

# EMBL-EBI Resources and tools for genomics and proteomics

## Expression Atlas: gene expression results across species and conditions

Laura Huerta, PhD

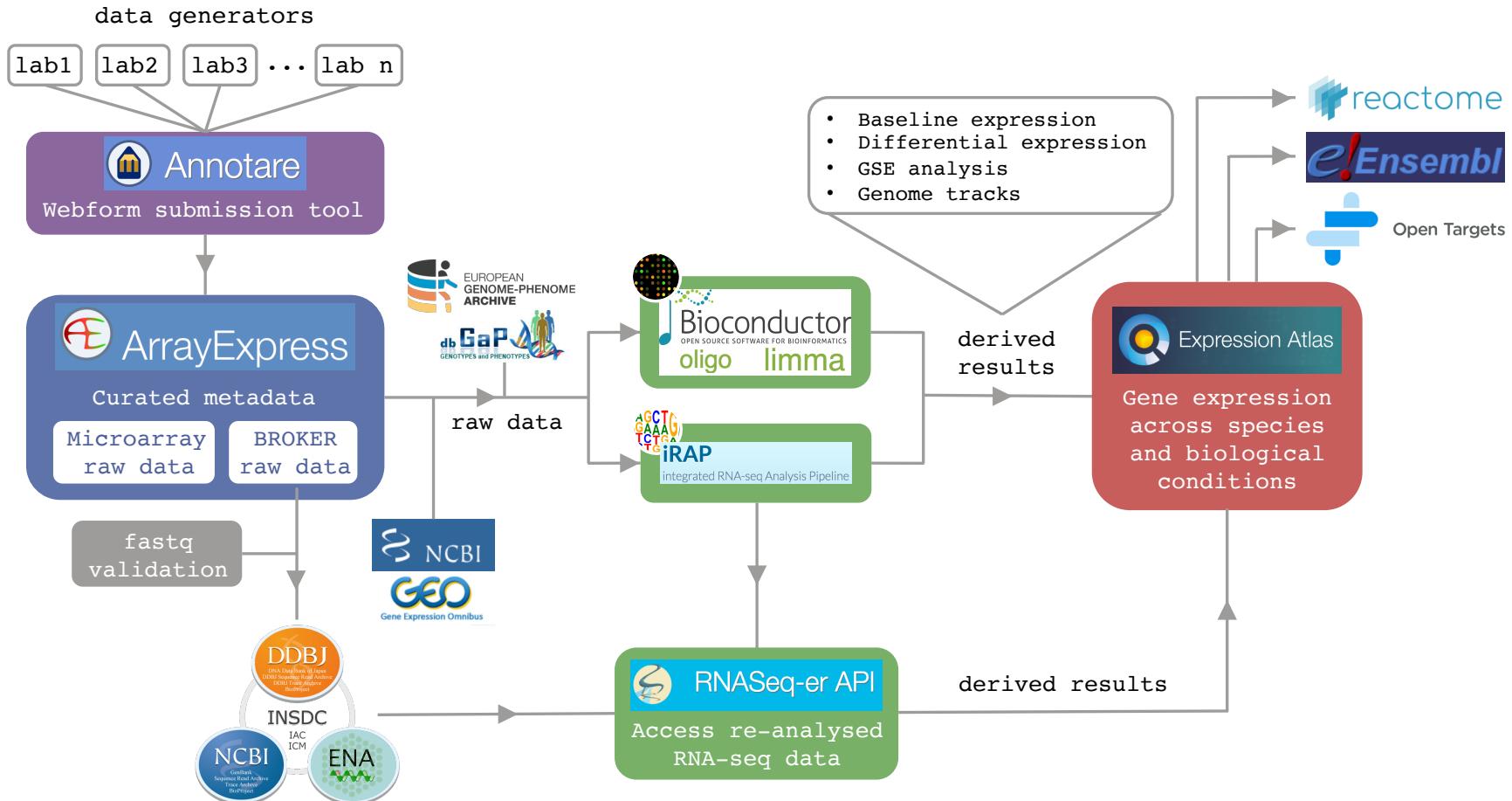
Senior Scientific Curator

[lauhuema@ebi.ac.uk](mailto:lauhuema@ebi.ac.uk)

6 December 2017



# Functional genomics resources at EMBL-EBI





# What is Expression Atlas?

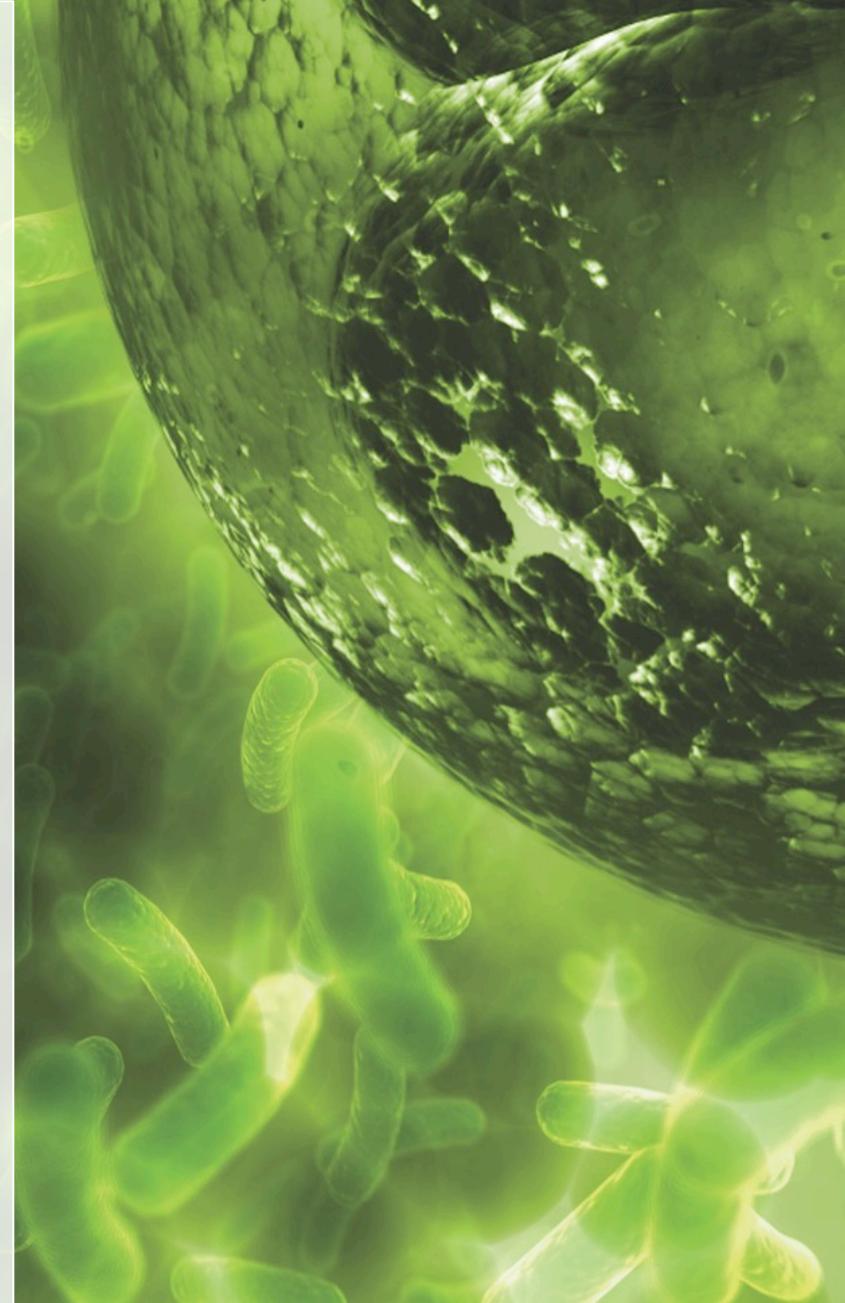


Reference resource for gene/protein expression  
across species and biological conditions

[www.ebi.ac.uk/gxa/home](http://www.ebi.ac.uk/gxa/home)

# OUR MISSION

To provide to the scientific community **freely available** information on the **abundance and localisation of RNA** (and proteins) across species and biological conditions





# What is Expression Atlas?

## Big Data



Generate and publish gene expression data



## Big Analysis



EBI infrastructure

## Big Results



Where is my favourite gene expressed?

How its expression changes in a disease?



# How ‘big’ is Expression Atlas data analysis?



<https://commonfund.nih.gov/gtex>

## Genetic effects on gene expression across human tissues

**GTEX Consortium, Lead analysts:, Laboratory, Data Analysis & Coordinating Center (LDACC):, NIH program management:, Biospecimen collection:, Pathology:, eQTL manuscript working group:, Alexis Battle, Christopher D. Brown, Barbara E. Engelhardt & Stephen B. Montgomery**

[Affiliations](#) | [Contributions](#) | [Corresponding authors](#)

*Nature* 550, 204–213 (12 October 2017) | doi:10.1038/nature24277

Received 08 September 2016 | Accepted 15 September 2017 | Published online 11 October 2017

