Platform: GPL30173 11 Samples  
[Download data: TXT](#)

Series Accession: GSE188705 ID: 200188705

- [The histologic phenotype of lung cancers is associated with transcriptomic features rather than](#)

## Top Organisms [Tree]

Homo sapiens (42684)

Mus musculus (12061)

Drosophila melanogaster (48)

Rattus norvegicus (45)

synthetic construct (25)

[More...](#)

## Find related data

Database: [Select](#) ▾[Find items](#)

## Search details

"lung neoplasms" [MeSH Terms] OR  
lung cancer [All Fields]

[Search](#)[See more...](#)

## Important Links

[GEO Home](#)[GEO Documentation](#)

<https://www.ncbi.nlm.nih.gov/gds/?term=lung+cancer>

**Search details**

("lung neoplasms"[MeSH Terms] OR lung cancer[All Fields]) AND "Homo sapiens"[porgn] AND ("gse"[Filter] AND "Expression profiling by array"[Filter])



## COVID-19 Information



[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

Entry type

clear

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾

Filters: [Manage Filters](#)

DataSets (0)

Samples (0)

Platforms (0)

Organism

Customize ...

Study type

clear

✓ Expression profiling by array

Methylation profiling by array

Customize ...

Author

Customize ...

Attribute name

tissue (245)

strain (6)

Customize ...

Publication dates

30 days

1 year

Custom range...

[Clear all](#)[Show additional filters](#)

## Search results

Items: 1 to 20 of 849

<< First < Prev Page  of 43 Next > Last >>

Filters activated: Series, Expression profiling by array. [Clear all](#) to show 42684 items.

- [The histologic phenotype of lung cancers is associated with transcriptomic features rather than genomic characteristics](#)

(Submitter supplied) Histology plays an essential role in therapeutic decision-making for **lung cancer** patients. However, the molecular determinants of **lung cancer** histology are largely unknown. We conducted whole-exome sequencing (WES) and microarray profiling on 19 micro-dissected tumor regions of different histologic subtypes from 9 patients with lung cancers of mixed histology. A median of 68.9% of point mutations and 83% of copy number aberrations were shared between different histologic components within the same tumors.  
[more...](#)

Organism: **Homo sapiens**  
 Type: **Expression profiling by array**  
 Platform: **GPL23159** 19 Samples  
 Download data: CEL, CHP  
 Series Accession: GSE188665 ID: 200188665  
[Analyze with GEO2R](#)

- [Microarray analysis of drug-tolerant cells in ALK-rearranged non-small cell lung cancer cell lines induced by ALK-TKIs](#)

(Submitter supplied) To investigate the molecular mechanisms underlying the emergence and maintenance of drug-tolerant (DT) cells in anaplastic lymphoma kinase (ALK)-rearranged **lung cancer**, we isolated DT cells from H2228 and A925L cells exposed to high doses of alectinib or lorlatinib for 9 days, and performed transcriptome analysis between parental and DT cells using a gene expression microarray.

## Top Organisms [Tree]

**Homo sapiens** (849)  
**Mus musculus** (19)  
**Rattus norvegicus** (2)  
**Human gammaherpesvirus 8** (1)  
**Murid gammaherpesvirus 4** (1)  
[More...](#)

## Find related data

Database: [Select](#)[Find items](#)

## Search details

```
("lung neoplasms"[MeSH Terms] OR
lung cancer[All Fields]) AND "Homo
sapiens"[porgn] AND ("gse"[Filter]
AND "Expression profiling by array"
[Filter])
```

[Search](#)[See more...](#)

## Important Links

**✓ Expression profiling by array**

- Methylation profiling by array
- Customize ...
- Author**
- Customize ...
- Attribute name**
- tissue (245)
- strain (6)
- Customize ...
- Publication dates**
- 30 days
- 1 year
- Custom range...
- [Clear all](#)
- [Show additional filters](#)

[Microarray analysis of induced by ALK TKIs](#)

**Additional filters**

- Entry type
- Organism
- Study type
- Author
- Attribute name
- Subset variable type
- Publication dates
- Sample count
- Supplementary file
- Search fields

[Show](#)

**1. The histologic primary genomic characteristic**

(Submitter supplied) Hist However, the molecular exome sequencing (WES) subtypes from 9 patients of copy number aberrations more... Organism: **Homo sapiens** Type: **Expression** Platform: GPL23159 19 S Download data: CEL, CFF Series Accession: GSE1 [Analyze with GEO2R](#)

30 days  
1 year  
Custom range...

**Custom range**

50 days  
1 year  
Custom range...

**Sample count**

Custom range...

Example: 10 to 100

[Apply](#) [Clear](#)

[Clear all](#)

[Show additional filters](#)

Organism: **Homo sapiens**  
Type: **Expression profiling by array**

**Custom range**

50 to 1000000

Example: 10 to 100

[Apply](#) [Clear](#)

Organism: **Homo sapiens**

61

GEO DataSets

GEO DataSets

(lung cancer) AND "Homo sapiens"[porgn:\_txid9606]

Search

Create alert Advanced

Help



## COVID-19 Information


[Public health information \(CDC\)](#) | 
[Research information \(NIH\)](#) | 
[SARS-CoV-2 data \(NCBI\)](#) | 
[Prevention and treatment information \(HHS\)](#) | 
[Español](#)

Entry type

Summary ▾ 20 per page ▾ Sort by Default order ▾

Send to: ▾

Filters: [Manage Filters](#)

DataSets (0)

✓ Series (217)

Samples (0)

Platforms (0)

Organism

Customize ...

Study type

clear

✓ Expression profiling by array

Methylation profiling by array

Customize ...

Author

Customize ...

Attribute name

tissue (100)

strain (2)

Customize ...

Publication dates

30 days

1 year

Custom range...

Sample count

clear

✓ From 50 to 1000000

[Clear all](#)[Show additional filters](#)

## Search results

Items: 1 to 20 of 217

<< First < Prev Page  of 11 Next > Last >>

Filters activated: Series, Expression profiling by array, Sample count from 50 to 1000000. [Clear all](#) to show 42684 items.

- [Transcriptomic analysis of PD-L1 positive and negative hot lung adenocarcinomas identifies a different molecular immunogenic landscape](#)

(Submitter supplied) Lung cancer is a major global health problem, as it is the leading cause of cancer-related deaths worldwide. Non-small-cell lung cancer (NSCLC), the most common form, is a heterogeneous disease with adenocarcinoma and squamous cell carcinoma being the predominant subtypes. Immune-inhibiting interaction of Programmed cell death-ligand 1 (PD-L1) with programmed cell death-protein 1 (PD-1) causes checkpoint mediated immune evasion and is, accordingly, an important therapeutic target in cancer.

[more...](#)

Organism: **Homo sapiens**  
Type: **Expression profiling by array**  
Platform: **GPL29738** 144 Samples  
Download data: RCC  
Series Accession: GSE180347 ID: 200180347  
[Analyze with GEO2R](#)

- [Response of SCLC to mutant KRAS or EGFR induction](#)

2. (Submitter supplied) Lineage transformation between lung cancer subtypes is a poorly understood phenomenon associated with resistance to treatment and poor patient outcomes. Here, we aimed to model this transition in order to define underlying biological mechanisms and identify potential avenues for therapeutic intervention. Small cell lung cancer (SCLC) is neuroendocrine in origin and, in contrast to non-SCLC (NSCLC), rarely contains mutations that drive the MAPK pathway. [more...](#)

Organism: **Homo sapiens**

## Top Organisms [Tree]

Homo sapiens (217)

Mus musculus (3)

Rattus norvegicus (1)

## Find related data

Database: [Select](#)[Find items](#)

## Search details

```
("lung neoplasms"[MeSH Terms] OR
lung cancer[All Fields]) AND "Homo
sapiens"[porgn] AND ("gse"[Filter]
AND "Expression profiling by array"
[Filter] AND ("50"[n_samples] :
```

[Search](#)[See more...](#)

## Important Links

[GEO Home](#)[GEO Documentation](#)[About GEO DataSets](#)[Construct a Query](#)

# Choose the one that best fits your question

- No treatment
- No control
- Several stages

<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE161533>

Status	Public on Nov 17, 2020
Title	Expression data from esophageal squamous cell carcinoma patients
Organism	<a href="#">Homo sapiens</a>
Experiment type	Expression profiling by array
Summary	<p>we conducted microarray experiments of 28 stage I-III ESCC patients based on Affymetrix Gene Chip Human Genome U133 plus 2.0 Array, performed enrichment analysis of differentially expressed genes (DEGs) as well as gene set enrichment analysis of all valid genes. Moreover, we summarized the secreted protein-encoding DEGs as well as esophagus-specific DEGs, hoping to offer some hints for early diagnosis and target for more efficacious treatment for ESCC in near future.</p>
Overall design	In total, there were 84 paired normal tissues, paratumor tissues, and tumor tissues from 28 ESCC patients were chosen to perform microarray analysis.
Contributor(s)	<a href="#">Qiu H, Li R, Li P, Xing W</a>
Citation missing	<i>Has this study been published? Please <a href="#">login</a> to update or <a href="#">notify GEO</a>.</i>
Submission date	Nov 16, 2020
Last update date	Nov 19, 2020
Contact name	Renling Li
E-mail(s)	<a href="mailto:1951668569@qq.com">1951668569@qq.com</a>
Phone	86-15811371935
Organization name	Shenzhen Technology University
Street address	3002 Lantian Road, Pingshan District, Shenzhen Guangdong, China
City	Shenzhen
ZIP/Postal code	518118
Country	China
Platforms (1)	<a href="#">GPL570</a> [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
Samples (84)	<a href="#">GSM4909553</a> normal tissue from patient 012
<a href="#">More...</a>	<a href="#">GSM4909554</a> normal tissue from patient 013
	<a href="#">GSM4909555</a> normal tissue from patient 014
<b>Relations</b>	
BioProject	<a href="#">PRJNA678656</a>

[Analyze with GEO2R](#)

# Index

1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. **GEO2R**
7. Enrichment analysis



COVID-19 is an emerging, rapidly evolving situation.

Get the latest public health information from CDC: <https://www.coronavirus.gov>.

Get the latest research from NIH: <https://www.nih.gov/coronavirus>.

Find NCBI SARS-CoV-2 literature, sequence, and clinical content: <https://www.ncbi.nlm.nih.gov/sars-cov-2/>.

Use GEO2R to compare two or more groups of Samples in order to identify genes that are differentially expressed across experimental conditions. Results are presented as a table of genes ordered by significance. [Full instructions](#) [YouTube](#)

GEO accession  Set Expression data from esophageal squamous cell carcinoma patients

Samples		Define groups										Selected 0 out of 84 samples		
Group	Accession	Title	Source name	Tissue	Stage	Age	Gender	Smoking history	Drinking history	Disease history	Columns	Set		
-	GSM4909553	normal tissue from patient 012	ESCC patient	normal tissue	IB	56	Male	None	None	None				
-	GSM4909554	normal tissue from patient 013	ESCC patient	normal tissue	IB	57	Male	30 years	Seldom	None				
-	GSM4909555	normal tissue from patient 014	ESCC patient	normal tissue	IB	51	Male	20 years	None	None				
-	GSM4909556	normal tissue from patient 015	ESCC patient	normal tissue	IB	64	Female	None	None	Hypertension				
-	GSM4909557	normal tissue from patient 016	ESCC patient	normal tissue	IB	54	Male	36 years	36 years	None				
-	GSM4909558	normal tissue from patient 111	ESCC patient	normal tissue	I	64	Female	None	None	None				
-	GSM4909559	normal tissue from patient 112	ESCC patient	normal tissue	IA	73	Female	None	None	Hypertension				
-	GSM4909560	normal tissue from patient 113	ESCC patient	normal tissue	I	73	Male	30 years	Seldom	None				
-	GSM4909561	normal tissue from patient 115	ESCC patient	normal tissue	I	61	Male	None	None	Hypertension				
-	GSM4909562	normal tissue from patient 021	ESCC patient	normal tissue	IIA	71	Male	50 years	Seldom	None				
		normal tissue from patient 023	ESCC patient	normal tissue	IIA	65	Male	30 years	None	None				

GEO accession

GSE161533

Set

Expression data from esophageal squamous cell carcinoma patients

Samples		Define groups									
Group	Accession	Title	Sample name	Tissue	Stage	Age	Gender	Smoking history	Drinking history	Disease history	
normal	GSM4909553	normal tissue	patient	normal tissue	IB	56	Male	None	None	None	
normal	GSM4909554	normal tissue from patient 013	ESCC patient	normal tissue	IB	57	Male	30 years	Seldom	None	
normal	GSM4909555	normal tissue from patient 014	ESCC patient	normal tissue	IB	51	Male	20 years	None	None	
normal	GSM4909556	normal tissue from patient 015	ESCC patient	normal tissue	IB	64	Female	None	None	Hypertension	
normal	GSM4909557	normal tissue from patient 016	ESCC patient	normal tissue	IB	54	Male	36 years	36 years	None	
normal	GSM4909558	normal tissue from patient 111	ESCC patient	normal tissue	I	64	Female	None	None	None	
normal	GSM4909559	normal tissue from patient 112	ESCC patient	normal tissue	IA	73	Female	None	None	Hypertension	
normal	GSM4909560	normal tissue from patient 113	ESCC patient	normal tissue	I	73	Male	30 years	Seldom	None	
normal	GSM4909561	normal tissue from patient 115	ESCC patient	normal tissue	I	61	Male	None	None	Hypertension	
normal	GSM4909562	normal tissue from patient 021	ESCC patient	normal tissue	IIA	71	Male	50 years	Seldom	None	
normal	GSM4909563	normal tissue from patient 023	ESCC patient	normal tissue	IIA	65	Male	30 years	None	None	
normal	GSM4909564	normal tissue from patient 024	ESCC patient	normal tissue	IIA	60	Female	None	None	None	
normal	GSM4909565	normal tissue from patient 025	ESCC patient	normal tissue	IIA	64	Female	None	None	Breast cancer	
normal	GSM4909566	normal tissue from patient 026	ESCC patient	normal tissue	IIB	69	Male	None	40 years	Hypertension	
normal	GSM4909567	normal tissue from patient 027	ESCC patient	normal tissue	IIB	65	Male	20 years	None	None	

Use GEO2R to compare two or more groups of Samples in order to identify genes that are differentially expressed across experimental conditions. Results are presented as a table of genes ordered by significance. [Full instructions](#) [YouTube](#)

GEO accession

GSE161533

Set

Expression data from esophageal squamous cell carcinoma patients

Samples		Define groups							Selected 56 out of 84 samples	
		Enter a group name: List							Columns	Set
tumor	GSM4909615	tumor tissue	patient	tumor tissue	I A	73	Female	None	None	Hypertension
tumor	GSM4909616	tumor tissue	patient	tumor tissue	I	73	Male	30 years	Seldom	None
tumor	GSM4909617	tumor tissue	patient	tumor tissue	I	61	Male	None	None	Hypertension
tumor	GSM4909618	tumor tissue from patient 021	ESCC patient	tumor tissue	II A	71	Male	50 years	Seldom	None
tumor	GSM4909619	tumor tissue from patient 023	ESCC patient	tumor tissue	II A	65	Male	30 years	None	None
tumor	GSM4909620	tumor tissue from patient 024	ESCC patient	tumor tissue	II A	60	Female	None	None	None
tumor	GSM4909621	tumor tissue from patient 025	ESCC patient	tumor tissue	II A	64	Female	None	None	Breast cancer
tumor	GSM4909622	tumor tissue from patient 026	ESCC patient	tumor tissue	II B	69	Male	None	40 years	Hypertension
tumor	GSM4909623	tumor tissue from patient 027	ESCC patient	tumor tissue	II B	65	Male	20 years	None	None
tumor	GSM4909624	tumor tissue from patient 121	ESCC patient	tumor tissue	II	63	Male	None	None	None
tumor	GSM4909625	tumor tissue from patient 122	ESCC patient	tumor tissue	II	56	Male	40 years	None	None
tumor	GSM4909626	tumor tissue from patient 124	ESCC patient	tumor tissue	II	64	Female	None	None	None
tumor	GSM4909627	tumor tissue from patient 125	ESCC patient	tumor tissue	II	64	Male	None	None	None
tumor	GSM4909628	tumor tissue from patient 126	ESCC patient	tumor tissue	II	57	Male	None	None	None
tumor	GSM4909629	tumor tissue from patient 031	ESCC patient	tumor tissue	III A	67	Male	None	Seldom	None
tumor	GSM4909630	tumor tissue from patient 032	ESCC patient	tumor tissue	III B	70	Male	40 years	40 years	None

GEO2R

Options

Profile graph

R script

The screenshot shows a software interface titled "GEO2R". At the top, there is a navigation bar with four tabs: "GEO2R" (highlighted in blue), "Options", "Profile graph", and "R script". Below the navigation bar, there is a section titled "▼ Quick start" which contains a bulleted list of instructions. After the list, there is a "How to use" section with an "Analyze" button.

▼ Quick start

- Specify a GEO Series accession and a Platform if prompted.
- Click 'Define groups' and enter names for the groups of Samples
- Assign Samples to each group. Highlight Sample rows then click Samples belong to which group.
- Click 'Analyze' to perform the calculation with default settings.
- You may change settings in the Options tab.