doi:10.1038/nature24267

doi:10.1038/nature24265

doi:10.1038/nature24041



# How ‘big’ is Expression Atlas data analysis?

... e.g. GTEx dataset



**550 donors**  
**53 tissues**

**19,000** paired  
RNA-sequencing libraries

**22 years**  
of processing time in a  
single machine

**78 terabytes**  
of EMBL-EBI storage capacity



# Large-scale RNA-seq experiments

Access gene expression results of large-scale datasets



Basic research



Key cell line  
models



Genentech

Cancer research



Proteomics



Zebrafish development



Prenatal human brain



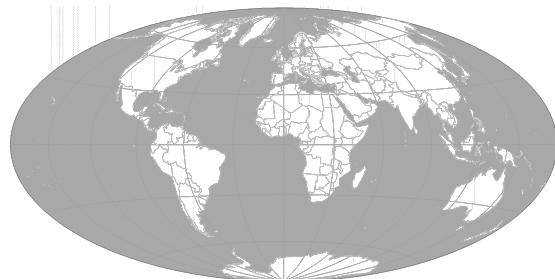
DMDD

Deciphering the Mechanisms  
of Developmental Disorders



Mouse models

# Expression Atlas contains thousands of selected and curated datasets



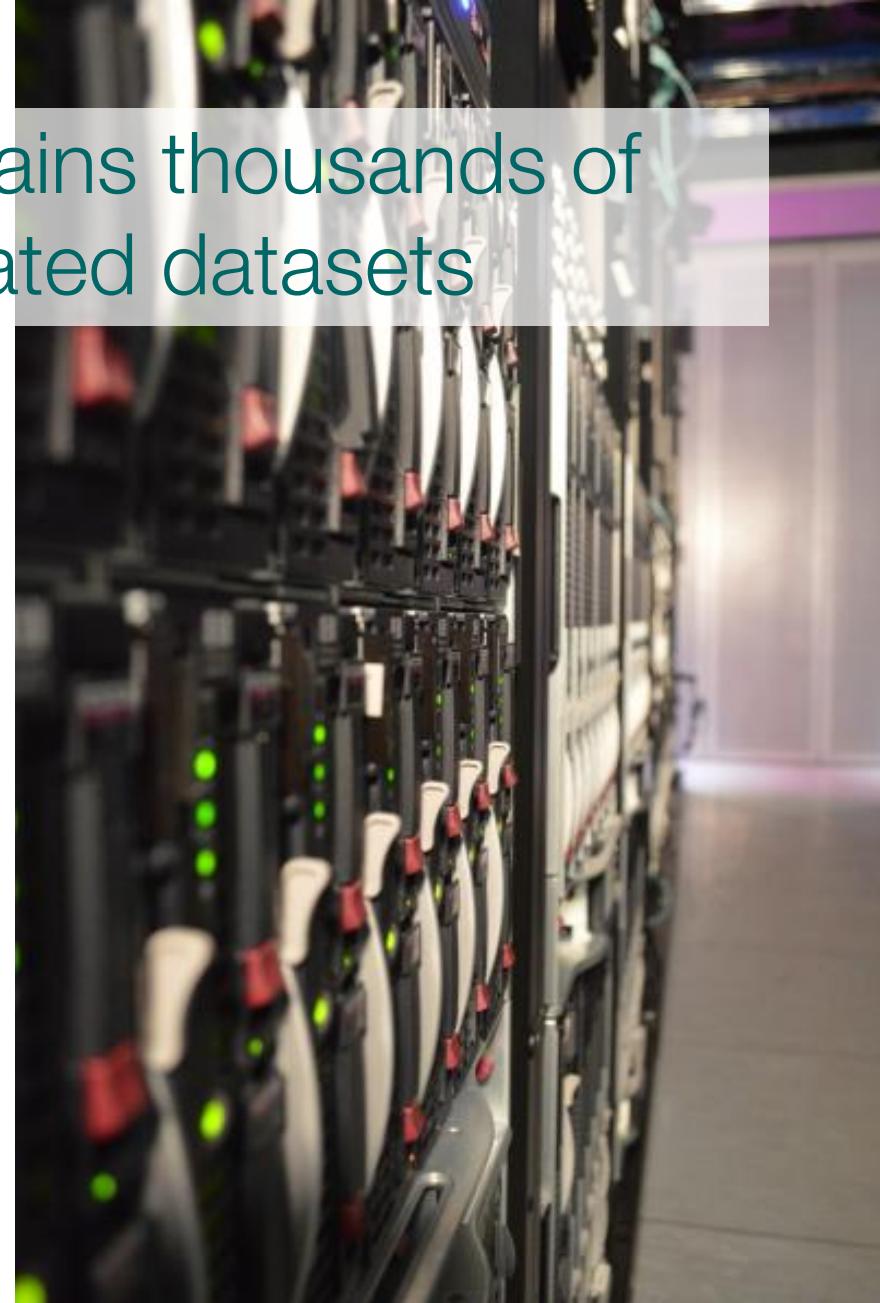
> 3,000 datasets



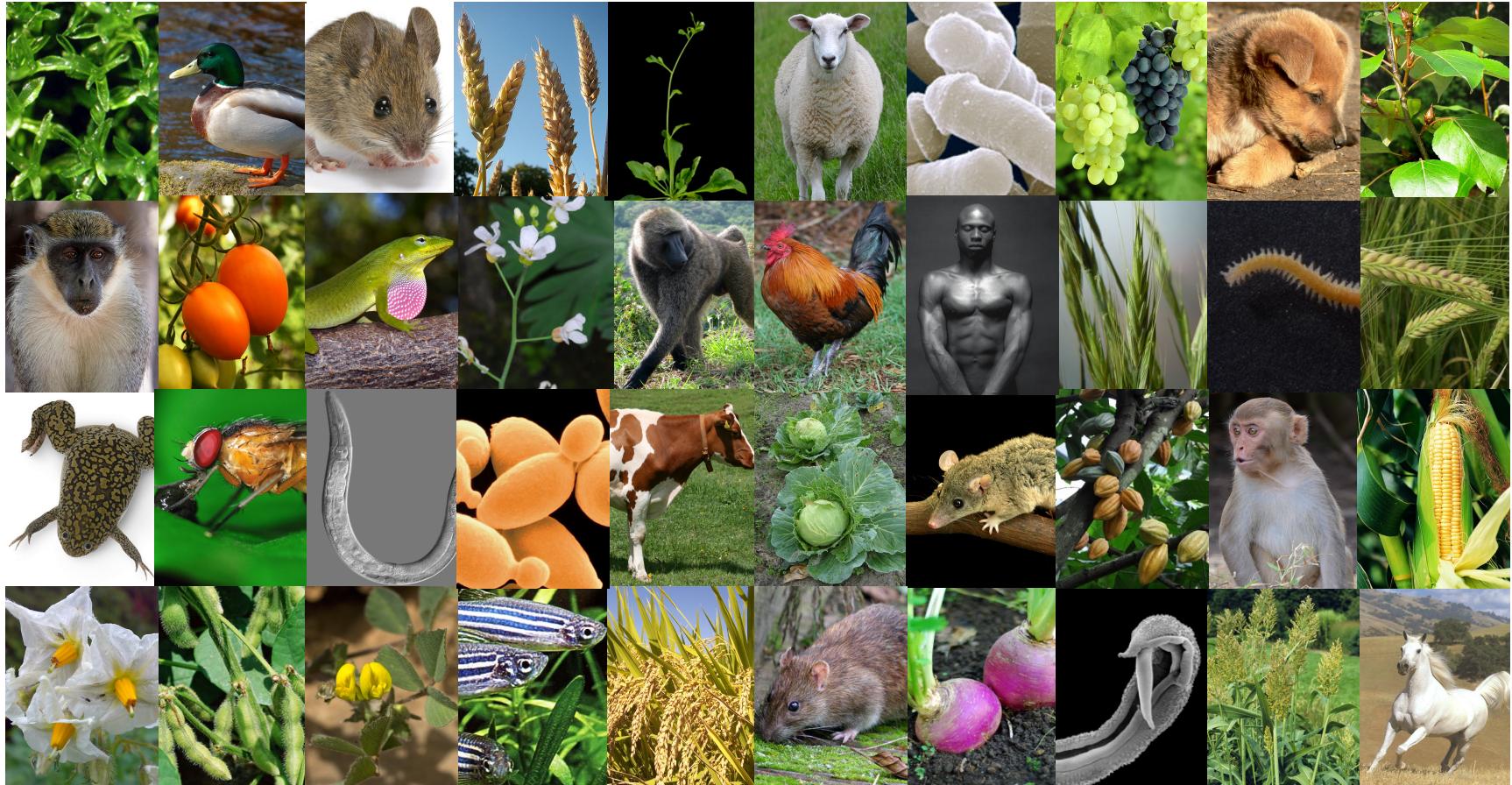
> 500 RNA-sequencing data



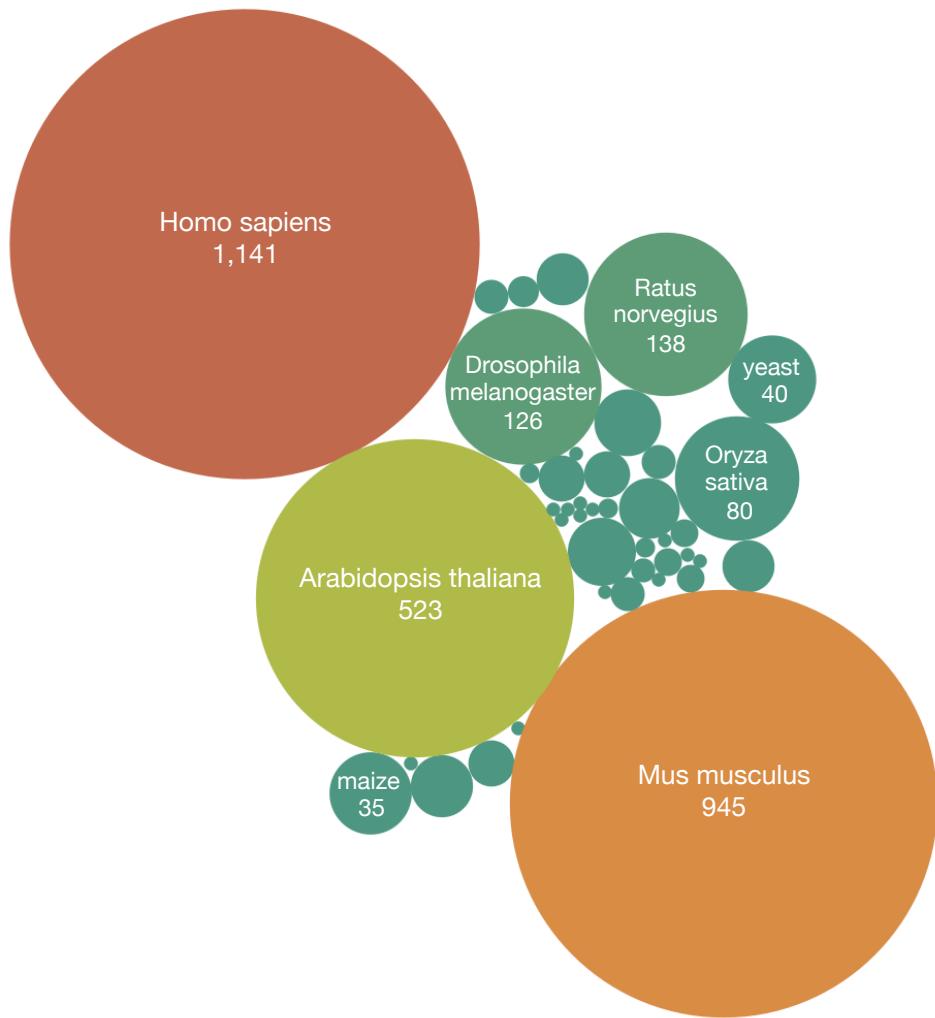
> 2,500 microarray data



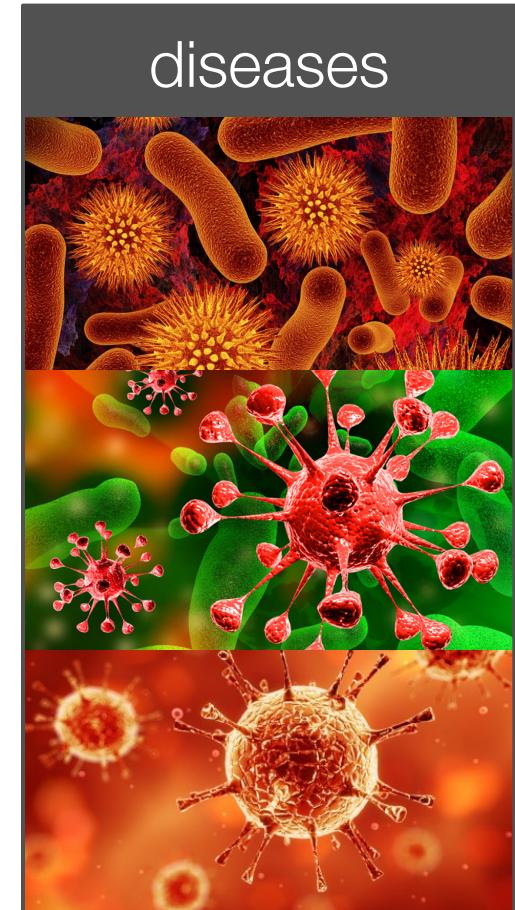
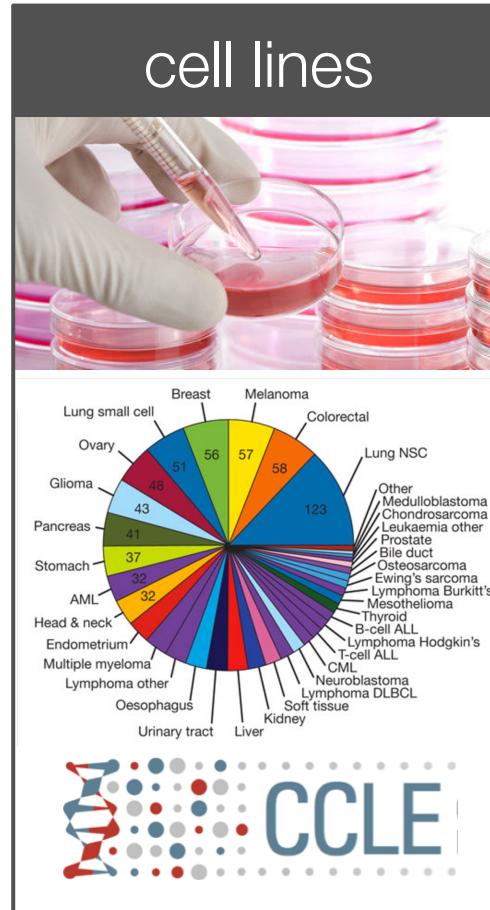
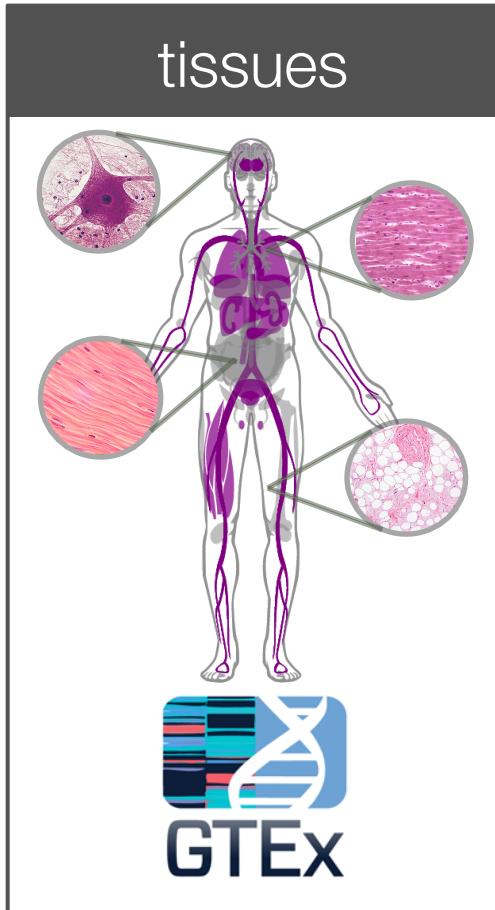
... across more than 40 species ...



... across more than 40 species ...



# ... under different biological conditions ...



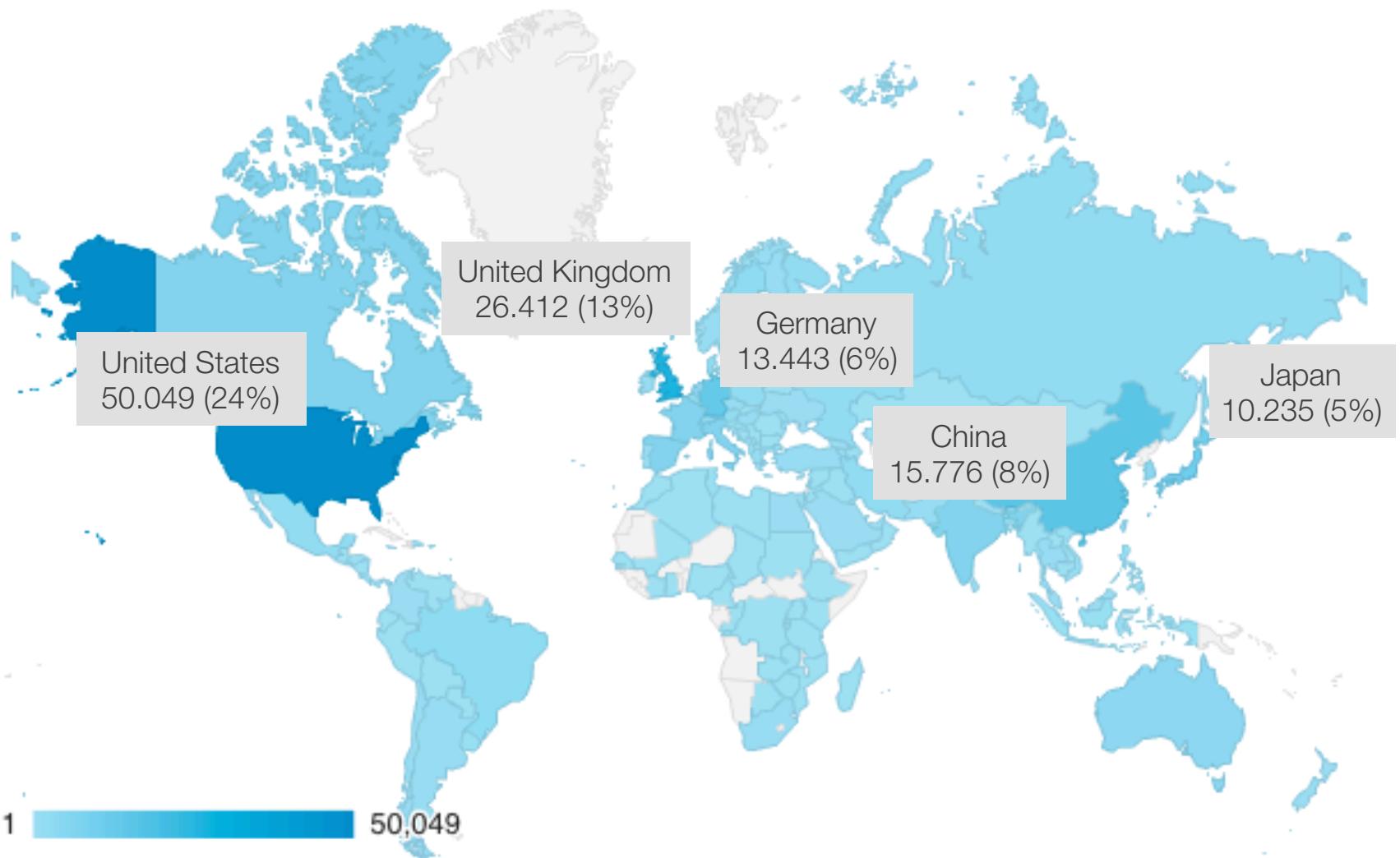
# ... to help scientists find answers ...

In which conditions  
is my favourite gene  
expressed?

What genes are  
differentially expressed in  
a particular condition?

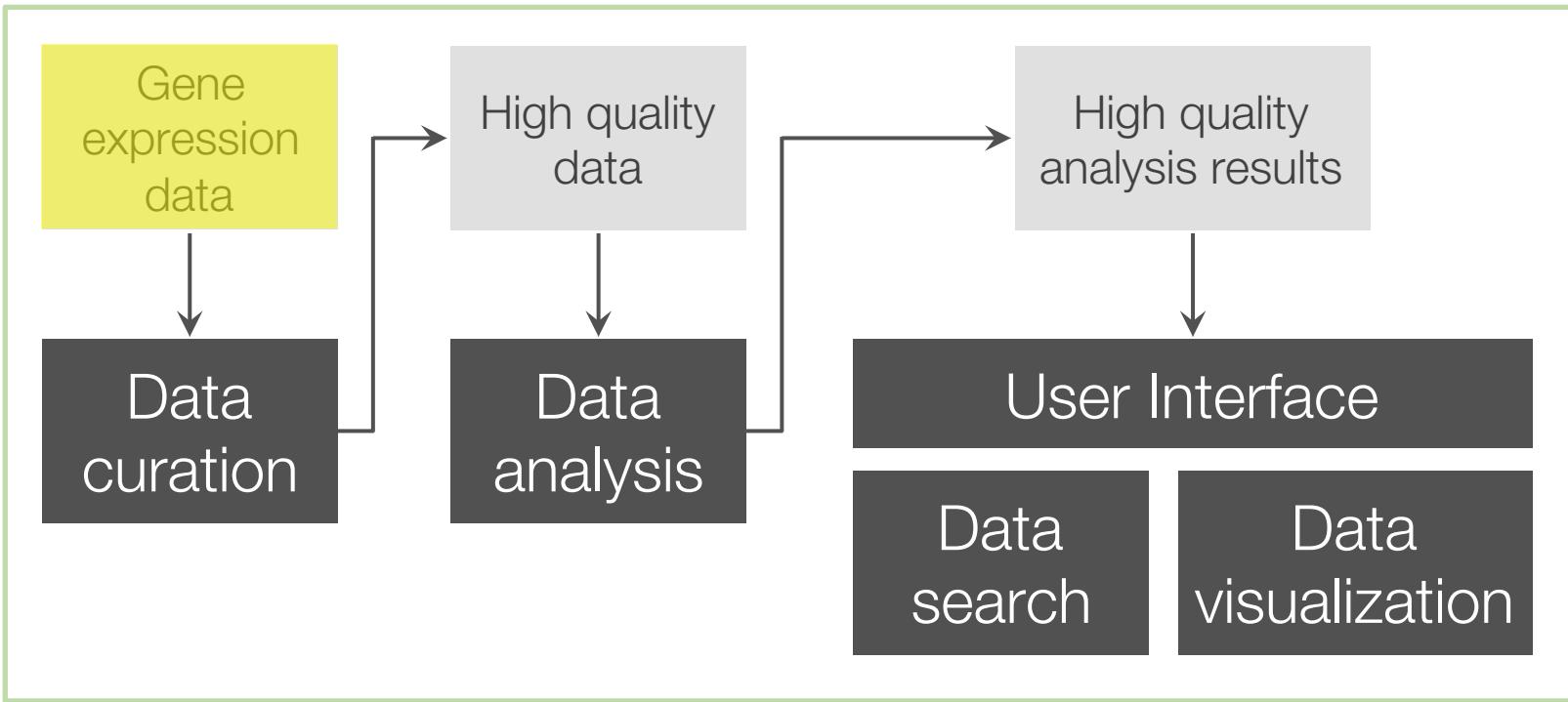
What genes are  
expressed in a  
particular condition?

... all across the world ...