How to use

Analyze

← → C ⌘ 🔒 ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE161533

Selected 56 out of 84 samples

Samples Define groups

GEO2R Options Profile graph R script

Quick start

- Specify a GEO Series accession and a Platform if prompted.
- Click 'Define groups' and enter names for the groups of Samples you plan to compare, e.g., test and control.
- Assign Samples to each group. Highlight Sample rows then click the group name to assign those Samples to the group. Use the Sample metadata (title, source and characteristics) columns to help determine which Samples belong to which group.
- Click 'Analyze' to perform the calculation with default settings.
- You may change settings in the Options tab.

How to use

Processing...

---

| NLM | NIH | Email GEO | Disclaimer | Accessibility

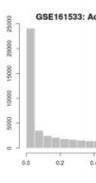
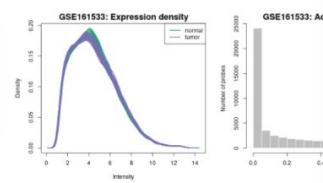
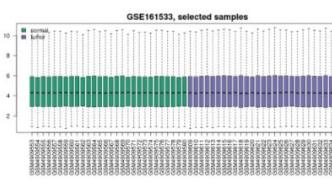
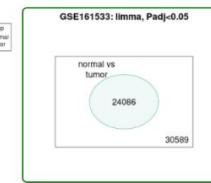
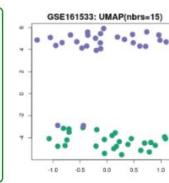
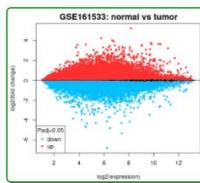
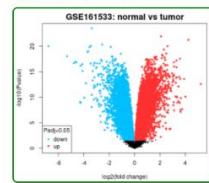
**GEO2R**

Options

Profile graph

R script

Log-transformation has been applied to the data. You can change this in the Options tab.

Reanalyze if you changed any options.
**Visualization** **Top differentially expressed genes**
[Download full table](#) [Select columns](#)

ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol	Gene.title
229802_at	1.67e-19	3.06e-24	-17.2	44.6	-3.379	WISP1	WNT1 inducible signaling ...
225548_at	3.33e-18	1.22e-22	15.9	41	2.085	SHROOM3	shroom family member 3
244056_at	1.05e-17	5.74e-22	15.4	39.5	4.239	SFTA2	surfactant associated 2
205479_s_at	6.24e-17	4.56e-21	-14.7	37.5	-2.547	PLAU	plasminogen activator, uro... WD repeat domain 72
236741_at	8.03e-17	9.58e-21	-14.5	36.8	-3.448	WDR72	WD repeat domain 72
228568_at	8.03e-17	1.01e-20	14.5	36.8	2.594	MYZAP	myocardial zonula adhere... matrix metallopeptidase 1
204475_at	8.03e-17	1.03e-20	-14.5	36.7	-6.779	MMP1	BH3 interacting domain de... collagen triple helix repeat
211725_s_at	9.58e-17	1.40e-20	-14.4	36.4	-1.882	BID	collagen triple helix repeat
225681_at	9.71e-17	1.60e-20	-14.3	36.3	-3.552	CTHRC1	inhibin beta A subunit
228400_at	2.46e-16	4.49e-20	14	35.3	2.311	SHROOM3	WD repeat domain 72
227140_at	4.87e-16	9.79e-20	-13.8	34.6	-3.663	INHBA	adipocyte plasma membra...
227174_at	6.11e-16	1.34e-19	-13.7	34.2	-3.694	WDR72	adipocyte plasma membra...
225354_s_at	1.18e-15	2.79e-19	13.4	33.5	2.496	SH3BGR1L2	SH3 domain binding gluta...
			-13.4	33.4	-1.083	APMAP	

[https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE161533#top\\_gen...](https://www.ncbi.nlm.nih.gov/geo/geo2r/?acc=GSE161533#top_gen...)

-13.4

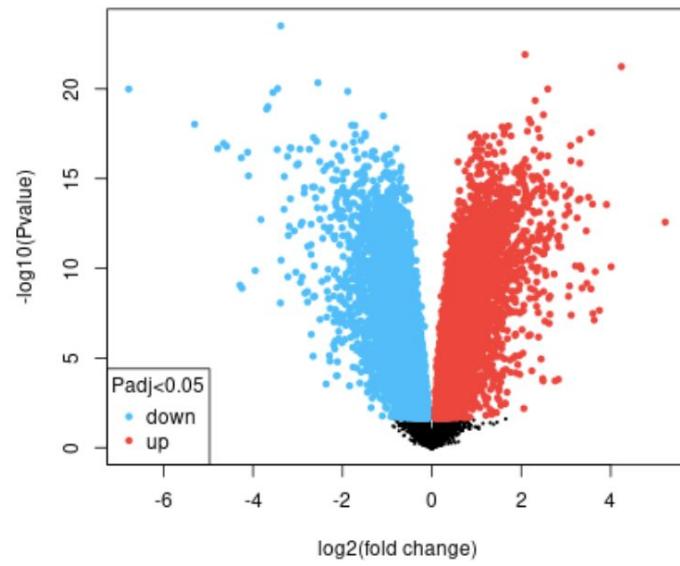
33.4

-1.083

### Volcano plot

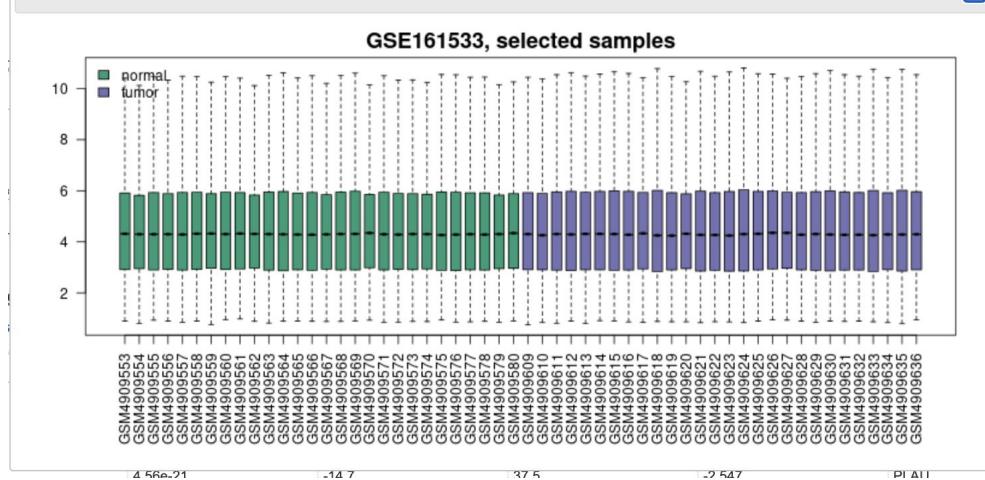
Explore and download

GSE161533: normal vs tumor



### Boxplot

GSE161533, selected samples



AutoSave OFF

GSE161533.top.table

**Home Insert Draw Page Layout Formulas Data Review View Tell me**

Paste **Calibri (Body) 12 A<sup>^</sup> A<sup>v</sup>** General Conditional Formatting Insert  
B I U **Format as Table** Delete Cell Styles

**Share Comments**

**Possible Data Loss** Some features might be lost if you save this workbook in the comma-delimited (.csv) format. To preserve these features, save it in an Excel file format.