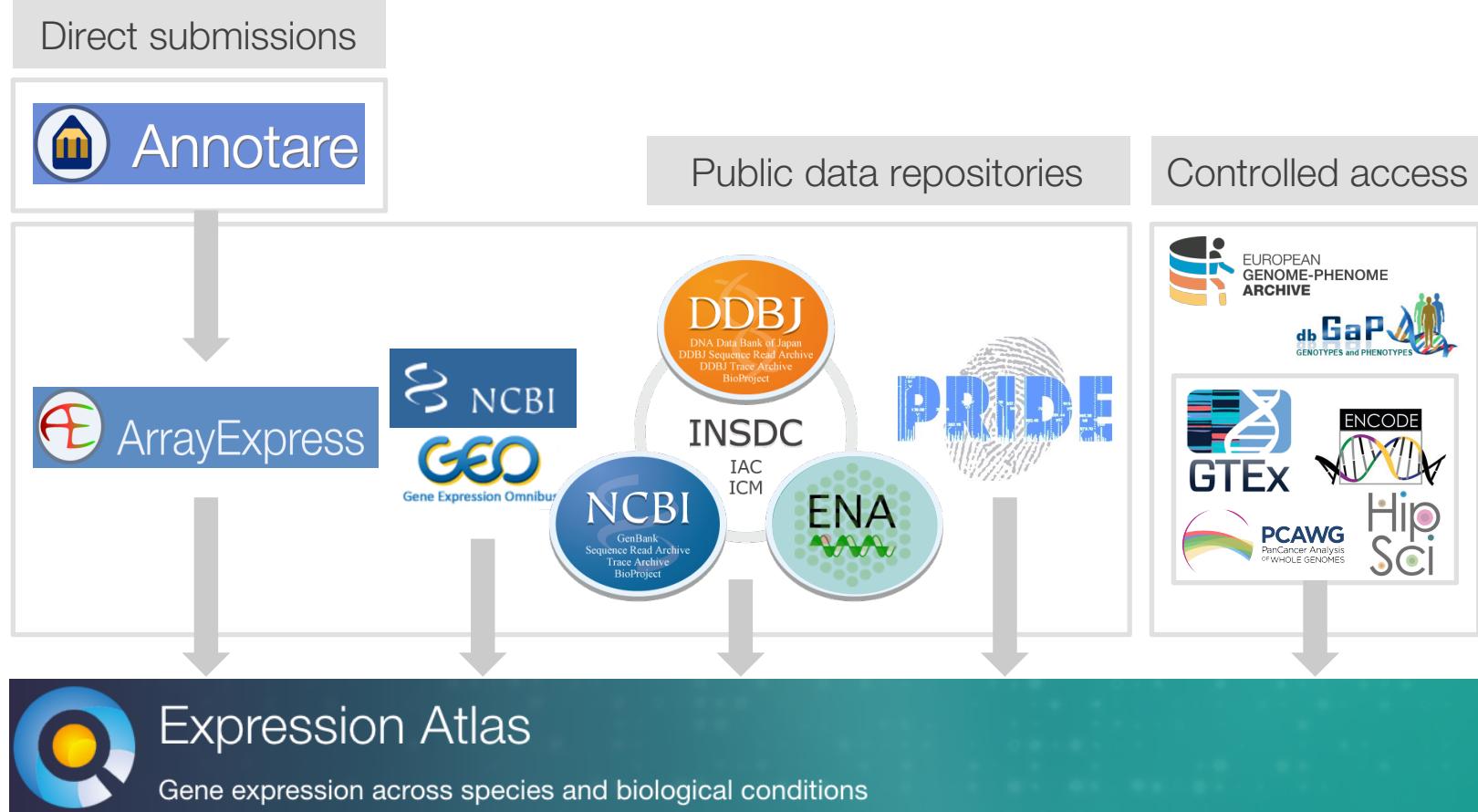
# How does Expression Atlas do it?



Discover and interpret gene expression analysis results quickly and easy

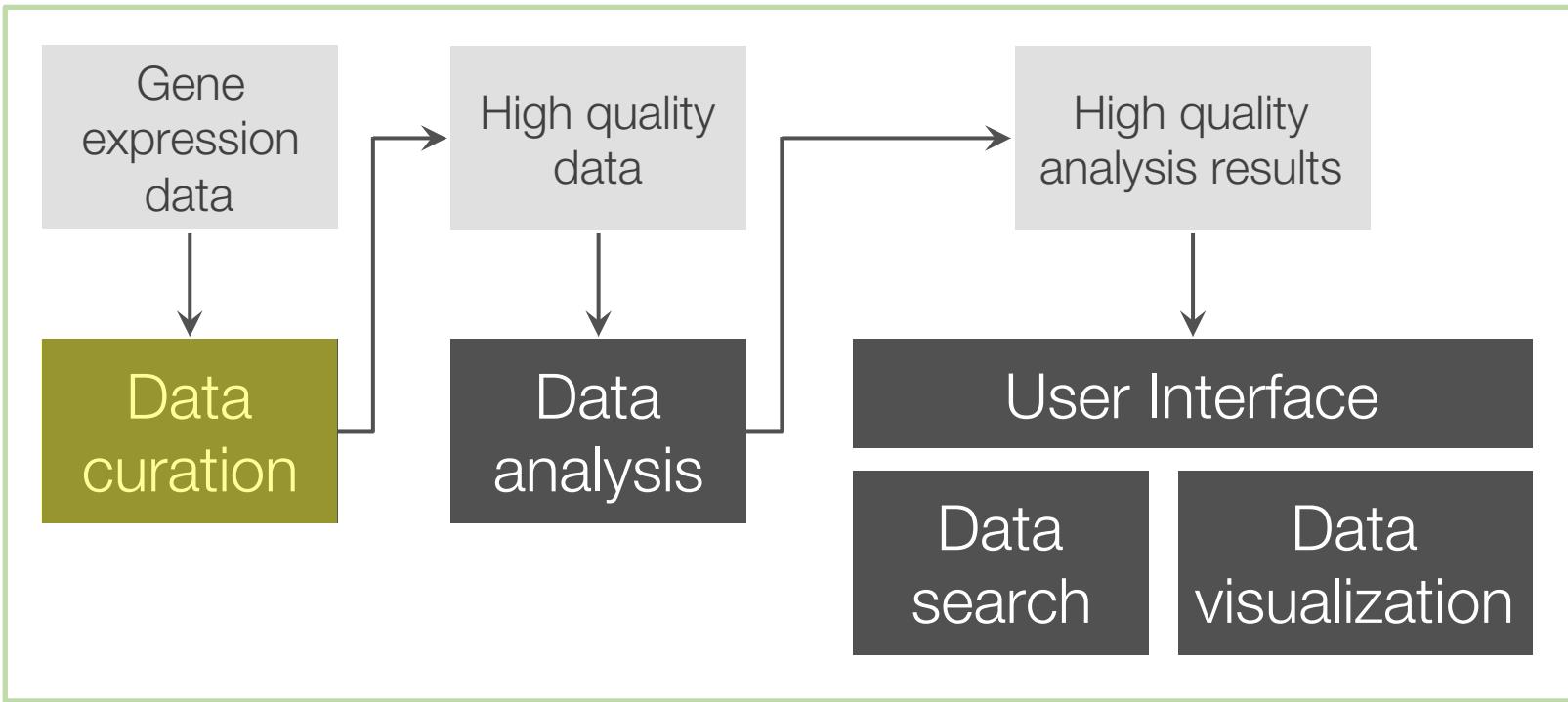


# Where does expression data come from?





# How does Expression Atlas do it?



Discover and interpret gene expression analysis results quickly and easy

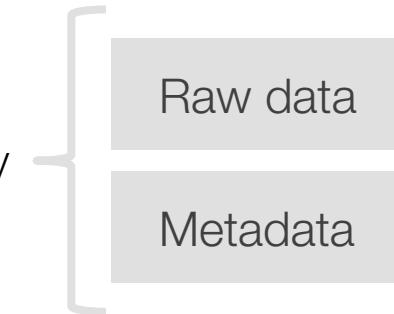


# Expression Atlas – data curation

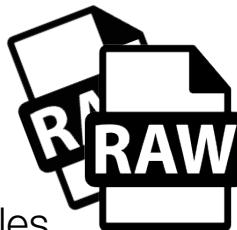
## Biocuration



Translation and integration of information relevant to biology into a database or resource



### Raw data



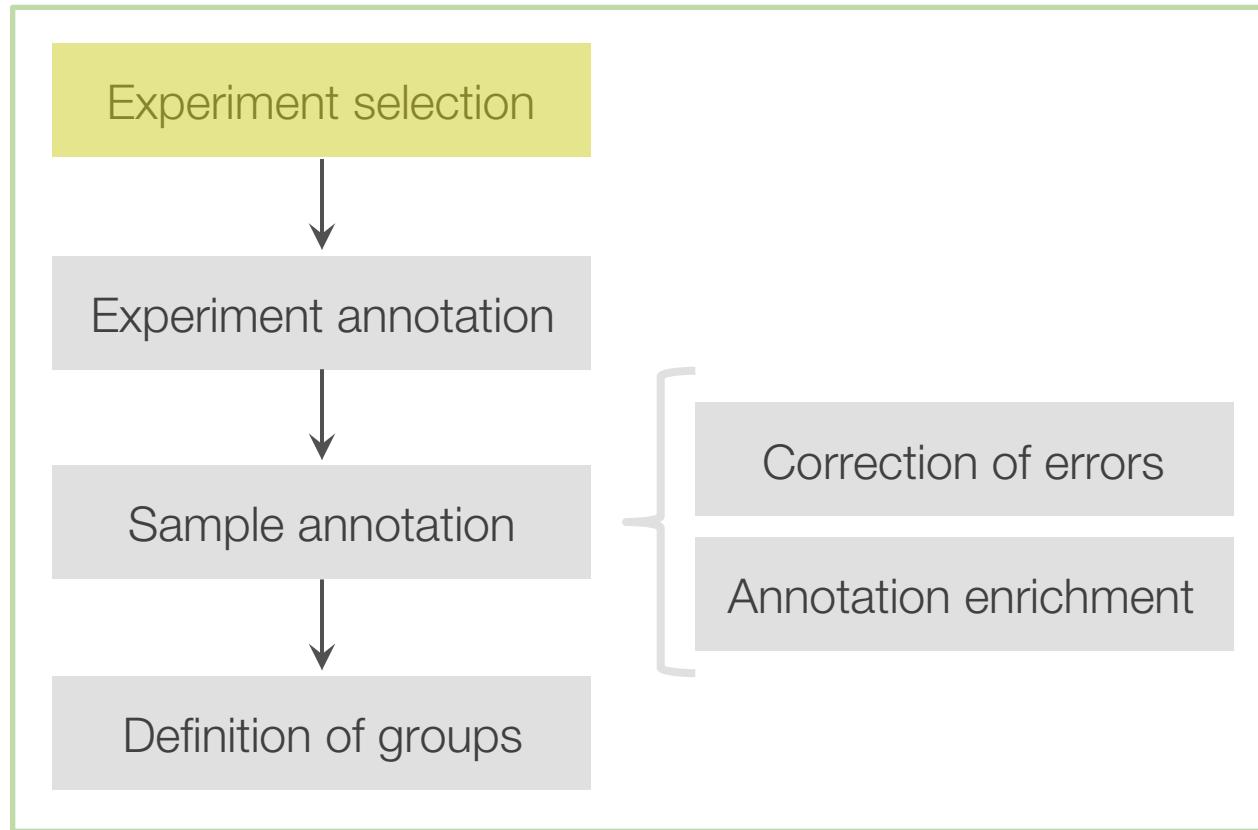
- unprocessed data files
- **Microarray:** files from the scanner (e.g. Affymetrix CEL files, Agilent feature extraction *txt* files, Illumina *idat* files)
- **Sequencing:** raw sequence read files (e.g. FASTQ files)

### Metadata

- Experiment description
- Experiment title
- Sample annotation
- Protocols
- Publication details (if any)
- Author contact details



# Expression Atlas – data curation





# Expression Atlas – data curation

## Experiment selection

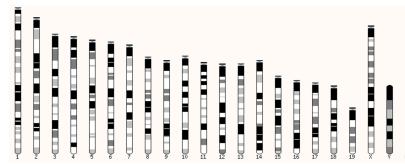


Raw data



Experiment type

[www.ensembl.org](http://www.ensembl.org)



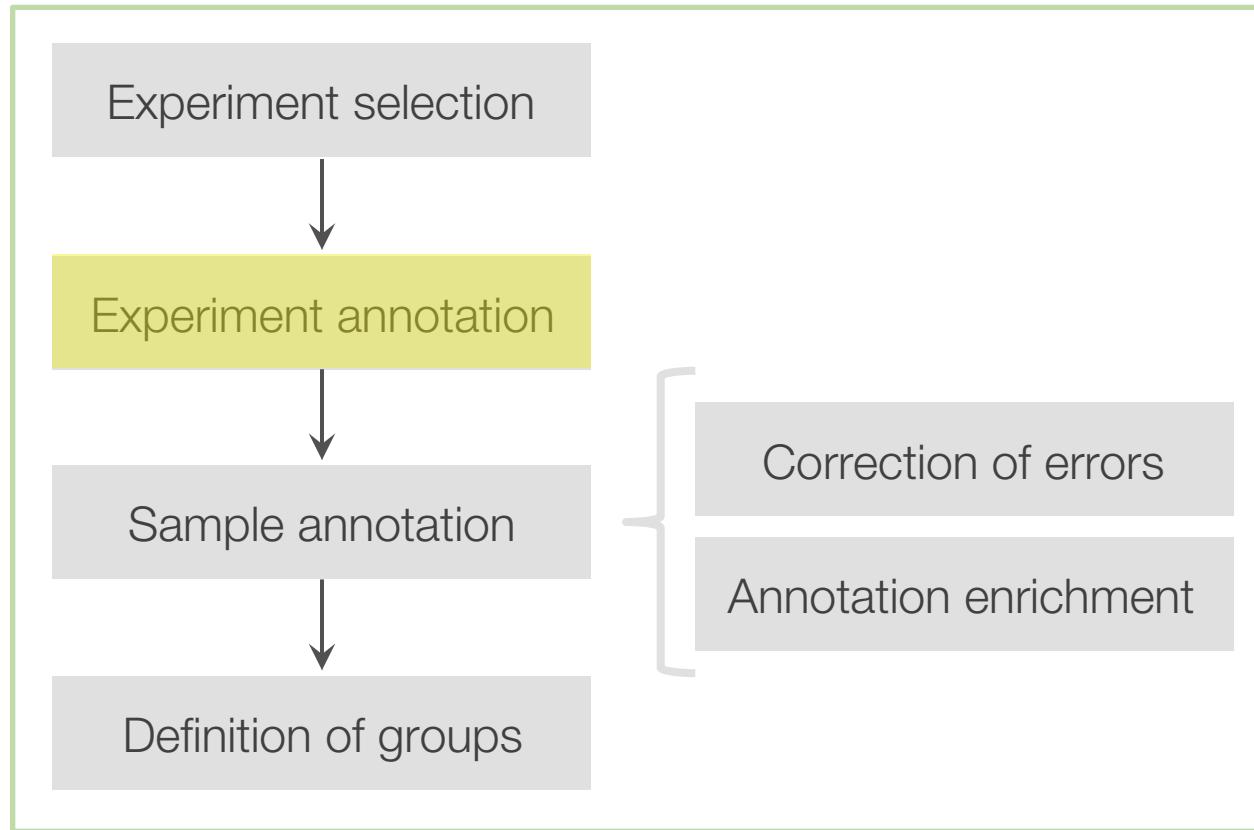
Reference genome



Biological replicates



# Expression Atlas – data curation





# Expression Atlas – data curation

## Experiment annotation



Clear experiment intent

<title> "rice transcriptome project" →

*"Transcription profiling by high throughput sequencing of rice oscerk1 mutants treated with two fungus-secreted substances"*

Informative title



Protocol  
description



Europe PMC

□ Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans.  
(PMID:25954001 PMID:PMC4547484)

Abstract · Citations · BioEntities · Related Articles · External Links

GTEx Consortium  
+ Collaborators (139)