**Save As...**

A1 ID adj.P.Val P.Value t B logFC Gene.symbol Gene.title

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	ID	adj.P.Val	P.Value	t	B	logFC	Gene.symbol	Gene.title										
2	229802_at	1.67E-19	3.06E-24	-17.242471	44.568682	-3.3791068	WISP1	WNT1 inducible signaling pathway protein 1										
3	225548_at	3.33E-18	1.22E-22	15.9488254	41.030561	2.0846304	SHROOM3	shroom family member 3										
4	244056_at	1.05E-17	5.74E-22	15.4240445	39.535625	4.23856836	SFTA2	surfactant associated 2										
5	205479_s_at	6.24E-17	4.56E-21	-14.73904	37.530673	-2.5472552	PLAU	plasminogen activator, urokinase										
6	236741_at	8.03E-17	9.58E-21	-14.498451	36.811845	-3.4480073	WDR72	WD repeat domain 72										
7	228568_at	8.03E-17	1.01E-20	14.4808231	36.758874	2.59372398	MYZAP	myocardial zona adherens protein										
8	204475_at	8.03E-17	1.03E-20	-14.475822	36.743838	-6.7785099	MMP1	matrix metallopeptidase 1										
9	211725_s_at	9.58E-17	1.40E-20	-14.37605	36.443185	-1.8817473	BID	BH3 interacting domain death agonist										
10	225681_at	9.71E-17	1.60E-20	-14.334012	36.316109	-3.5519325	CTHRC1	collagen triple helix repeat containing 1										
11	228400_at	2.46E-16	4.49E-20	14.0047996	35.312753	2.3109943	SHROOM3	shroom family member 3										
12	227140_at	4.87E-16	9.79E-20	-13.75958	34.555908	-3.6633428	INHBA	inhibin beta A subunit										
13	227174_at	6.11E-16	1.34E-19	-13.661189	34.249953	-3.6943889	WDR72	WD repeat domain 72										
14	225354_s_at	1.18E-15	2.79E-19	13.4331293	33.535733	2.49586797	SH3BGRL2	SH3 domain binding glutamate rich protein like 2										
15	206656_s_at	1.25E-15	3.21E-19	-13.390707	33.402101	-1.0826746	APMAP	adipocyte plasma membrane associated protein										
16	228865_at	1.82E-15	5.00E-19	13.2538835	32.969429	2.17319879	C1orf116	chromosome 1 open reading frame 116										
17	219722_s_at	2.69E-15	7.88E-19	13.1144415	32.525859	2.21794702	GDPD3	glycerophosphodiester phosphodiesterase domain containing 3										
18	204580_at	3.04E-15	9.45E-19	-13.059048	32.348917	-5.3039917	MMP12	matrix metallopeptidase 12										
19	212013_at	3.14E-15	1.07E-18	-13.020645	32.226001	-1.7923972	PXDN	peroxidasin										
20	211981_at	3.14E-15	1.09E-18	-13.014806	32.207294	-1.7229894	COL4A1	collagen type IV alpha 1 chain										
21	219100_at	3.22E-15	1.18E-18	12.9918756	32.133788	1.71748717	OBFC1	oligonucleotide/oligosaccharide binding fold containing 1										
22	209679_s_at	3.38E-15	1.31E-18	12.9586067	32.027013	1.56752933	SMAGP	small cell adhesion glycoprotein										
23	207655_s_at	3.38E-15	1.36E-18	12.9478864	31.992575	1.66569828	BLNK	B-cell linker										
24	228653_at	4.11E-15	1.73E-18	12.8758347	31.760708	2.37342381	SAMD5	sterile alpha motif domain containing 5										
25	228221_at	5.24E-15	2.30E-18	12.7891395	31.480778	2.1378179	SLC44A3	solute carrier family 44 member 3										
26	234998_at	5.50E-15	2.51E-18	12.7627148	31.395252	1.62039122	RAB11A	RAB11A, member RAS oncogene family										
27	214621_at	5.80E-15	2.76E-18	12.7348478	31.304956	3.56756634	GYS2	glycogen synthase 2										
28	221538_s_at	6.20E-15	3.06E-18	-12.703142	31.202091	-1.4925317	PLXNA1	plexin A1										
29	217845_x_at	6.45E-15	3.33E-18	12.6777193	31.119514	0.96733583	HIGD1A	HIG1 hypoxia inducible domain family member 1A										

```
# Differential expression analysis with limma
library(GEOquery)
library(limma)
library(umap)

# load series and platform data from GEO

gset <- getGEO("GSE161533", GSEMatrix =TRUE, AnnotGPL=TRUE)
if (length(gset) > 1) idx <- grep("GPL570", attr(gset, "names")) else idx <- 1
gset <- gset[[idx]]

# make proper column names to match toptable
fvarLabels(gset) <- make.names(fvarLabels(gset))

# group membership for all samples
gsms <- paste0("00000000000000000000000000XXXXXX",
               "XXXXXX1111111111111111111111111111")
sml <- strsplit(gsms, split="")[[1]]

# filter out excluded samples (marked as "X")
sel <- which(sml != "X")
sml <- sml[sel]
gset <- gset[,sel]
```

```

# log2 transformation
ex <- exprs(gset)
qx <- as.numeric(quantile(ex, c(0., 0.25, 0.5, 0.75, 0.99, 1.0), na.rm=T))
LogC <- (qx[5] > 100) ||
  (qx[6]-qx[1] > 50 && qx[2] > 0)
if (LogC) { ex[which(ex <= 0)] <- NaN
  exprs(gset) <- log2(ex) }

# assign samples to groups and set up design matrix
gs <- factor(sml)
groups <- make.names(c("normal","tumor"))
levels(gs) <- groups
gset$group <- gs
design <- model.matrix(~group + 0, gset)
colnames(design) <- levels(gs)

fit <- lmFit(gset, design) # fit linear model

# set up contrasts of interest and recalculate model coefficients
cts <- paste(groups[1], groups[2], sep="-")
cont.matrix <- makeContrasts(contrasts=cts, levels=design)
fit2 <- contrasts.fit(fit, cont.matrix)

# compute statistics and table of top significant genes
fit2 <- eBayes(fit2, 0.01)
tT <- topTable(fit2, adjust="fdr", sort.by="B", number=250)

tT <- subset(tT, select=c("ID","adj.P.Val","P.Value","t","B","logFC","Gene.symbol","Gene.title"))
write.table(tT, file=stdout(), row.names=F, sep="\t")

```

# Differentially expressed gene

P-values < 0.001 & ( LogFC < -2 or LogFC > -2 )

MMP1, MMP3, CCL20, FCRL5, KRT14, MMP9, IGK, MMP10, SFRP4, MMP13, CXCL10, COL1A2, MMP12, VCAN, COL1A1, CXCL11, CXCL13, SHISA2, CYP24A1, MFAP2, COL5A2, KRT17, RBP1, CHI3L1, CXCL8, SHOX2, TYMP, ADAMDEC1, SPRR2G, SLC44A5, BCL2A1, CCL3L3, SERPINE1, TNFAIP6, LINC01296, CXCL1, LAMC2, TREM1, CXCL6, LOXL2, WDR66, CXCL5, WDR72, COL10A1, SPP1, MAGEA6, MAGEA11, DUXAP10, POSTN, IGFBP3, ADAM12, IL36G, INHBA, ISG15, PTHLH, APOC1, APOBEC3A\_B, HOXB7, APOBEC3B, SIX1, LOC100129518, PLAU, ABCA13, PLA2G7, WISP1, HOXD10, S100A7A, CST1, ADAMTS2, TDO2, ZIC2, COL11A1, CCL18, CTHRC1

# Index

1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

# Enrichment analysis

W90–W97 Nucleic Acids Research, 2016, Vol. 44, Web Server issue  
doi: 10.1093/nar/gkw377

Published online 3 May 2016

## Enrichr: a comprehensive gene set enrichment analysis web server 2016 update

Maxim V. Kuleshov<sup>1</sup>, Matthew R. Jones<sup>1</sup>, Andrew D. Rouillard<sup>1</sup>, Nicolas F. Fernandez<sup>1</sup>, Qiaonan Duan<sup>1</sup>, Zichen Wang<sup>1</sup>, Simon Koplev<sup>1</sup>, Sherry L. Jenkins<sup>1</sup>, Kathleen M. Jagodnik<sup>2</sup>, Alexander Lachmann<sup>1</sup>, Michael G. McDermott<sup>1</sup>, Caroline D. Monteiro<sup>1</sup>, Gregory W. Gundersen<sup>1</sup> and Avi Ma'ayan<sup>1,\*</sup>

<sup>1</sup>Department of Pharmacology and Systems Therapeutics, BD2K-LINCS Data Coordination and Integration Center, Icahn School of Medicine at Mount Sinai, One Gustave L. Levy Place Box 1215, New York, NY 10029, USA and

<sup>2</sup>Fluid Physics and Transport Processes Branch, NASA Glenn Research Center, 21000 Brookpark Rd., Cleveland, OH 44135, USA

Received January 29, 2016; Revised April 9, 2016; Accepted April 25, 2016

**PNAS**  
Proceedings of the  
National Academy of Sciences  
of the United States of America

Institution: Centre de  
Medecine

Keyword, Author, c

Home Articles Front Matter News Podcasts Authors

RESEARCH ARTICLE

Check for updates

### Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles

Aravind Subramanian, Pablo Tamayo, Vamsi K. Mootha, Sayan Mukherjee, Benjamin L. Ebert, ...  
+ See all authors and affiliations

PNAS October 25, 2005 102 (43) 15545–15550; first published September 30, 2005;  
<https://doi.org/10.1073/pnas.0506580102>

Contributed by Eric S. Lander, August 2, 2005

Article Figures & SI Info & Metrics PDF



enrichr



All

Images

Videos

Books

News

More

Tools

About 168,000 results (0.48 seconds)

<https://maayanlab.cloud> › Enrichr

⋮

## Enrichr - Ma'ayan Laboratory - Computational Systems Biology

Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC

Bioinformatics. 2013; 128(14). Kuleshov MV, Jones MR, Rouillard AD, ...

[Help](#) · [Login](#) · [FishEnrichr](#) · [FlyEnrichr](#)

### People also ask ⋮

What is Enrichr used for?



What is Enrichr analysis?



How Enrichr works?



What is an enriched pathway?