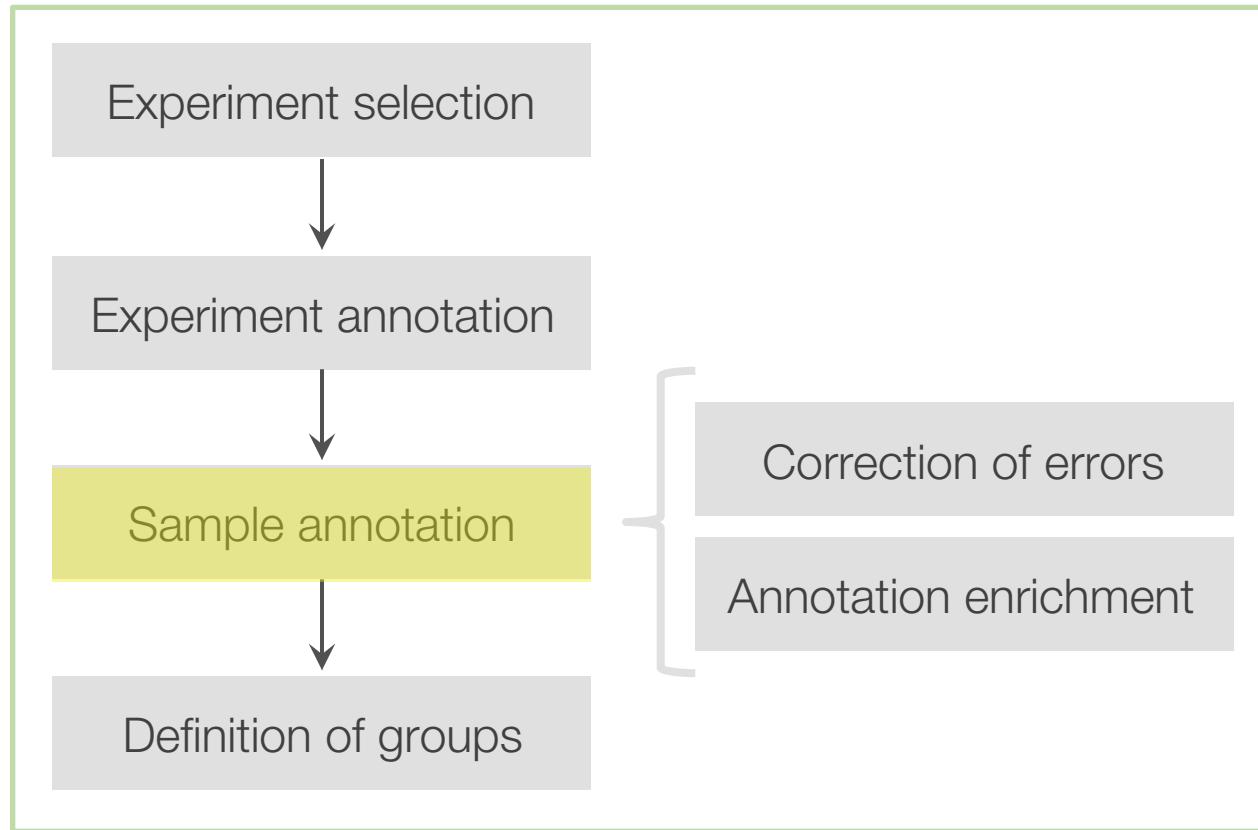
Science (New York, N.Y.) [07 May 2015, 348(6235):648-660]

Type: Research Support, Non-U.S. Gov't, research-article, journal Article, Research Support, N.I.H., Extramural  
DOI: 10.1126/science.1262110

Publication details

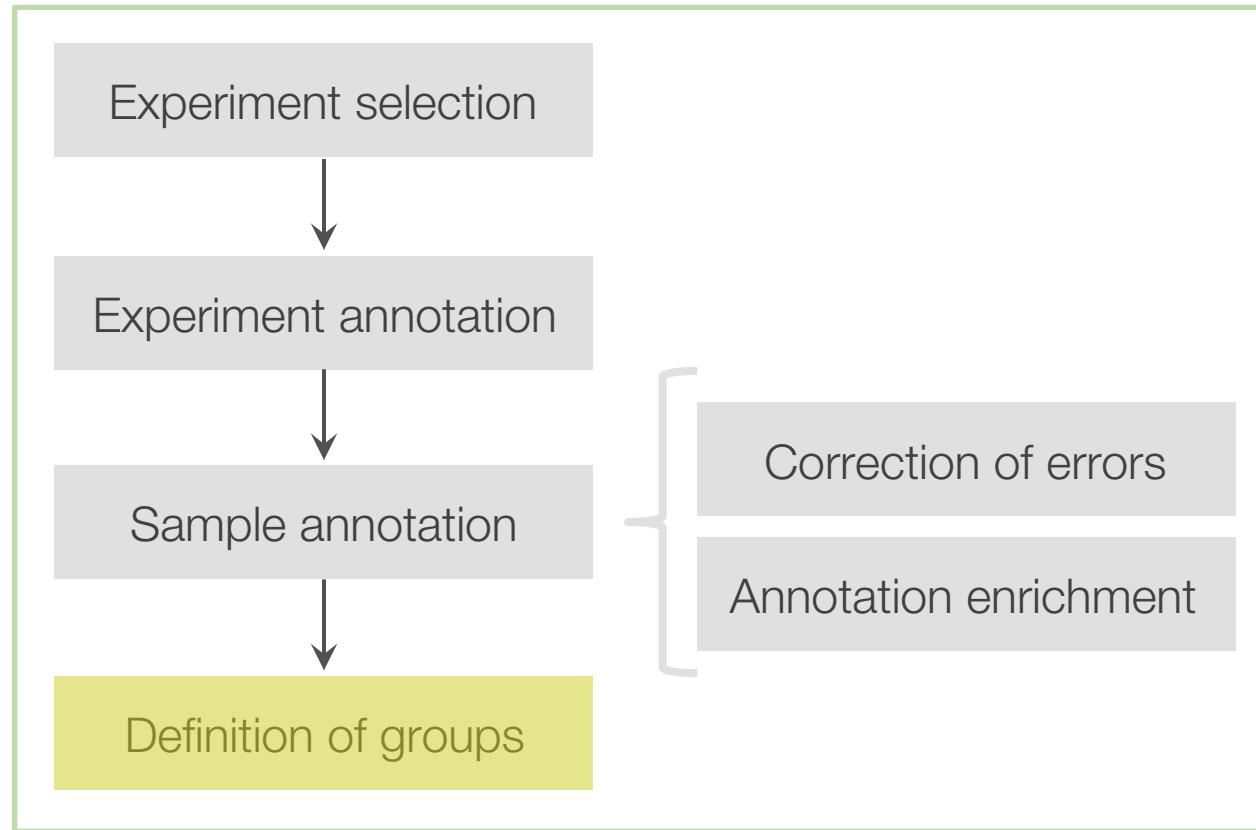


# Expression Atlas – data curation



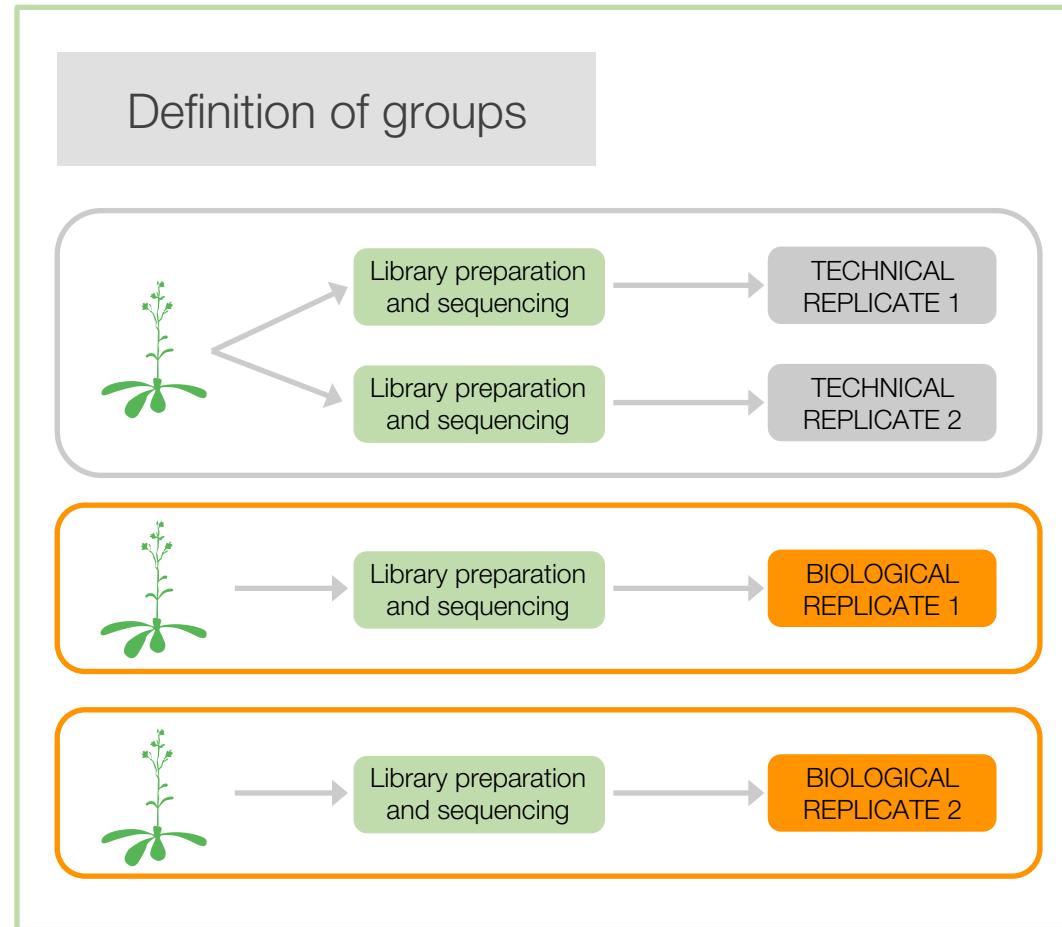


# Expression Atlas – data curation



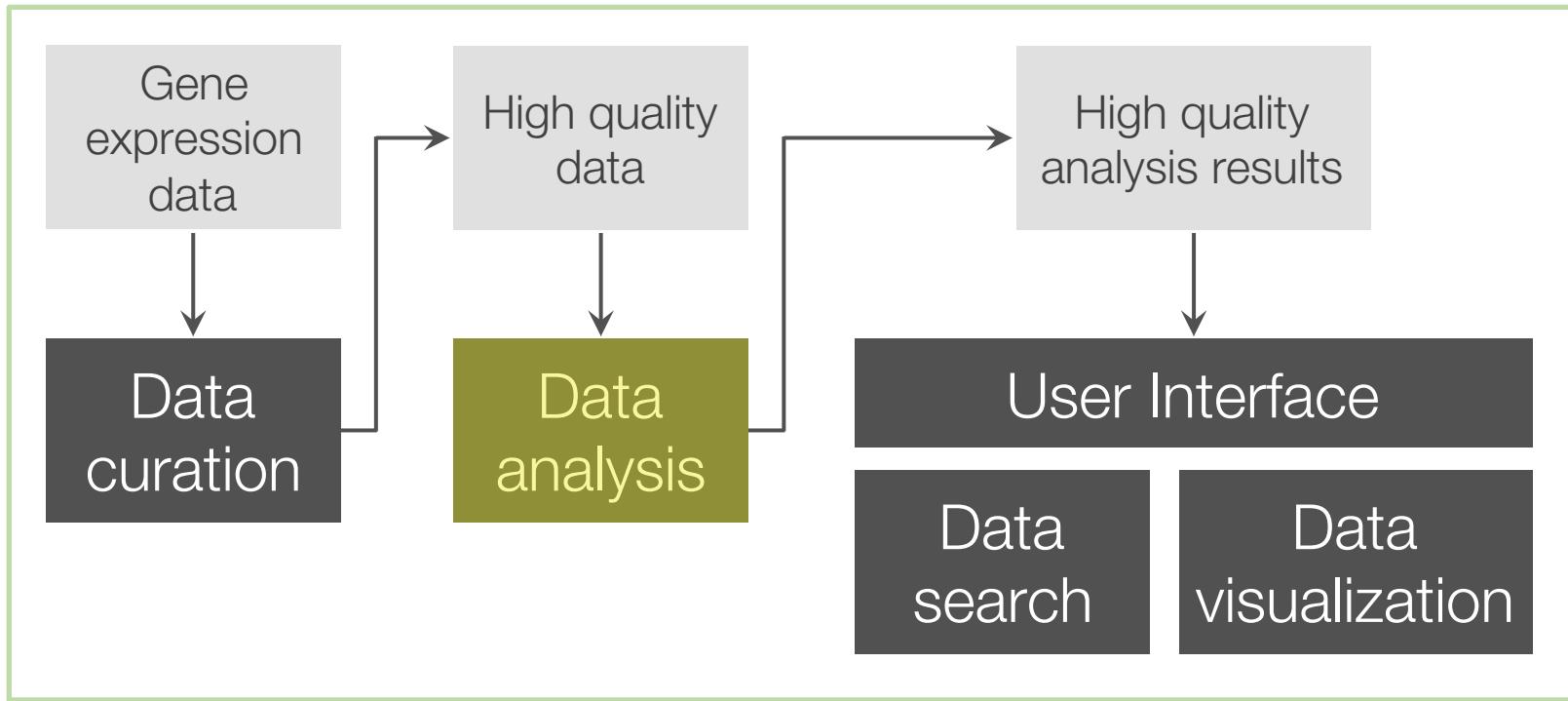


# Expression Atlas – data curation



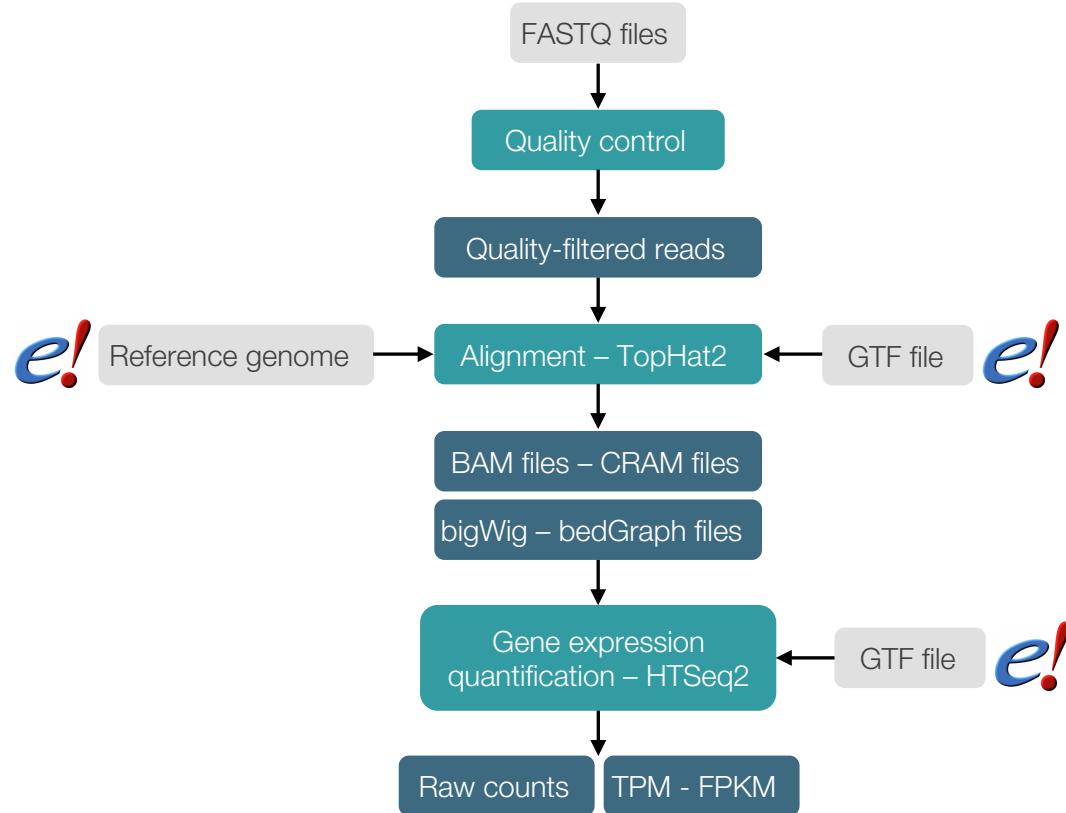


# How does Expression Atlas do it?



Discover and interpret gene expression analysis results quickly and easy

# RNA-seq data analysis



<https://github.com/nunofonseca/irap/>

# RNA-seq data analysis

## Pages

[New Page](#)

<a href="#">1 News</a>	Last updated on Jan 20
<a href="#">2 How to install and update iRAP</a>	Last updated on May 11
<a href="#">3 How to run iRAP</a>	Last updated on Jan 31
<a href="#">4 Output files and directory structure</a>	Last updated on Mar 23
<a href="#">5 Tools supported</a>	Last updated on Dec 14, 2016
<a href="#">6 Virtual Machines</a>	Last updated on Feb 16, 2016
<a href="#">7 Quick Example</a>	Last updated on Dec 20, 2016
<a href="#">8 Configuration file</a>	Last updated on Aug 21
<a href="#">9 Gene set analysis in iRAP</a>	Last updated on Mar 19, 2015
<a href="#">FAQ</a>	Last updated on Jun 3, 2016
<a href="#">GSE</a>	Last updated on Feb 16, 2016
<a href="#">iRAP single library</a>	Last updated on Dec 20, 2016
<a href="#">LSF</a>	Last updated on Feb 16, 2016

<https://github.com/nunofonseca/irap/wiki>



# RNA-seq data analysis

## Train online

[Training](#)[Train online](#)[About Train online](#)[Glossary](#)[Support and feedback](#)[Login/register](#)

Functional genomics (II):  
Common technologies and  
data analysis methods

[Introduction](#)[Real-time PCR](#)[Microarrays](#)[Next Generation Sequencing \(NGS\)](#)[RNA sequencing](#)[Biological interpretation of gene expression data](#)[Genotyping, epigenetic and DNA/RNA-protein interaction methods](#)[Summary](#)[Quiz: Check your learning](#)[Your feedback](#)[Learn more](#)[References](#)[Contributors](#)

## Functional genomics (II): Common technologies and data analysis methods

This is the second part of our functional genomics course. This course will introduce you to common technologies in functional genomics studies, including microarrays and next generation sequencing (NGS), with a special focus on RNA-seq. We will look at important considerations when designing your experiments, data analysis methods, and discuss when to use one technology over another.