Feedback



## Input data

Expand a gene, a term, or a variant into a gene set:

e.g. STAT3, breast cancer, or rs28897756



Try an example [STAT3](#) [breast cancer](#) [rs28897756](#)

Include the top 100 most relevant genes



Paste a set of valid Entrez gene symbols on each row in the text-box below. [Try a gene set example.](#)

Paste a set of valid Entrez gene symbols (e.g. STAT3) on each row in the text-box

0 gene(s) entered

In order to enable others to search your set please enter a brief description of it.

Contribute your set so it can be searched by others

Submit

Please acknowledge Enrichr in your publications by citing the following references:

Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, Clark NR, Ma'ayan A.

Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*. 2013; 128(14).

Kuleshov MV, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, Koplev S, Jenkins SL, Jagodnik KM, Lachmann A, McDermott MG, Monteiro CD, Gundersen GW, Ma'ayan A.

Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*. 2016; gkw377.



Transcription Pathways Ontologies Diseases/Drugs Cell Types Misc Legacy Crowd

Description No description available (74 genes)



### ChEA 2016

RELA 24523406 ChIP-Seq FIBROSARCOMA H  
BACH1 22875853 ChIP-PCR HELA AND SCP4  
SMAD4 19686287 ChIP-ChIP HaCaT Human  
SUZ12 27294783 Chip-Seq ESCs Mouse  
EED 16625203 ChIP-ChIP MESCs Mouse

### ENCODE and ChEA Consensus TFs from

SUZ12 CHEA  
EZH2 CHEA  
SMAD4 CHEA  
NFE2L2 CHEA  
SOX2 CHEA

### ARCHS4 TFs Coexp

SNAI2 human tf ARCHS4 coexpression  
TWIST1 human tf ARCHS4 coexpression  
TBX5 human tf ARCHS4 coexpression  
FBNT human tf ARCHS4 coexpression  
MTF1 human tf ARCHS4 coexpression

### TF Perturbations Followed by Expression

NEUROG1 OE HUMAN GSE18296 CREEDSID  
EGR3 KD HUMAN GSE2108 CREEDSID GENE  
TP63 OE HUMAN GSE20286 CREEDSID GENE  
FOSL1 KO MOUSE GSE43695 CREEDSID GEN  
FOSL1 KO MOUSE GSE43695 CREEDSID GEN

### TRRUST Transcription Factors 2019

RELA human  
NFKB1 human  
JUN mouse  
ETS1 human  
SP1 human

### IncHUB lncRNA Co-Expression

VCAN-AS1  
LINC01614  
GORAB-AS1  
LINC02544  
LINC01929

### Enrichr Submissions TF-Gene Cooccurrence

ZBED2  
HOXC10  
BNC1  
FOXS1

### TRANSFAC and JASPAR PWMs

NFAT2 (human)  
SRF (human)  
MYB (human)  
TCF4 (human)

### Epigenomics Roadmap HM ChIP-seq

H3K27me3 Mobilized CD34 Primary Cells  
H3K27me3 CD34 Primary Cells  
H3K27me3 Brain Hippocampus Middle  
H3K27me3 Skeletal Muscle



il17 signaling pathway lung cancer



All

Images

News

Videos

More

Tools

About 678,000 results (0.57 seconds)

## Scholarly articles for **il17 signaling pathway lung cancer**

... /GIV **signaling pathway** in non-small-cell **lung cancer** - Pan - Cited by 112

... -17 promoted metastasis of non-small-cell **lung cancer** ... - Li - Cited by 86

... -Th1/Th17 **pathway**: an attractive target for **lung cancer** ... - Joerger - Cited by 24

Numerous studies have shown that IL-17 directly or indirectly promotes tumour angiogenesis and cell proliferation and that it inhibits apoptosis via the activation of inflammatory signalling pathways. Therefore, IL-17 contributes to the metastasis **and progression** of lung cancer.

<https://www.ncbi.nlm.nih.gov> › articles › PMC5107223

⋮

**The Role of Interleukin-17 in Lung Cancer - NCBI**



About featured snippets



Feedback

[nature](#) > [scientific reports](#) > [articles](#) > [article](#)

Open Access | Published: 03 November 2015

## Interleukin-17 promotes angiogenesis by stimulating VEGF production of cancer cells via the STAT3/GIV signalling pathway in non-small-cell lung cancer

Bo Pan, Jing Shen, Jingyan Cao, Yongxu Zhou, Lihua Shang, Shi Jin, Shoubo Cao, Dehai Che, Fang Liu  
& Yan Yu

[Scientific Reports](#) 5, Article number: 16053 (2015) | [Cite this article](#)

3846 Accesses | 77 Citations | [Metrics](#)

[Download PDF](#)[Sections](#)[Figures](#)[Abstract](#)[Introduction](#)[Results](#)[Discussion](#)[Materials and Methods](#)[Additional Information](#)

[Transcription](#)
[Pathways](#)
[Ontologies](#)
[Diseases/Drugs](#)
[Cell Types](#)
[Misc](#)
[Legacy](#)
[Crowd](#)
[Description](#)

No description available (74 genes)



### BioPlanet 2019



- [Interleukin-1 regulation of extracellular mat](#)
- [Oncostatin M](#)
- [Extracellular matrix organization](#)
- [TWEAK regulation of gene expression](#)
- [Binding of chemokines to chemokine recept](#)

### WikiPathway 2021 Human



- [IL-18 signaling pathway WP4754](#)
- [Matrix Metalloproteinases WP129](#)
- [SARS-CoV-2 innate immunity evasion and ce](#)
- [Senescence and Autophagy in Cancer WP61!](#)
- [Endochondral Ossification with Skeletal Dys](#)

### KEGG 2021 Human



- [IL-17 signaling pathway](#)
- [Viral protein interaction with cytokine and c](#)
- [Cytokine-cytokine receptor interaction](#)
- [Rheumatoid arthritis](#)
- [Chemokine signaling pathway](#)

### ARCS4 Kinases Coexp



- [IRAK2 human kinase ARCS4 coexpression](#)
- [DDR2 human kinase ARCS4 coexpression](#)
- [PDGFRα human kinase ARCS4 coexpressio](#)
- [NEK7 human kinase ARCS4 coexpression](#)
- [PDGFRβ human kinase ARCS4 coexpressio](#)

### Elsevier Pathway Collection



- [Proteins Involved in Atherosclerosis](#)
- [Proteins Involved in Osteoarthritis](#)
- [Proteins with Altered Expression in Cancer N](#)
- [Proteins Involved in Chronic Obstructive Pul](#)
- [IL17 Signaling in Psoriasis](#)

### MSigDB Hallmark 2020



- [Epithelial Mesenchymal Transition](#)
- [TNF-alpha Signaling via NF-κB](#)
- [KRAS Signaling Up](#)
- [Angiogenesis](#)
- [Inflammatory Response](#)

### BioCarta 2016



- [Platelet Amyloid Precursor Protein Pathway](#)
- [Fibrinolysis Pathway Homo sapiens h fibrin](#)
- [Inhibition of Matrix Metalloproteinases Hom](#)
- [Regulators of Bone Mineralization Homo saj](#)

### Reactome 2016



- [Extracellular matrix organization Homo sapi](#)
- [Assembly of collagen fibrils and other multi](#)
- [Collagen formation Homo sapiens R-HSA-14](#)
- [Chemokine receptors bind chemokines Hom](#)

### HumanCyc 2016



- [tryptophan degradation to 2-amino-3-carbo](#)
- [retinoate biosynthesis I Homo sapiens PWY-](#)
- [tryptophan degradation Homo sapiens TRYF](#)
- [NAD de novo biosynthesis Homo sapiens N/](#)

[Login](#) | [Register](#)[Transcription](#) [Pathways](#) [Ontologies](#) [Diseases/Drugs](#) [Cell Types](#) [Misc](#) [Legacy](#) [Crowd](#)**Description** No description available (74 genes)**GO Biological Process 2021**

extracellular structure organization (GO:004  
external encapsulating structure organizatio  
extracellular matrix organization (GO:00301  
chemokine-mediated signaling pathway (GO  
cellular response to chemokine (GO:199086

**GO Molecular Function 2021**

chemokine activity (GO:0008009)  
CXCR chemokine receptor binding (GO:0045  
chemokine receptor binding (GO:0042379)  
cytokine activity (GO:0005125)  
metallopeptidase activity (GO:0008237)

**GO Cellular Component 2021**

collagen-containing extracellular matrix (GO  
endoplasmic reticulum lumen (GO:0005788;  
intracellular organelle lumen (GO:0070013)  
tertiary granule lumen (GO:1904724)  
high-density lipoprotein particle (GO:003436

**MGI Mammalian Phenotype Level 4 2021**

abnormal trabecular bone morphology MP:  
decreased skin tensile strength MP:0003089  
short femur MP:0003109  
decreased bone strength MP:0004991  
decreased susceptibility to injury MP:000516

**Human Phenotype Ontology**

Atrophic scars (HP:0001075)  
Osteoarthritis (HP:0002758)  
Bruising susceptibility (HP:0000978)  
Soft skin (HP:0000977)  
Tibial bowing (HP:0002982)

**Jensen TISSUES**

Stromal cell  
Myofibroblast  
Immune system  
Synovial tissue  
Bone matrix

**Jensen COMPARTMENTS**

Thrombospondin complex  
Extracellular space  
VEGF-A complex  
Proteinaceous extracellular matrix

**Jensen DISEASES**

Arthritis  
Ehlers-Danlos syndrome  
Periodontal disease  
Lung disease

NCBI Resources How To Sign in to I

PMC Advanced Journal list Search

**COVID-19 Information**

[Public health information \(CDC\)](#) | [Research information \(NIH\)](#) | [SARS-CoV-2 data \(NCBI\)](#) | [Prevention and treatment information \(HHS\)](#) | [Español](#)

**Try out PMC Labs and tell us what you think. [Learn More.](#)**

Journal List &gt; Int J Mol Sci &gt; v.19(10); 2018 Oct &gt; PMC6213383

*Int J Mol Sci.* 2018 Oct; 19(10): 3028.Published online 2018 Oct 4. doi: [10.3390/ijms19103028](https://doi.org/10.3390/ijms19103028)

PMCID: PMC6213383

PMID: [30287763](#)

## Role of Extracellular Matrix in Development and Cancer Progression

Cameron Walker,<sup>1,†</sup> Elijah Mojares,<sup>1,†</sup> and Armando del Río Hernández<sup>1,\*</sup>
[Author information](#) [Article notes](#) [Copyright and License information](#) [Disclaimer](#)

This article has been cited by other articles in PMC.

### Abstract

Go to: 

The immense diversity of extracellular matrix (ECM) proteins confers distinct biochemical and biophysical properties that influence cell phenotype. The ECM is highly dynamic as it is constantly deposited, remodelled, and degraded during development until maturity to maintain tissue homeostasis. The ECM's composition and organization are spatiotemporally regulated to control cell behaviour and differentiation, but dysregulation of ECM dynamics leads to the development of diseases such as cancer. The chemical cues presented by the ECM have been appreciated as key drivers for both development and cancer

### Formats:

[Article](#) | [PubReader](#) | [ePub \(beta\)](#) | [PDF \(3.6M\)](#) | [Cite](#)

### Share

 [Facebook](#)  [Twitter](#)  [Google+](#)

### Save items

### Similar articles in PubMed

The matrix environmental and cell mechanical properties reg

cell migration and contribute to the invasive [Rep Prog Phys.

The Tumor Microenvironment: Focus on Extracellular Matrix.

[Adv Exp Med Biol.

Charting the unexplored extracellular matrix in cancer.

[Int J Exp Pathol.

Esophageal Cancer Development: Crucial Clues Arising from  
Extracellular Matrix. [Cells.Roles of Proteoglycans and Glycosaminoglycans in Cancer  
Development and Progression. [Int J Mol Sci.



stringdb



All



News



Videos



Tools

About 443,000 results (0.50 seconds)

<https://string-db.org> ::

## STRING: functional protein association networks

Welcome to STRING. Protein-Protein Interaction Networks Functional Enrichment Analysis.

Organisms 14094; Proteins 67.6 mio; Interactions

### About

STRING is a database of known and predicted protein-protein ...

### Access

736'429 proteins from 179 organisms; 12'014'052 ...

### Info

For most types of evidence, there are two types of scores: the ...

### Credits

Credits · Institutional Support(manpower, hardware, office ...

[More results from string-db.org »](#)

The screenshot shows the STRING website. The top navigation bar includes links for All, Images, News, Videos, More, and Tools. Below the search bar, it says "About 443,000 results (0.50 seconds)". The main content area features a large heading "STRING: functional protein association networks". Below this is a sub-headline "Welcome to STRING. Protein-Protein Interaction Networks Functional Enrichment Analysis." followed by statistics: "Organisms 14094; Proteins 67.6 mio; Interactions". To the right, there's a large network graph with nodes labeled with protein names like "TSHZ", "GAPDH", "HSP90AA1", etc., and four panels (A, B, C, D) showing different network components. Further down, there's a section titled "Access" with stats "736'429 proteins from 179 organisms; 12'014'052 ...", a "Info" section with "For most types of evidence, there are two types of scores: the ...", a "Credits" section with "Credits · Institutional Support(manpower, hardware, office ...)", and a link "More results from string-db.org »". At the bottom, there's a "STRING" logo and a share icon.

My Drive - Google Drive | Enrichr | Geo\_Workshop - Google Slides | STRING: functional protein ass

string-db.org Version: 11.5 LOGIN REGISTER

# STRING

Welcome to STRING

Protein-Protein Interaction Networks  
Functional Enrichment Analysis

ORGANISMS 14094 | PROTEINS 67.6 mio | INTERACTIONS >20 bln

SEARCH

88

[Search](#)[Download](#)[Help](#)[My Data](#)[Protein by name](#)[Protein by sequence](#)[Multiple proteins](#)[Multiple sequences](#)[Proteins with Values/Ranks](#)[Organisms](#)[Protein families \("COGs"\)](#)[Examples](#)[Random entry](#)

## SEARCH

### Multiple Proteins by Names / Identifiers

List Of Names:

(one per line; examples: #1 #2 #3)

... or, upload a file:

 [Browse ...](#)

Organism:

[Advanced Settings](#)[SEARCH](#)

[Search](#)[Download](#)[Help](#)[My Data](#)

## organism selection ...

Your input contains 74 lines.

Please specify the organism below, then click 'Continue' to proceed.