We recommend that you look at [Functional genomics \(I\): Introduction and designing experiments](#) before starting this course.

### About this course

**Author(s):** Laura Huerta, Melissa Burke

Gene Expression

Beginner

2 hours

[Start the course](#)

### Learning objectives:

- Evaluate different functional genomics technologies and know when to use them
- List some of the considerations when designing your own functional genomics experiments using these technologies
- Describe the steps involved in data analysis of microarray and RNA-seq experiments

<https://www.ebi.ac.uk/training/online>



# RNA-seq data analysis

Functional genomics (II):  
Common technologies and  
data analysis methods

Introduction
Real-time PCR
▪ Microarrays
Next Generation Sequencing (NGS)
▪ RNA sequencing
Advantages of RNA-seq over hybridisation-based approaches
Applications of RNA-seq
▪ Performing a RNA-seq experiment
▪ Design considerations
Number of replicates
Laboratory performance
▪ Data analysis
Quality control
Read mapping or alignment
Quantification
Differential gene expression analysis
▪ Biological interpretation of gene expression data
▪ Genotyping, epigenetic and DNA/RNA-protein interaction methods
Summary
Quiz: Check your learning
Your feedback
Learn more
References
Contributors



All course materials in Train online  
are free cultural works licensed  
under a Creative Commons  
Attribution-ShareAlike 4.0  
International license

## Performing a RNA-seq experiment

There is no common pipeline for all the different types of RNA-seq. Both the experimental design and the analysis procedures will be different depending on the organism being studied and the research goals. As the primary objective of many biological studies is a comparison of gene expression between samples, we outline a possible workflow for detecting differential expression using RNA-seq (Figure 7).

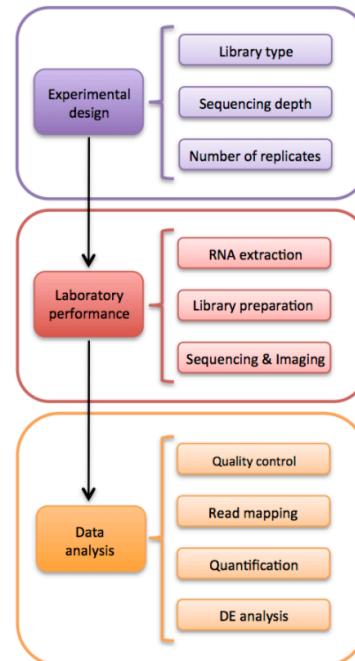


Figure 7 Overview of a RNA-seq experiment for detecting differential expression (DE).

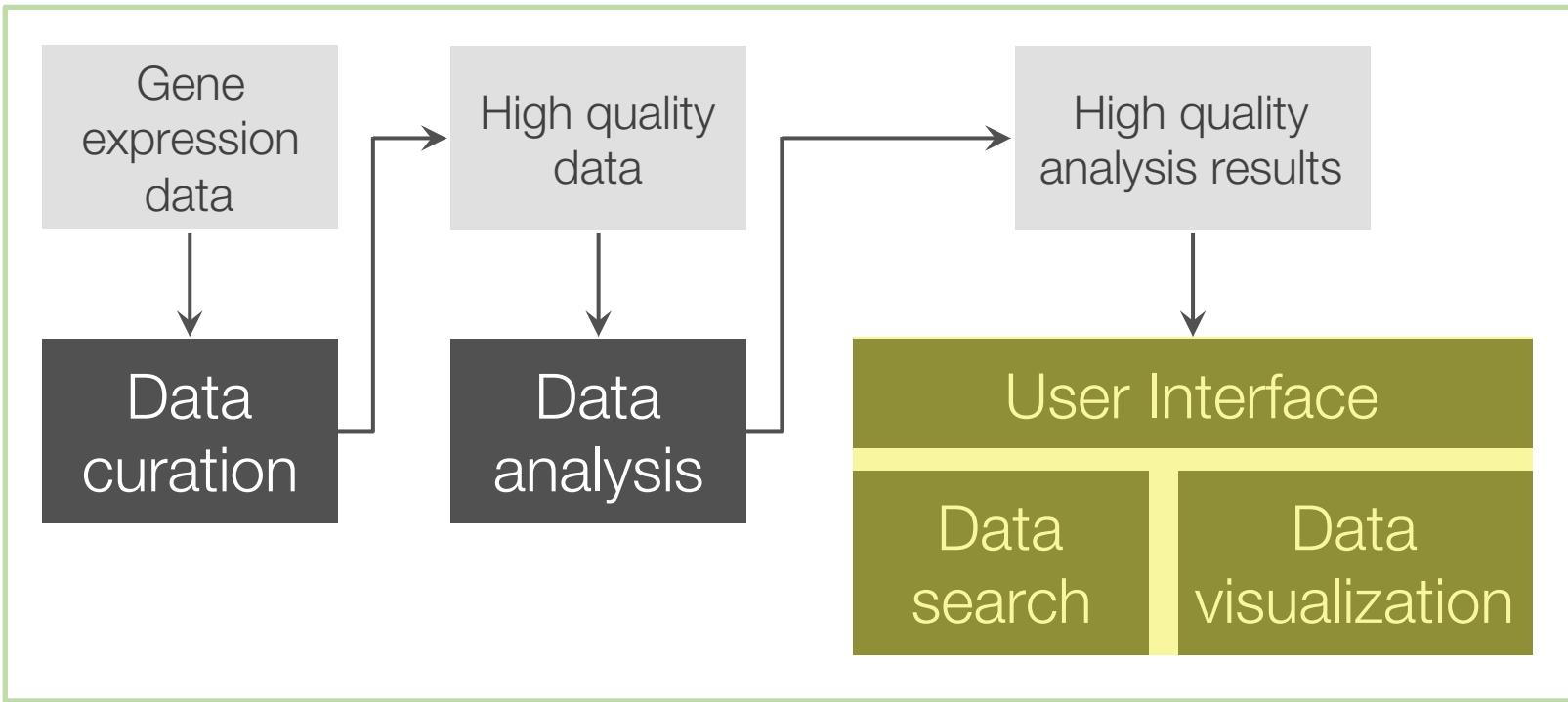
« Applications of RNA-seq

Next: Design considerations »

<https://www.ebi.ac.uk/training/online>



# How does Expression Atlas do it?



Discover and interpret gene expression analysis results quickly and easy



# Let's try Expression Atlas

## Expression Atlas

Gene expression across species and biological conditions

Home | Download | Release notes | FAQ | Help | Licence | About

Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how its expression changes in a disease?'.

[Read more about Expression Atlas](#)

Search Gene set enrichment

Gene / Gene properties  
CFHR2 [x](#)  
Examples: REG1B, zinc finger  
[Search](#) [Clear](#)

Organism  
Homo sapiens [▼](#)

Biological conditions  
Enter condition query...  
Examples: lung, leaf, valproic acid, cancer

[wwwdev.ebi.ac.uk/gxa/home](http://wwwdev.ebi.ac.uk/gxa/home)



# Let's try Expression Atlas

## Expression Atlas

Gene expression across species and biological conditions

Home | Download | Release notes | FAQ | Help | Licence | About

Exploring gene expression results across species under different biological conditions

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others. Expression Atlas aims to help answering questions such as 'where is a certain gene expressed?' or 'how its expression changes in a disease?'.

[Read more about Expression Atlas](#)

**In which conditions is my favourite gene expressed?**

Organism

- ✓ Homo sapiens
- Mus musculus
- Arabidopsis thaliana
- Rattus norvegicus
- Drosophila melanogaster
- Zea mays

Any

- Anas platyrhynchos
- Anolis carolinensis
- Anopheles gambiae
- Arabidopsis lyrata
- Arabidopsis thaliana
- Aspergillus fumigatus
- Bos taurus
- Brachypodium distachyon
- Brassica oleracea
- Brassica rapa
- Caenorhabditis elegans

Biological conditions

Enter condition query...

Examples: lung, leaf, valproic acid, cancer

Discover and interpret gene expression analysis results quickly and easy



# Baseline expression results

Baseline expression

Differential expression

Show anatomograms

Filter your results

Homo sapiens

Organism part

Cell line

Danio rerio

Developmental stage

Mus musculus

Organism part

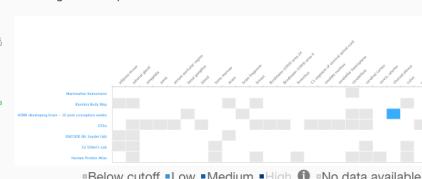
Cell type

Developmental stage

Homo sapiens — Organism part



Showing 7 of 7 experiments found:



All or  
which

Expression levels are  
visualised using heatmaps

Danio rerio — Developmental stage

Showing 1 of 1 experiment found:

Experiment: 18 White et al  
Experimental condition: larval day 5  
Expression level: 26 FPKM

By experiment type  Filters  Download table content

Mus musculus — Organism part



Showing 23 of 23 experiments found:



Below cutoff

Low

Medium

High



No data available



# Baseline expression results

Results for CFHR2

Baseline expression      Differential expression

Show anatomograms

Organism part

Filter your results

Homo sapiens

Organism part

Cell line

Danio rerio

Developmental stage

Mus musculus

Organism part

Cell type

Developmental stage

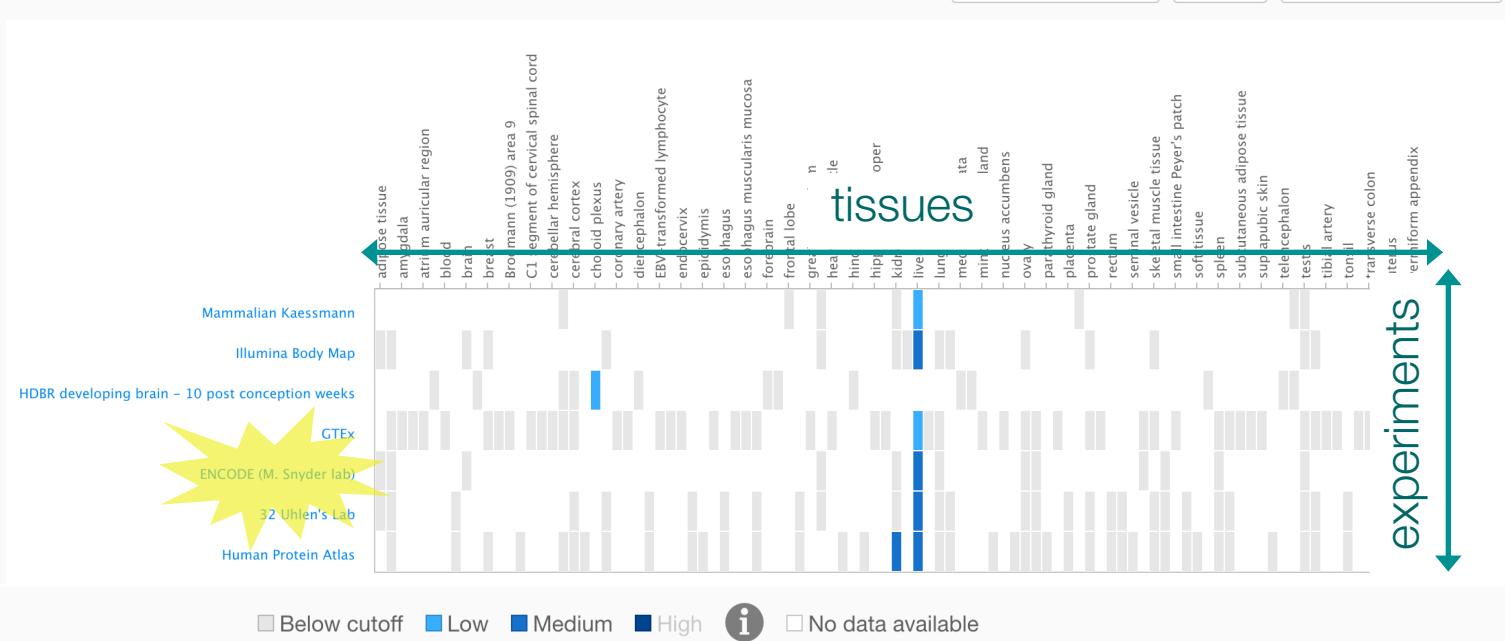
**HEATMAP:** For each species, *CFHR2* gene expression values for the selected condition (columns) in all experiments (rows) in which *CFHR2* gene is expressed

Showing 7 of 7 experiments found:

By experiment type ▾

Filters

Download table content





# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues  
RNA-Seq mRNA Baseline  
Organism: *Homo sapiens*  
Raw Data Provider: The Human Protein Atlas

Results   Experiment Design   Supplementary Information   Download

Genes



Showing 50 of 34,068 genes found:

Expression level in FPKM

0 3.659

PRM2  
MHN2  
THP1  
PRK27  
ANHBD7  
FLC2  
CAUS14  
F2  
C10orf68  
TUBA1C  
SMCP  
F9  
RODIL2  
CYP1A2  
ULP1  
KRT77  
SLC10A1  
LCZ28  
HAO1  
FATE1  
CTD-212N2N3.1  
PCK2  
SERPINA11  
GDF1  
AKAP4  
SPATA8  
PRB2  
SLC10A1  
ERVH48-1  
SERPINA12  
TSHR  
SLC12A1  
LCZ1C  
S100G  
LCZ1A  
C16orf82  
SP93  
MBL2  
C10orf62  
UBRN1  
LCZ6A  
HHLB  
ACTL7A  
TH  
HMGB4  
DDX4  
C10orf12  
ACVR2C  
SLC34A1  
PRK17

Specific checked

Most specific

Expression value

0.5

See distribution

Data units

TPM  
 FPKM

Organism parts

Select

Selected: 32 / 32

**HEATMAP:** top 50 genes (rows)  
most specifically expressed  
across all tissues (columns)

Genes expressed only  
in one tissue first  
Within that, genes with  
highest expression first



# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues  
RNA-Seq mRNA baseline  
Organism: Homo sapiens  
Raw Data Provider: The Human Protein Atlas

Results   Experiment Design   Supplementary Information   Download

Genes

Specific unchecked

Most specific

Expression value

0.5

See distribution

Data units

TPM

FPKM

Organism parts

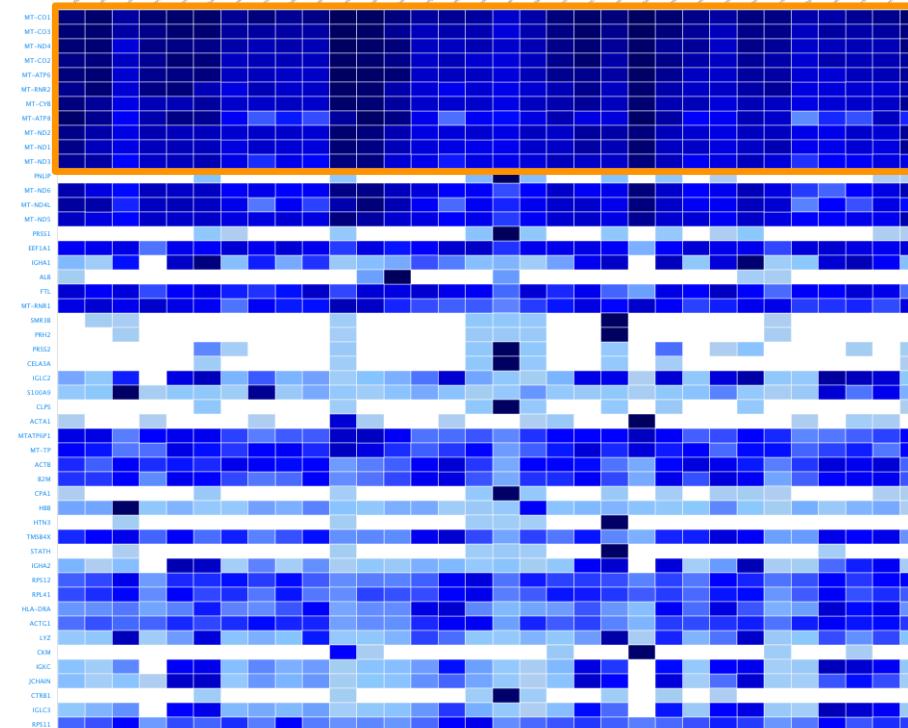
Select

Selected: 32 / 32



Showing 50 of 34,068 genes found:

Expression level in FPKM  
0      95.074



**HEATMAP:** top 50 genes (rows)  
most highly expressed across all  
tissues studied (columns)

Genes with highest  
expression first



# Baseline experiment page

## General information

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues

### RNA-Seq mRNA baseline

Organism: *Homo sapiens*

Raw Data Provider: The Human Protein Atlas

Results

Experiment Design

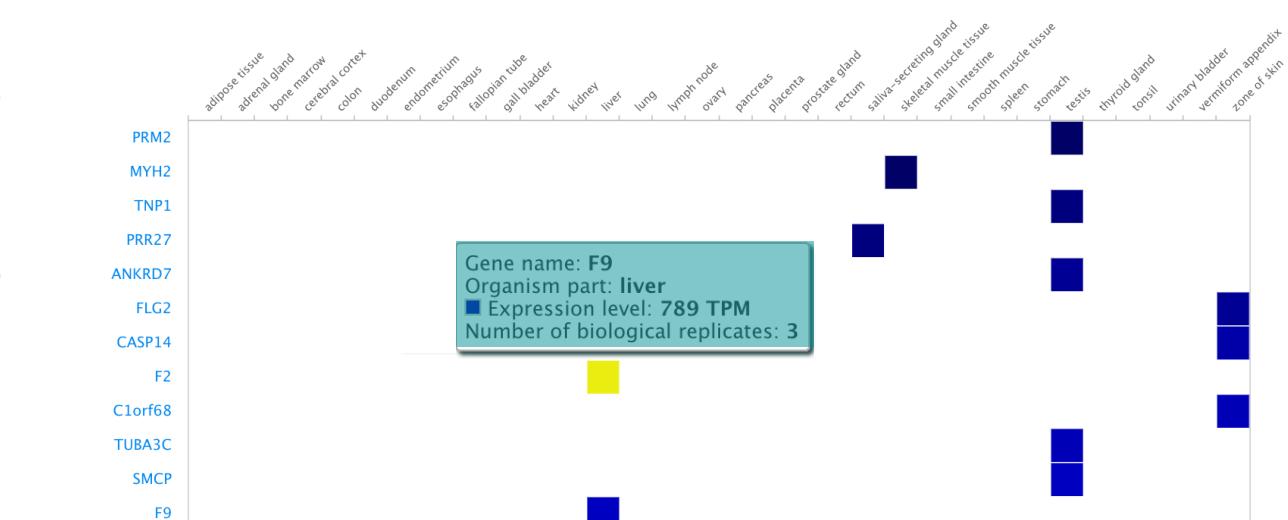
Supplementary Information

Downloads

### Genes expressed above 0.5 TPM

Showing 50 of 41,210 genes found:

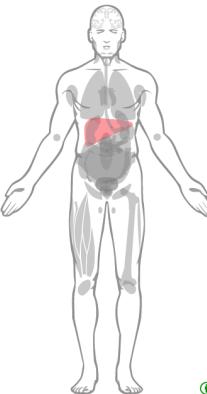
Expression level in TPM  
0 1,112



Download displayed results

[Download table content](#)

## Genes



Apply

Clear

Most specific

## Expression value

0.5

[See distribution](#)

## Data units

TPM  
 FPKM

## Organism parts

[Select](#)

Selected: 32 / 32





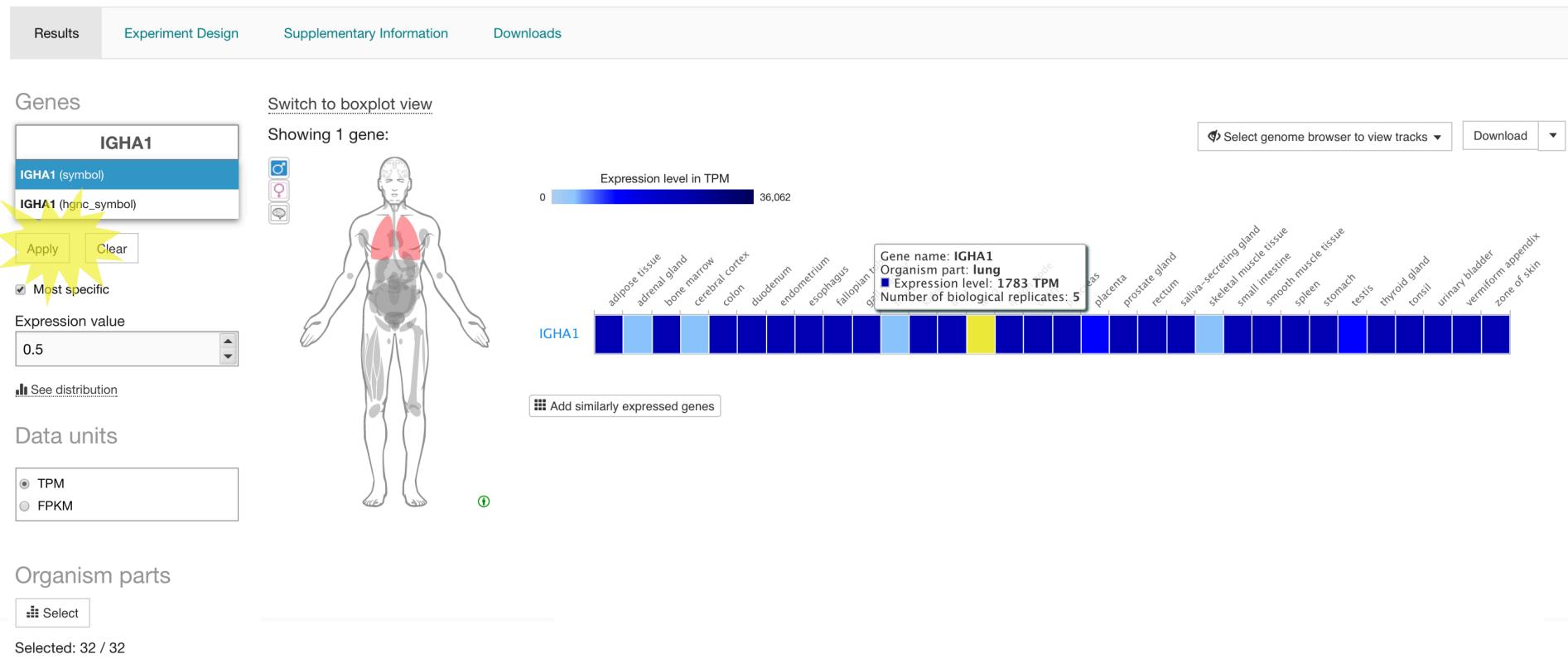
# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues

## RNA-Seq mRNA baseline

Organism: *Homo sapiens*

Raw Data Provider: The Human Protein Atlas





# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues

## RNA-Seq mRNA baseline

Organism: *Homo sapiens*

Raw Data Provider: The Human Protein Atlas



### Genes

**IGHA1**

IGHA1 (symbol)

IGHA1 (hgnc\_symbol)

Apply  

Switch to boxplot view

Showing 1 gene:



Apply

Clear

Most specific

Expression value

0.5

See distribution

Data units

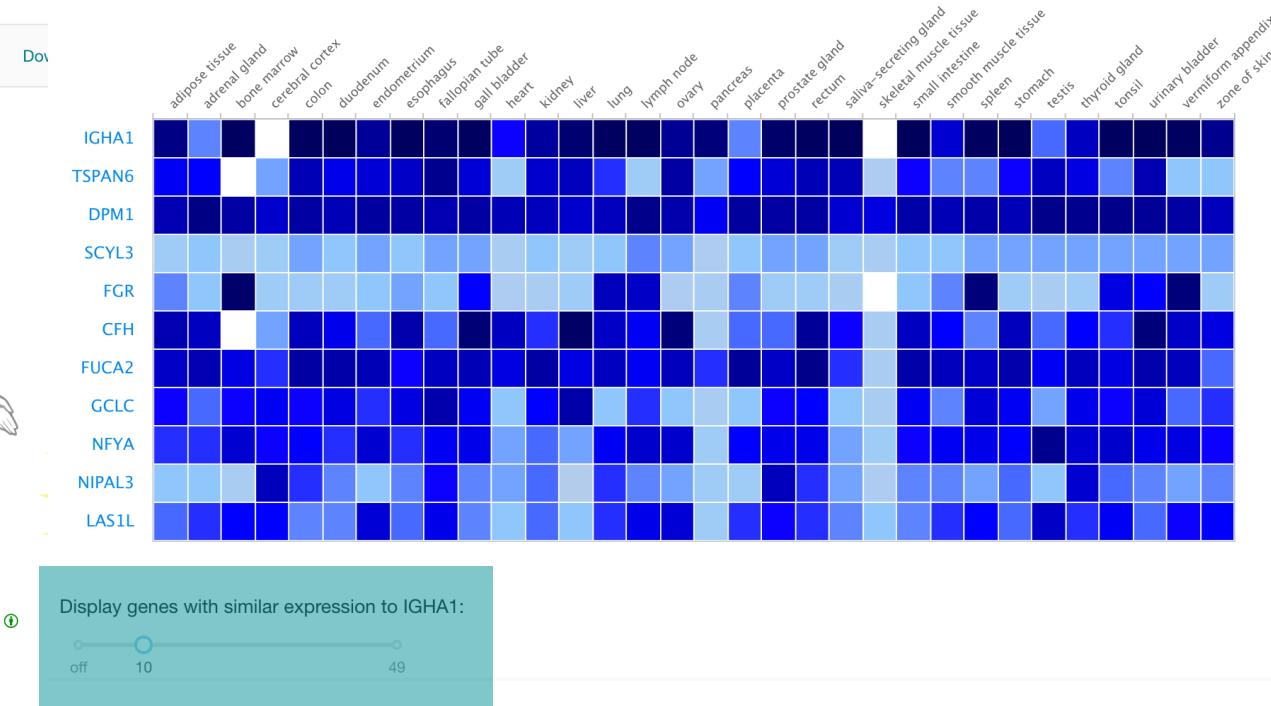
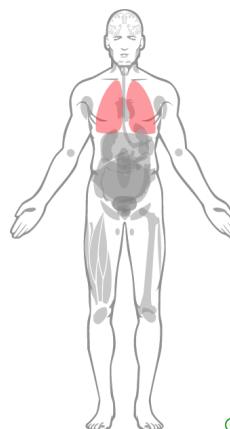
TPM

FPKM

Organism parts

Select

Selected: 32 / 32





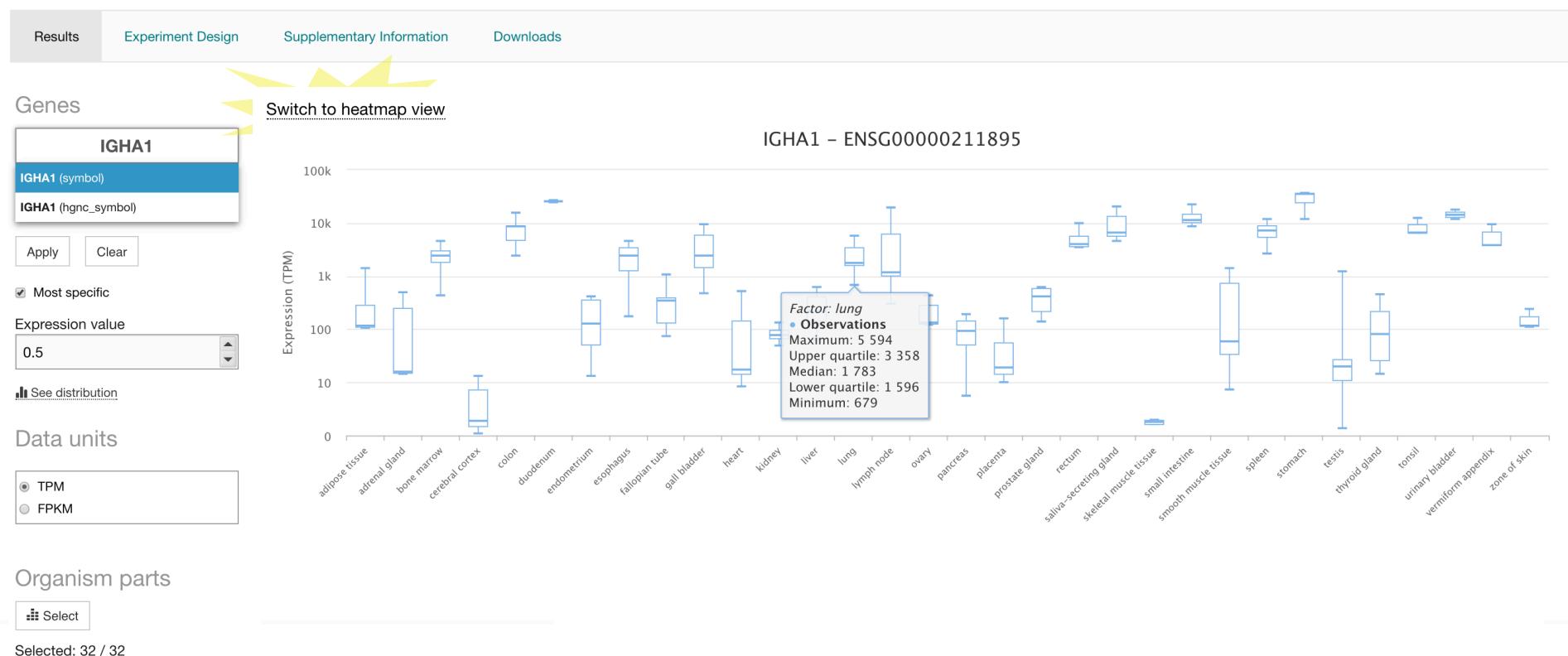
# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues

## RNA-Seq mRNA baseline

Organism: *Homo sapiens*

Raw Data Provider: The Human Protein Atlas



## Organism parts

Select

Selected: 32 / 32



# Baseline experiment page

Location: 14:105,707,118-105,708,665 Gene: IGAH1 Transcript: IGAH1-001

Human (GRCh38.p10) ▾

Location-based displays

- Whole genome
- Chromosome summary
- Region overview
- Region in detail**
- Comparative Genomics
- Synteny
- Alignments (image)
- Alignments (text)
- Region Comparison
- Genetic Variation
- Resequencing
- Linkage Data
- Markers
- Other genome browsers
- UCSC
- NCBI
- Vega
- Ensembl GRCh37

Configure this page

Custom tracks

Export data

Share this page

Bookmark this page

BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Bio | Mirrors

Search all species...

Assembly exceptions

Chr 14: 105,707,118-105,708,665

Region in detail

Chromosome bands

Contigs

Genes (Comprehensive set from GENCODE 26)

Gene Legend

- Ensembl protein coding
- IG C gene
- IG D gene
- processed transcript
- RNA gene

Location 14:105707118-105708665 Gene: IGAH1 Go

Assembly exceptions

Chromosome bands

Contigs

Genes (Comprehensive set from GENCODE 26)

Human cDNAs (RefSeq/ENA)

IGAH1 expression level (FPKM) in the context of the genomic location of the gene



# Baseline experiment page

RNA-seq of coding RNA from tissue samples of 122 human individuals representing 32 different tissues

RNA-Seq mRNA baseline

Organism: *Homo sapiens*

Raw Data Provider: The Human Protein Atlas

Results

Experiment Design

Supplementary Information

Downloads

Expression values across all genes (TPM)

Expression values across all genes (FPKM)

Summary of the expression results for this experiment ready to view in R

- Heatmap of aggregated expression data

Experiment Design (tsv)



# Let's try Expression Atlas

Baseline expression

## Hands-on activity

Look up expression of *TNNI3* gene in  
normal human tissues



In pairs



# Let's try Expression Atlas

The screenshot shows the Expression Atlas homepage. At the top left is the logo and the title "Expression Atlas". Below it is the subtitle "Gene expression across species and biological conditions". The top right features a search bar with the placeholder "Enter gene query..." and a "Search" button. Below the search bar are examples of queries: "ASPM, Apoptosis, ENSMUSG00000021789, zinc finger". The main content area has a heading "Exploring gene expression results across species under different biological conditions" with a subtext explaining the resource's purpose. A large orange box contains the question "In which conditions is my favourite gene differentially expressed?". An orange arrow points from this box to the "Organism" dropdown menu, which lists various species including Homo sapiens, Mus musculus, Arabidopsis thaliana, Rattus norvegicus, Drosophila melanogaster, and Zea mays. The "Any" option is selected. To the left, there are sections for "Gene / Gene properties" (showing CFHR2 details) and "Gene set enrichment". On the right, there is a "Biological conditions" search bar with examples like "lung, leaf, valproic acid, cancer". A large teal box at the bottom right contains the text "Discover and interpret gene expression analysis results quickly and easy".

# Discover and interpret gene expression analysis results quickly and easy



# Differential expression results

Results for CFHR2 (symbol)

Baseline expression Differential expression



Display log<sub>2</sub>-fold change

Filter your results

Experimental variables

- Growth condition
- Time
- Genotype
- Treatment
- Cell type
- Diet
- Disease
- Sampling time
- Infect
- Phenotype

Log <sub>2</sub> -fold	Species	Gene	Comparison	Experimental variables	Experiment name
Adjusted p-value	Log <sub>2</sub> -fold change				
1.575 × 10 <sup>-17</sup>	6.5		'lung carcinoma' vs 'normal' in 'CD11b-CD11c- epithelial cell'	cell type, disease	Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq
		CFH	Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq 'lung carcinoma' vs 'normal' in 'lung cell'		
		CFH	Property Test value (N=3) Reference value (N=3)	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFH	cell type lung cell lung cell	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFH	disease lung carcinoma normal	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFH	organism Mus musculus Mus musculus	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFH	organism part lung lung	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFHR2	strain C57BL/6 C57BL/6	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFHR2	'pancreatic adenocarcinoma' vs 'normal'	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFHR2	'non-small cell lung carcinoma' vs 'normal'	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		CFHR2	'hepatobiliary carcinoma' vs 'normal'	block, disease	RNA-seq of blood platelets from six tumor types and healthy donors
		Cfhr2	'lung carcinoma' vs 'normal' in 'lung cell'	cell type, disease	Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq

All organisms and all conditions in which *CFHR2* gene is differentially expressed



# Differential experiment page

## General information

Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq

RNA-Seq mRNA differential

Organism: *Mus musculus*

Reference(s): 25704820 (Filter by genes in paper)

Results

Plots

Experiment Design

Supplementary Information

Downloads

## Differentially expressed genes

Showing 50 of 7,557 genes found:



Genes

Download displayed results

Igkv14-126  
Ighv1-64  
Muc5ac  
Gm43745  
Aldh1a3  
Dio3  
Otc  
BC016579  
D930028M14Rik  
Hoxb6  
Dlk1  
Twist1

'lung carcinoma' vs 'normal' in 'CD11b+Ly6G+ neutrophil'

'lung carcinoma' vs 'normal' in 'Epcam+CD11b-CD11c- epithelial cell'

'lung carcinoma' vs 'normal' in 'lung cell'

Gene name: BC016579  
Disease: lung carcinoma vs normal  
■ Log<sub>2</sub>-fold change: 5.4  
Adjusted p-value:  $4.5153 \times 10^{-11}$   
Click on the cell to show expression in the Genome Browser

## Criteria to select DE genes

Log<sub>2</sub>-fold change

Adjusted p-value

## Comparisons

Selected: 3 / 3



# Differential experiment page

Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq

RNA-Seq mRNA differential

Organism: *Mus musculus*

Reference(s): 25704820 (Filter by genes in paper)

Results

Plots

Experiment Design

Supplementary Information

Downloads

Genes

Showing 50 of 3,792 genes found:

[Download table content](#)

Log<sub>2</sub>-fold change

-5.5 0



Specific unchecked

Most specific

Up- or downregulated  
 Upregulated only  
 Downregulated only

Log<sub>2</sub>-fold change

1.0

Adjusted p-value

0.05

Comparisons

Select

Selected: 3 / 3



# Differential experiment page

Transcriptome analysis of isolated stromal cells and tumor epithelial cells in mouse lung cancer by RNA-Seq

RNA-Seq mRNA differential

Organism: *Mus musculus*

Reference(s): 25704820 (Filter by genes in paper)

Results

Plots

Experiment Design

Supplementary Information

Downloads

'lung carcinoma' vs 'normal' in 'CD11b+Ly6G+ neutrophil'

GO terms enrichment analysis

Interpro domains enrichment analysis

Reactome pathways enrichment analysis

MA plot for the contrast across all genes

'lung carcinoma' vs 'normal' in 'Epcam+CD11b-CD11c- epithelial cell'

GO terms enrichment analysis

Interpro domains enrichment analysis

Reactome pathways enrichment analysis

MA plot for the contrast across all genes

'lung carcinoma' vs 'normal' in 'lung cell'

GO terms enrichment analysis

Interpro domains enrichment analysis

Reactome pathways enrichment analysis

MA plot for the contrast across all genes



piano

Platform for integrative analysis of omics data

Bioconductor version: Release (3.5)

Piano performs gene set analysis using various statistical methods, from different gene level statistics and a wide range of gene-set collections. Furthermore, the Piano package contains functions for combining the results of multiple runs of gene set analyses.

Author: Leif Varemo <piano.rpkg@gmail.com> and Intawat Nookaew <piano.rpkg@gmail.com>

Maintainer: Leif Varemo <piano.rpkg@gmail.com>

Citation (from within R, enter `citation("piano")`):

Varemo L, Nielsen J and Nookaew I (2013). "Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods." *Nucleic Acids Research*, **41**(8), pp. 4378-4391. doi: [10.1093/nar/gkt111](https://doi.org/10.1093/nar/gkt111).



# Differential experiment page

Transcriptome analysis of isolated stromal cells and tumor

RNA-Seq mRNA differential

Organism: *Mus musculus*

Reference(s): 25704820 (Filter by genes in paper)

Results Plots Experiment Design Supplementary Information Downloads

'lung carcinoma' vs 'normal' in 'CD11b+Ly6G+ neutrophil'

GSE GO terms enrichment analysis

GSE Interpro domains enrichment analysis

GSE Reactome pathways enrichment analysis

MA MA plot for the contrast across all genes

'lung carcinoma' vs 'normal' in 'Epcam+CD11b-CD11c- epithelial cell'

GSE GO terms enrichment analysis

GSE Interpro domains enrichment analysis

GSE Reactome pathways enrichment analysis

MA MA plot for the contrast across all genes

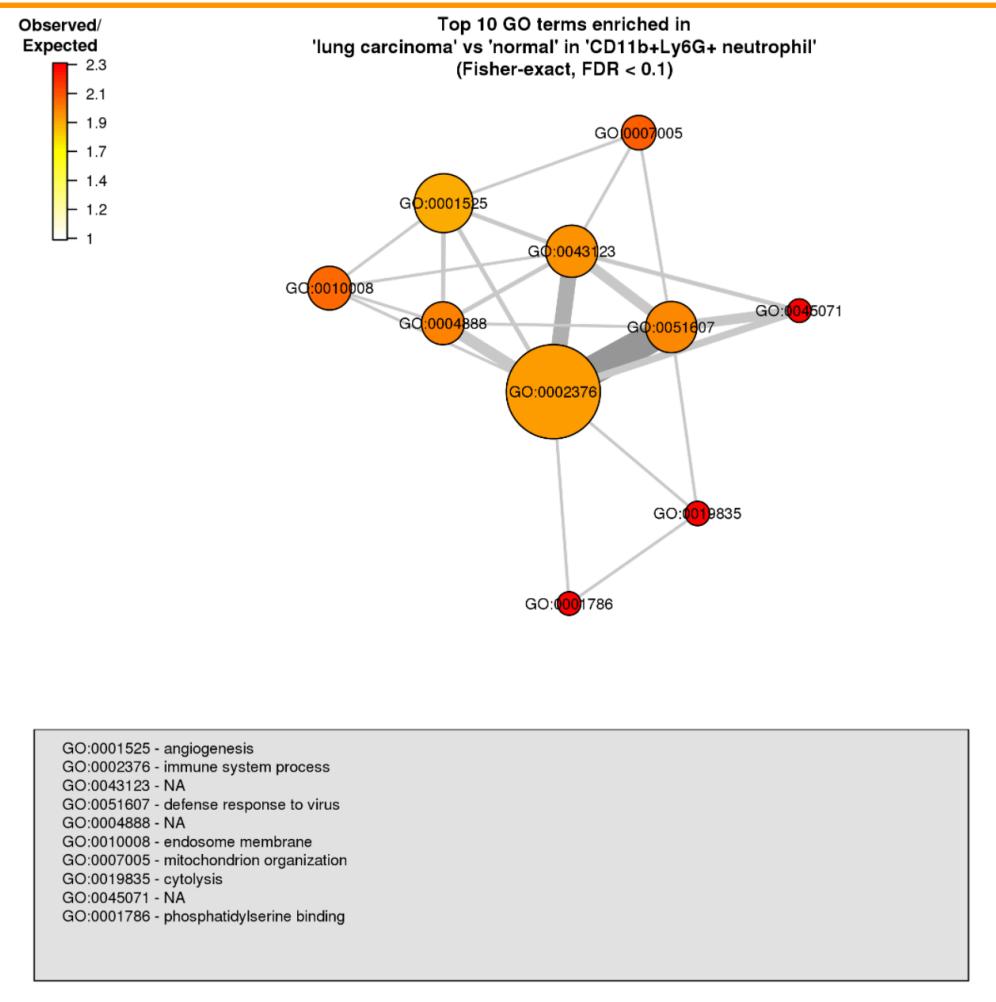
'lung carcinoma' vs 'normal' in 'lung cell'

GSE GO terms enrichment analysis

GSE Interpro domains enrichment analysis

GSE Reactome pathways enrichment analysis

MA MA plot for the contrast across all genes





# More information about *CFHR2* gene

Login/Register

**e!Ensembl** BLAST/BLAT | BioMart | Tools | Downloads | Help & Documentation | Blog | Mirrors

Human (GRCh38.p10) ▾

Location: 1:196,943,772-196,959,226

Gene-based displays

- Summary
- Splice variants
- Transcript comparison
- Gene alleles

- Sequence
  - Secondary Structure

- Comparative Genomics
  - Genomic alignments
  - Gene tree
  - Gene gain/loss tree
  - Orthologues
  - Paralogues
  - Ensembl protein families

- Ontologies
  - GO: Cellular component
  - GO: Molecular function
  - GO: Biological process

- Phenotypes

- Genetic Variation
  - Variant table
  - Variant image
  - Structural variants

- Gene expression
- Regulation
- External references
- Supporting evidence

Gene: CFHR2 ENSG00000080910

Description complement factor H related 2 [Source:HGNC Symbol;Acc:[HGNC:4890](#)]

Synonyms FHR2, HFL3, CFHL2

Location Chromosome 1: 196,943,772-196,959,226 forward strand.  
GRCh38:CM000663.2

About this gene This gene has 6 transcripts (splice variants), 16 orthologues, 11 paralogues, is a member of [1 Ensembl protein family](#) and is associated with [3 phenotypes](#).

Transcripts

[Hide transcript table](#)

Show/hide columns (1 hidden)

Name	Transcript ID	bp	Protein	Biotype	CCDS	UniProt	RefSeq	Flags
CFHR2-001	<a href="#">ENST00000367415.6</a>	1059	270aa	Protein coding	<a href="#">CCDS30959</a>	<a href="#">P36980</a>	NM_005666 NP_005657	TSL:1 GENCODE basic APPRIS P2
CFHR2-005	<a href="#">ENST00000476712.6</a>	1020	254aa	Protein coding	-	<a href="#">V9GYE7</a>	-	TSL:3 GENCODE basic APPRIS ALT2
CFHR2-006	<a href="#">ENST00000496448.5</a>	811	No protein	Processed transcript	-	-	-	TSL:2
CFHR2-002	<a href="#">ENST00000473386.1</a>	599	No protein	Processed transcript	-	-	<a href="#">NM_001312672</a>	TSL:1
CFHR2-004	<a href="#">ENST00000489703.1</a>	424	No protein	Processed transcript	-	-	-	TSL:3
CFHR2-003	<a href="#">ENST00000485647.1</a>	376	No protein	Processed transcript	-	-	-	TSL:5



# Let's try Expression Atlas

Differential expression

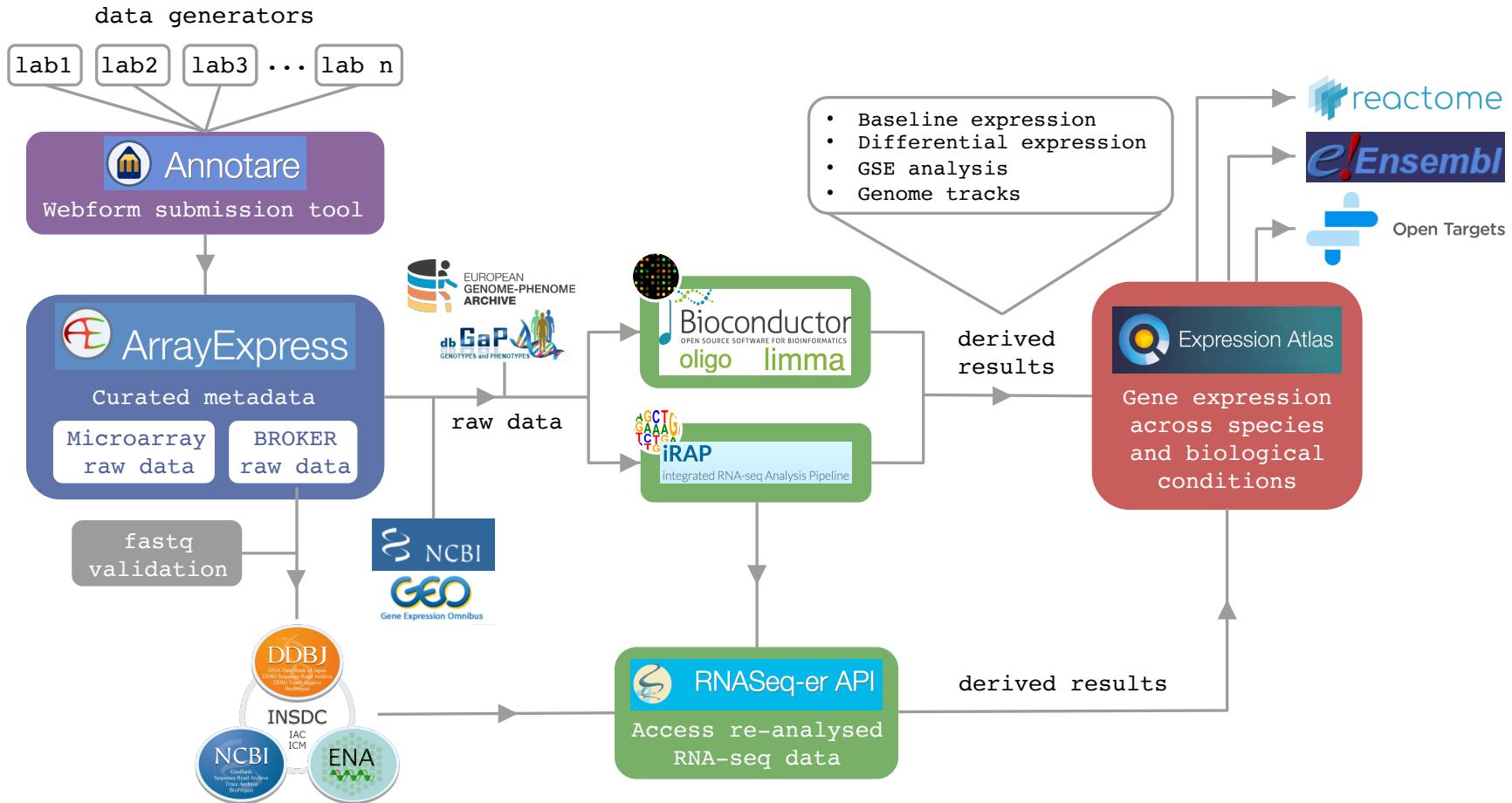
## Hands-on activity

Find differentially expressed genes in patients with hepatocellular carcinoma



In pairs

# Functional genomics resources at EMBL-EBI





# RNASeq-er API



## RNASeq-er API

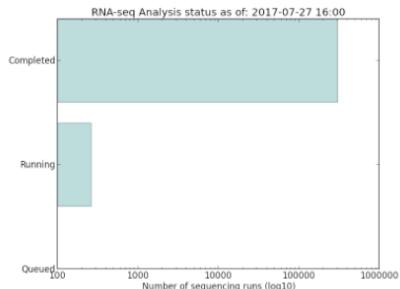
[Home](#) | [Documentation](#)

[Feedback](#)

### Welcome to RNASeq-er API - a gateway to systematically updated analysis of public RNA-Seq data

The RNASeq-er REST API provides easy access to the results of the systematically updated and continually growing analysis of public RNA-seq data in [European Nucleotide Archive \(ENA\)](#). The analysis of each sequencing run is performed by the EMBL-EBI's Gene Expression Team using the [iRAP pipeline](#).