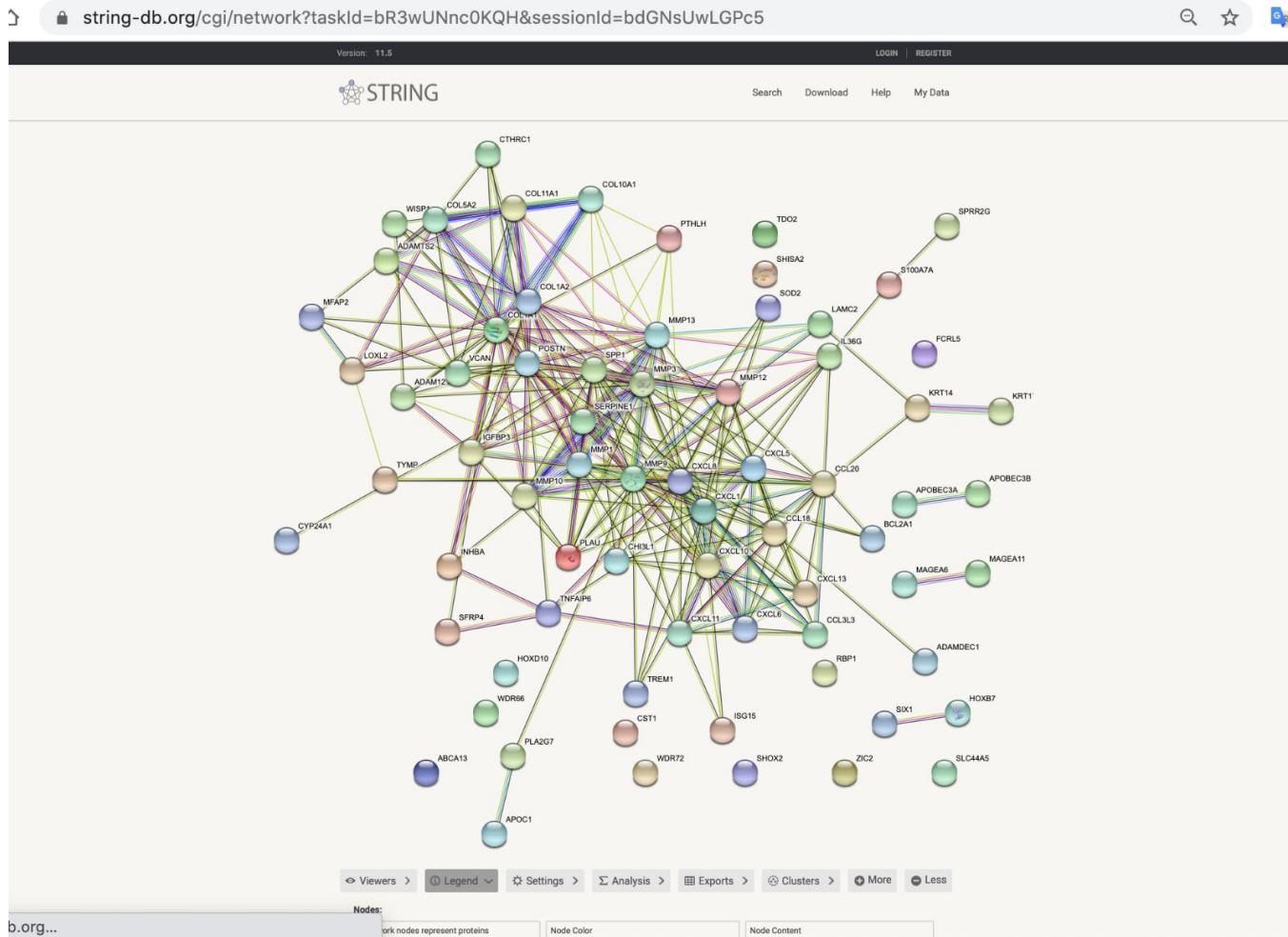
Homo sapiens

Homo sapiens

Homo sapiens

<- BACK

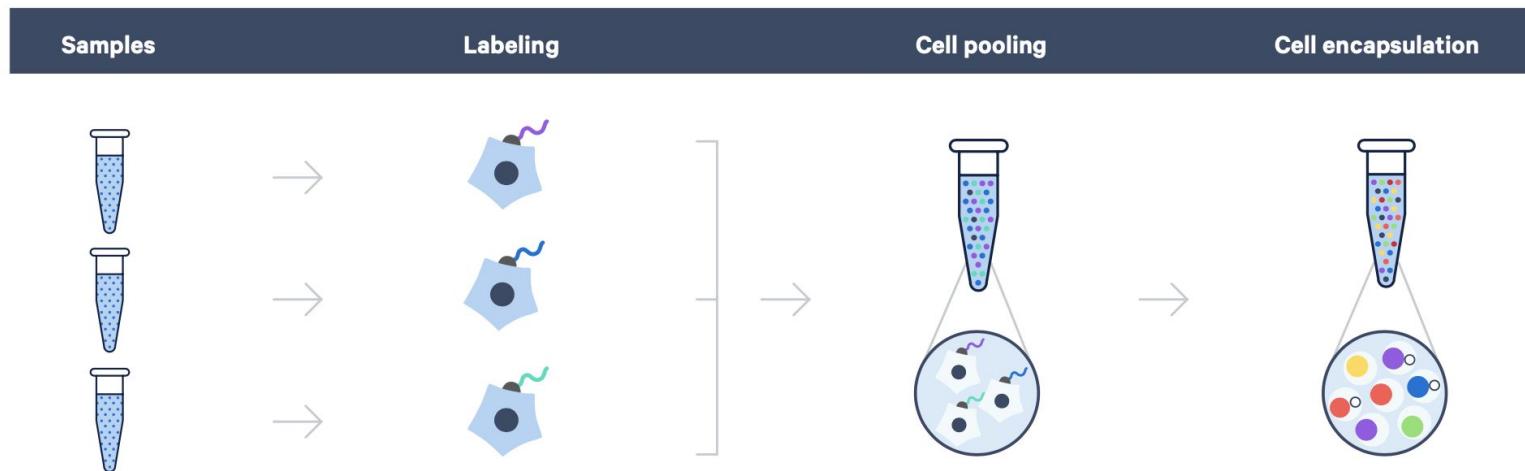
CONTINUE ->



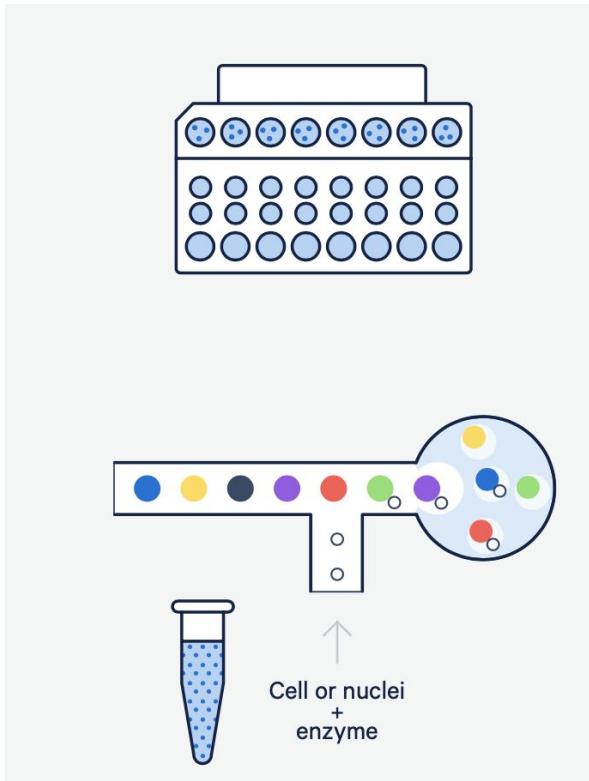
# Index

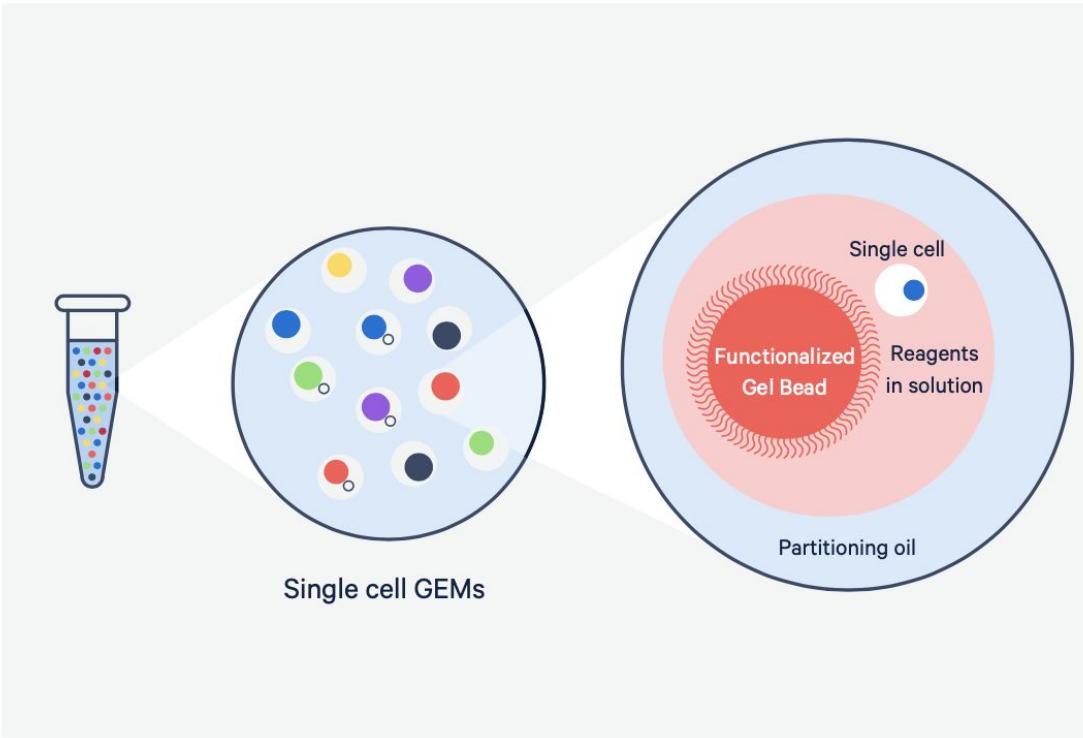
1. Basic biology
2. Microarray technology
3. RNA-seq technology
4. GEO website
5. Lung cancer dataset
6. GEO2R
7. Enrichment analysis

# Single cell sequencing

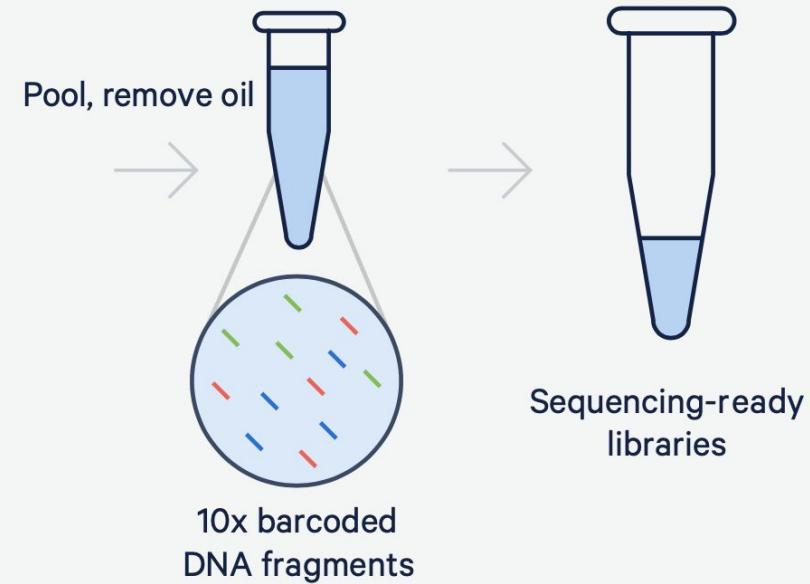
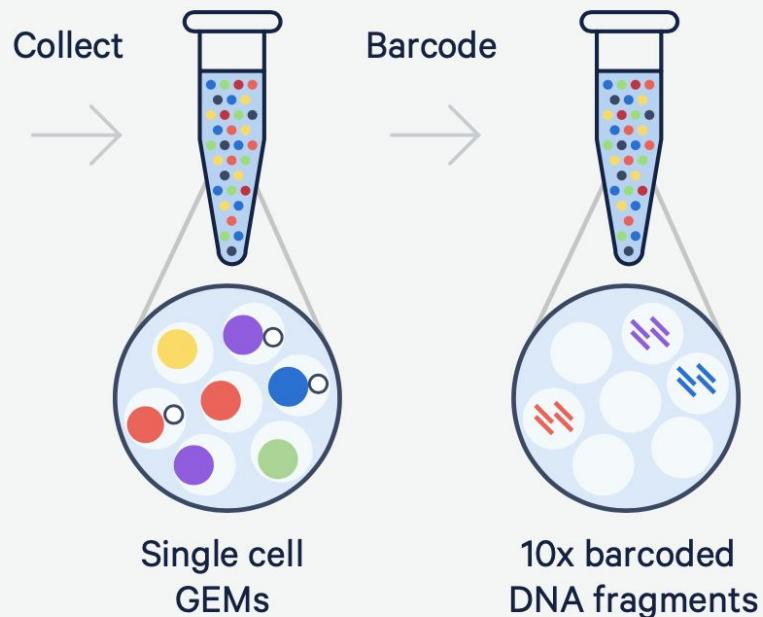


- Gel Beads, cells or nuclei, enzymes, and partitioning oil are loaded onto a chip.
- Barcoded gel beads are mixed with the cells or nuclei, enzymes, and partitioning oil.

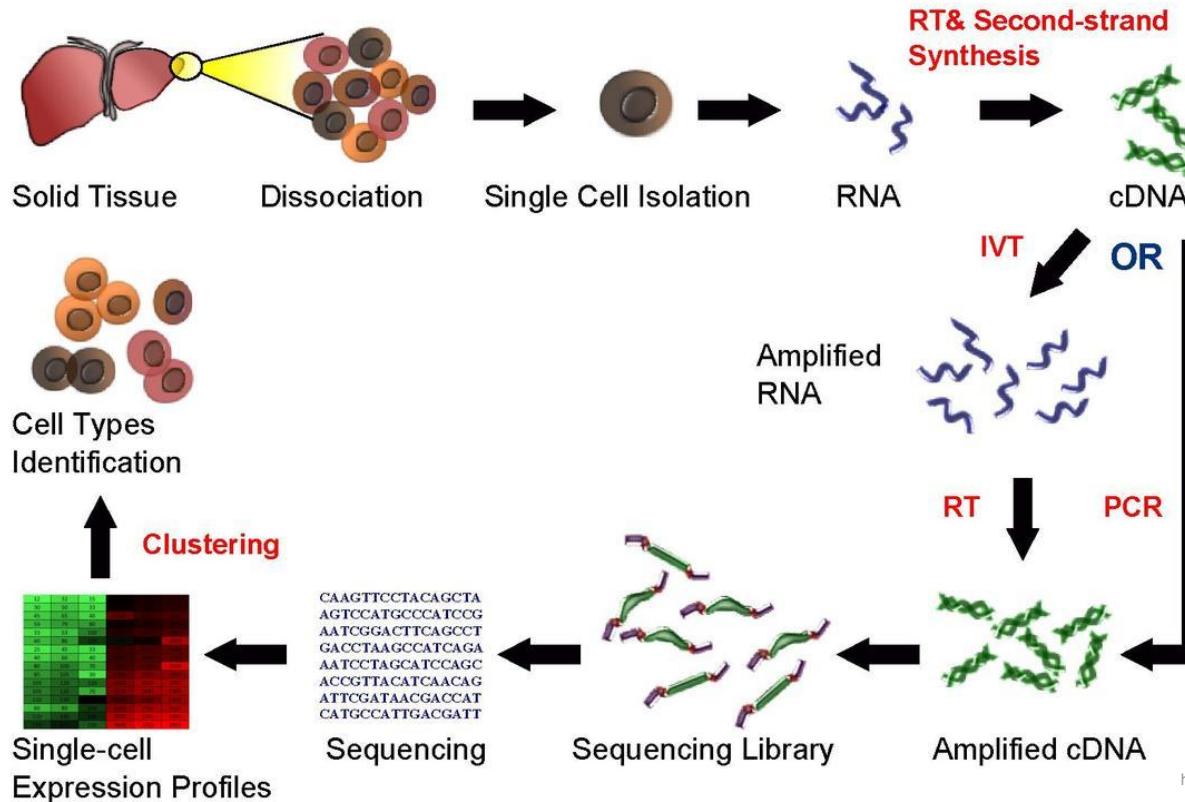


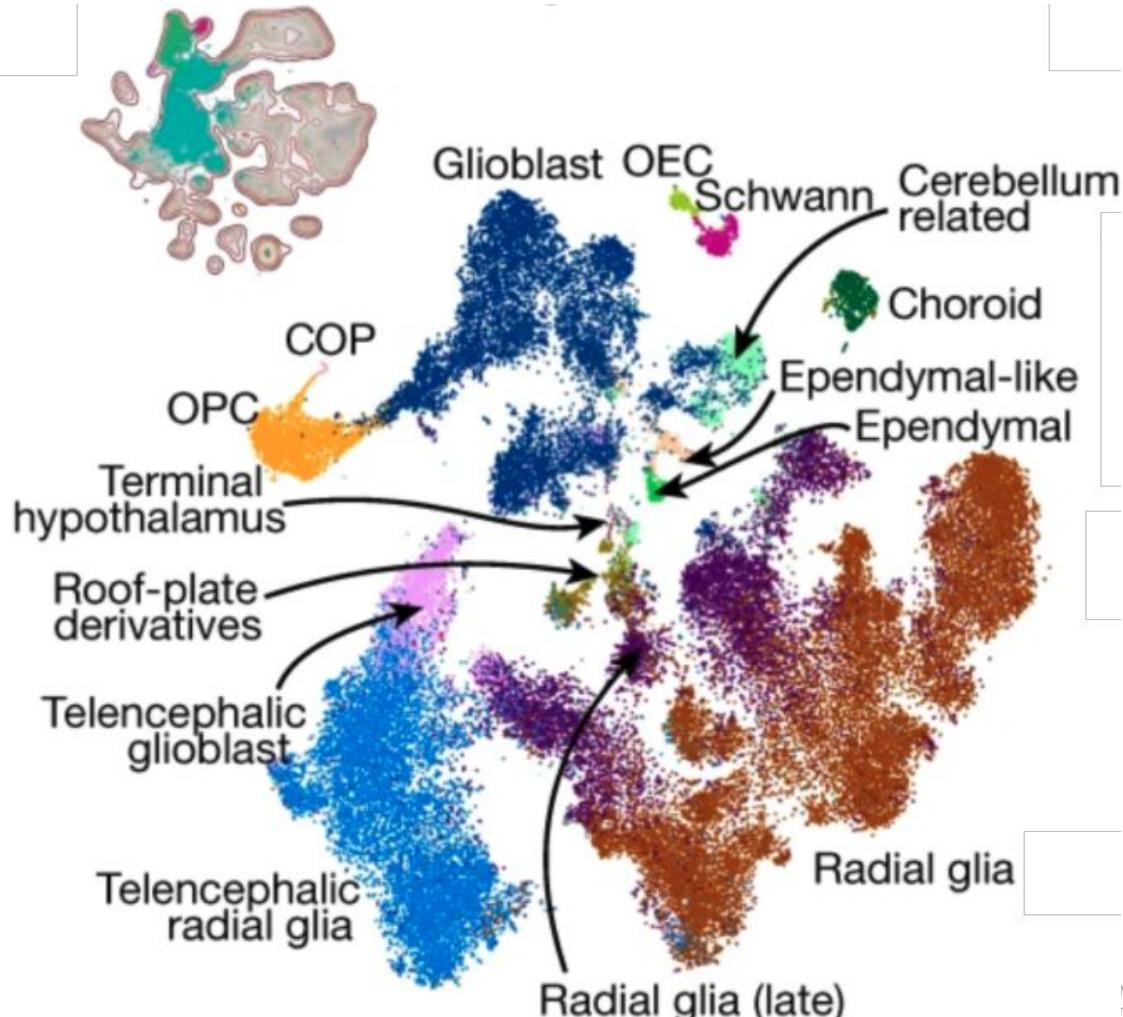


Single cell GEMs



## Single Cell RNA Sequencing Workflow





Thanks for your attention!



Sina.Majidian  
@gmail.com