- [1. Get Started](#)
- [2. What does the RNASeq-er pipeline do?](#)
- [3. How is the RNASeq-er performed?](#)
- [4. How to use the RNASeq-er API?](#)
- [5. What are the main classes of API calls?](#)
  - [5.1. Analysis results Per Run](#)
    - [5.1.1. Making per-run API calls](#)
    - [5.1.2. Results of per-run API calls](#)
  - [5.2. Analysis results Per Study](#)
    - [5.2.1. Making per-study API calls](#)
    - [5.2.2. Results of per-study API calls](#)
  - [5.3. Sample Attributes Per Run](#)
    - [5.3.1. Making sample attributes per-run API calls](#)
    - [5.3.2. Results of sample attributes per-run API calls](#)
  - [5.4. Baseline expression By Gene](#)
    - [5.4.1. Making baseline expression per-gene API calls](#)
    - [5.4.2. Results of baseline expression per-gene API calls](#)
  - [5.5. Mapping Quality Statistics For All Organisms](#)
    - [5.5.1. Retrieving the mean and standard deviation of mapping quality for all organisms](#)



<https://www.ebi.ac.uk/fg/rnaseq/>

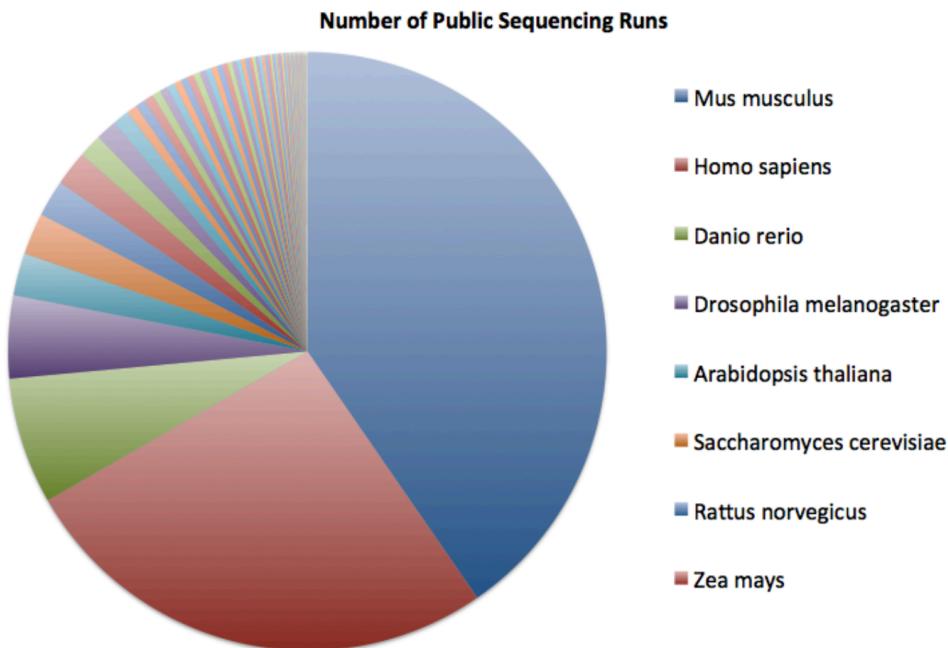


# RNASeq-er API

> 270,000 analysed runs

> 15,000 studies

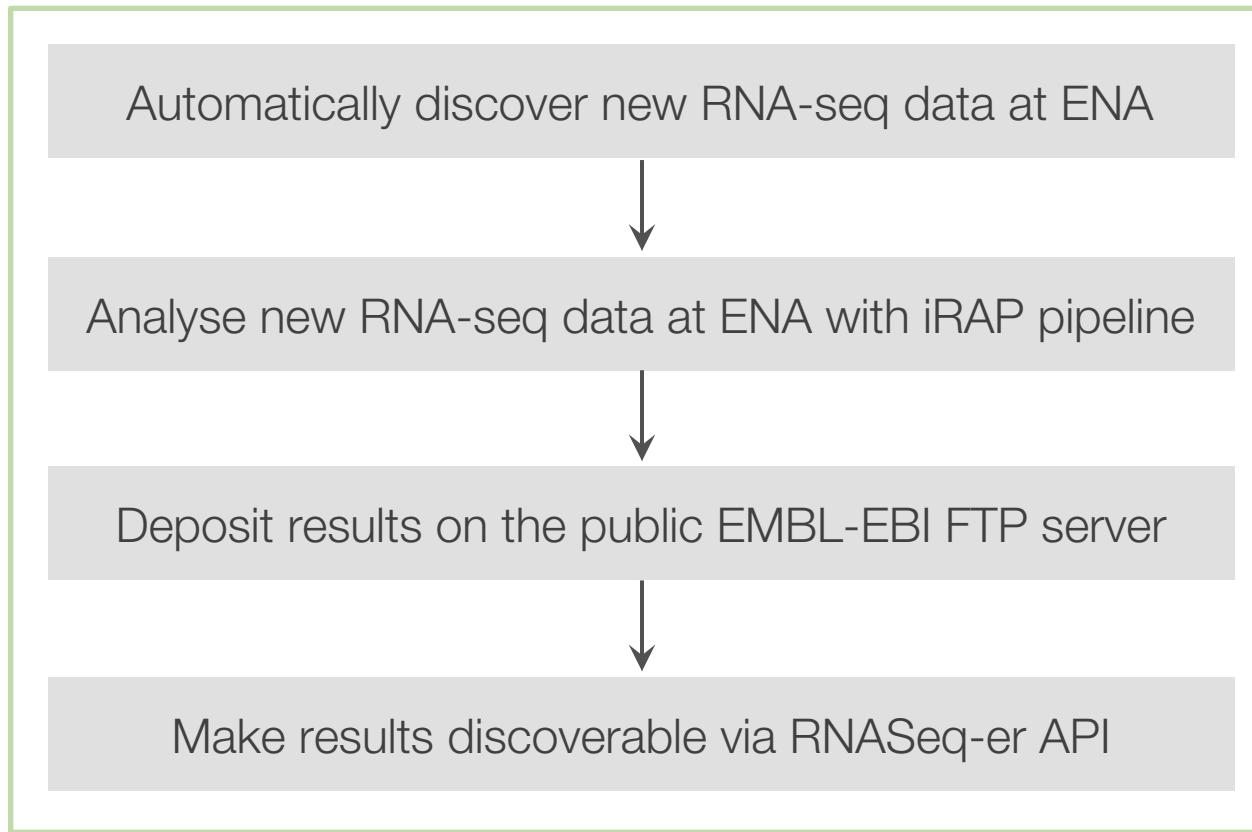
> 300 species (163 genome references)



*800,000 runs not analysed from 6,500 species, not available via API (no reference genome/gene annotation available in public databases)*



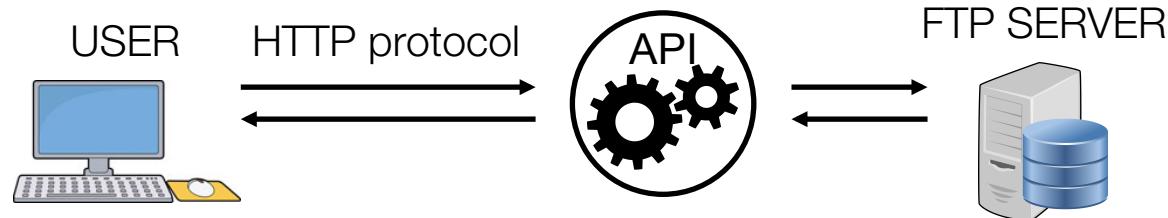
# RNASeq-er API



<https://doi.org/10.1093/bioinformatics/btx143>

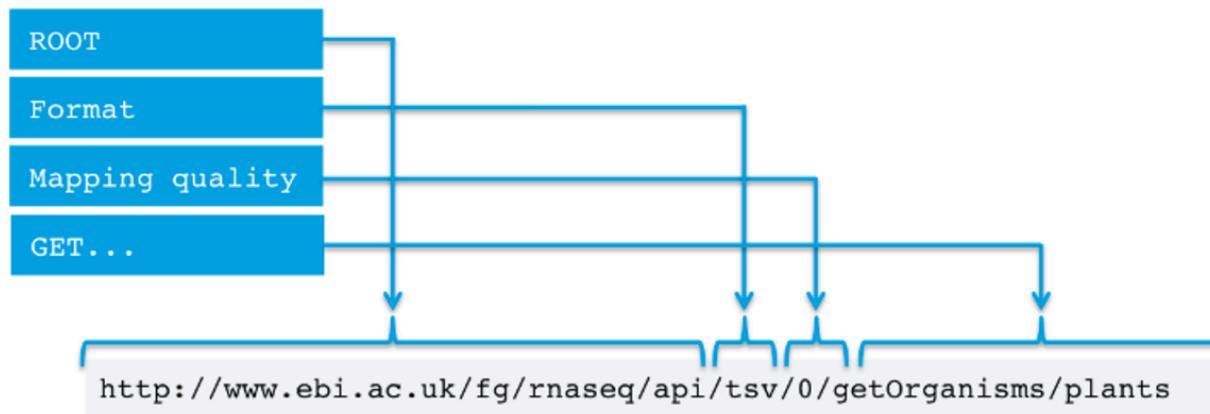


# RNASeq-er API



How to built API calls?

How to built the URL?





# RNASeq-er API

## Accessing the results of alignments

*Access alignments of RNA-seq data involving human autism samples*

[https://www.ebi.ac.uk/fg/rnaseq/api/tsv/0/getRunsByOrganismCondition/homo\\_sapiens/autism](https://www.ebi.ac.uk/fg/rnaseq/api/tsv/0/getRunsByOrganismCondition/homo_sapiens/autism)

Secure | https://www.ebi.ac.uk/fg/rnaseq/api/tsv/0/getRunsByOrganismCondition/homo\_sapiens/autism

STUDY_ID	SAMPLE_IDS	BIOREP_ID	RUN_IDS	ORGANISM	REFERENCE_ORGANISM	STATUS	ASSEMBLY_USED	ENA_LAST_UPDATED	LAST_PROCESSED_DATE	CRAM_LOCATION	BEDGRAPH_LOCATION
SRP007483	SAMN00668205	SRR309133	SRR309133	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sat Aug 22 2015 10:42:42		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.bw	96			disease autism
SRP007483	SAMN00668205	SRR309134	SRR309134	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Aug 10 2015 23:07:36		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.bw	96			disease autism
SRP007483	SAMN00668206	SRR309135	SRR309135	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Aug 24 2015 23:30:14		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.bw	95			disease autism
SRP007483	SAMN00668206	SRR309136	SRR309136	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Jul 27 2015 07:34:38		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.bw	95			disease autism
SRP007483	SAMN00668207	SRR309137	SRR309137	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Tue Aug 11 2015 22:40:15		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.bw	94			disease autism
SRP007483	SAMN00668207	SRR309138	SRR309138	homo_sapiens	homo_sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sat Aug 15 2015 13:44:31		
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.cram				ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.bedgraph
							ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.bw	95			disease autism

*alignment results in EMBL-EBI FTP server*



# RNASeq-er API

## Accessing the results of alignments

*Access alignment results of all samples of a RNA-seq study*

<https://www.ebi.ac.uk/fg/rnaseq/api/tsv/0/getRunsByStudy/SRP007483>

Secure | <https://www.ebi.ac.uk/fg/rnaseq/api/tsv/0/getRunsByStudy/SRP007483>

STUDY_ID	SAMPLE_IDS	BIOREP_ID	RUN_IDS	ORGANISM	REFERENCE_ORGANISM	STATUS	ASSEMBLY_USED	ENA_LAST_UPDATED	LAST_PROCESSED_DATE	CRAM_LOCATION	BEDGRAPH_LOCATION
<b>BIGWIG LOCATION MAPPING_QUALITY</b>											
SRP007483	SAMN00668205	SRR309133	SRR309133	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sat Aug 22 2015 10:42:42		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309133/SRR309133.bw	96
SRP007483	SAMN00668205	SRR309134	SRR309134	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Aug 10 2015 23:07:36		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309134/SRR309134.bw	96
SRP007483	SAMN00668206	SRR309135	SRR309135	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Aug 24 2015 23:30:14		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309135/SRR309135.bw	95
SRP007483	SAMN00668206	SRR309136	SRR309136	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Jul 27 2015 07:34:38		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309136/SRR309136.bw	95
SRP007483	SAMN00668207	SRR309137	SRR309137	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Tue Aug 11 2015 22:40:15		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309137/SRR309137.bw	94
SRP007483	SAMN00668207	SRR309138	SRR309138	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sat Aug 15 2015 13:44:31		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309138/SRR309138.bw	95
SRP007483	SAMN00668208	SRR309139	SRR309139	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Thu Aug 13 2015 04:56:53		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309139/SRR309139.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309139/SRR309139.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309139/SRR309139.bw	94
SRP007483	SAMN00668208	SRR309140	SRR309140	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Aug 24 2015 21:30:41		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309140/SRR309140.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309140/SRR309140.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309140/SRR309140.bw	94
SRP007483	SAMN00668209	SRR309141	SRR309141	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Mon Jul 27 2015 00:06:37		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309141/SRR309141.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309141/SRR309141.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309141/SRR309141.bw	96
SRP007483	SAMN00668209	SRR309142	SRR309142	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sun Aug 16 2015 03:50:59		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309142/SRR309142.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309142/SRR309142.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309142/SRR309142.bw	96
SRP007483	SAMN00668210	SRR309143	SRR309143	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Sun Aug 23 2015 10:36:16		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309143/SRR309143.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309143/SRR309143.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309143/SRR309143.bw	96
SRP007483	SAMN00668210	SRR309144	SRR309144	homo sapiens	homo sapiens	Complete	GRCh38	Fri Jun 19 2015 20:17:06	Fri Jul 24 2015 20:40:35		
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309144/SRR309144.cram	ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309144/SRR309144.bedgraph
										ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/SRR309/SRR309144/SRR309144.bw	96



# RNASeq-er API

## Accessing sample metadata

<https://www.ebi.ac.uk/fg/rnaseq/api/tsv/getSampleAttributesPerRunByStudy/SRP007483>

STUDY_ID	RUN_ID	TYPE	VALUE_EFO_URL
SRP007483	SRR309133		clinical information brain region BA22 (temporal cortex) NA
SRP007483	SRR309133		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309133		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309133		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309133		clinical information brain region BA22 (temporal cortex) NA
SRP007483	SRR309134		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309134		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309134		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309134		clinical information brain region BA41 (temporal cortex) NA
SRP007483	SRR309134		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309134		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309134		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309137		clinical information brain region BA41 (temporal cortex) NA
SRP007483	SRR309137		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309137		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309137		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309138		clinical information brain region BA41 (temporal cortex) NA
SRP007483	SRR309138		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309138		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309138		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309139		clinical information brain region BA09 (frontal cortex) NA
SRP007483	SRR309139		disease autism http://www.ebi.ac.uk/efo/EFO_0003758
SRP007483	SRR309139		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309139		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309140		clinical information brain region BA09 (frontal cortex) NA
SRP007483	SRR309140		disease normal http://purl.obolibrary.org/obo/PATO_0000461
SRP007483	SRR309140		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309140		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309144		clinical information brain region BA41 (temporal cortex) NA
SRP007483	SRR309144		disease normal http://purl.obolibrary.org/obo/PATO_0000461
SRP007483	SRR309144		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309144		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955
SRP007483	SRR309144		clinical information brain region BA41 (temporal cortex) NA
SRP007483	SRR309144		disease normal http://purl.obolibrary.org/obo/PATO_0000461
SRP007483	SRR309144		organism Homo sapiens http://purl.obolibrary.org/obo/NCBITaxon_9606
SRP007483	SRR309144		organism part brain http://purl.obolibrary.org/obo/UBERON_0000955

samples from  
patients with autism

samples from  
normal individuals



# RNASeq-er API

## Accessing gene expression quantification results

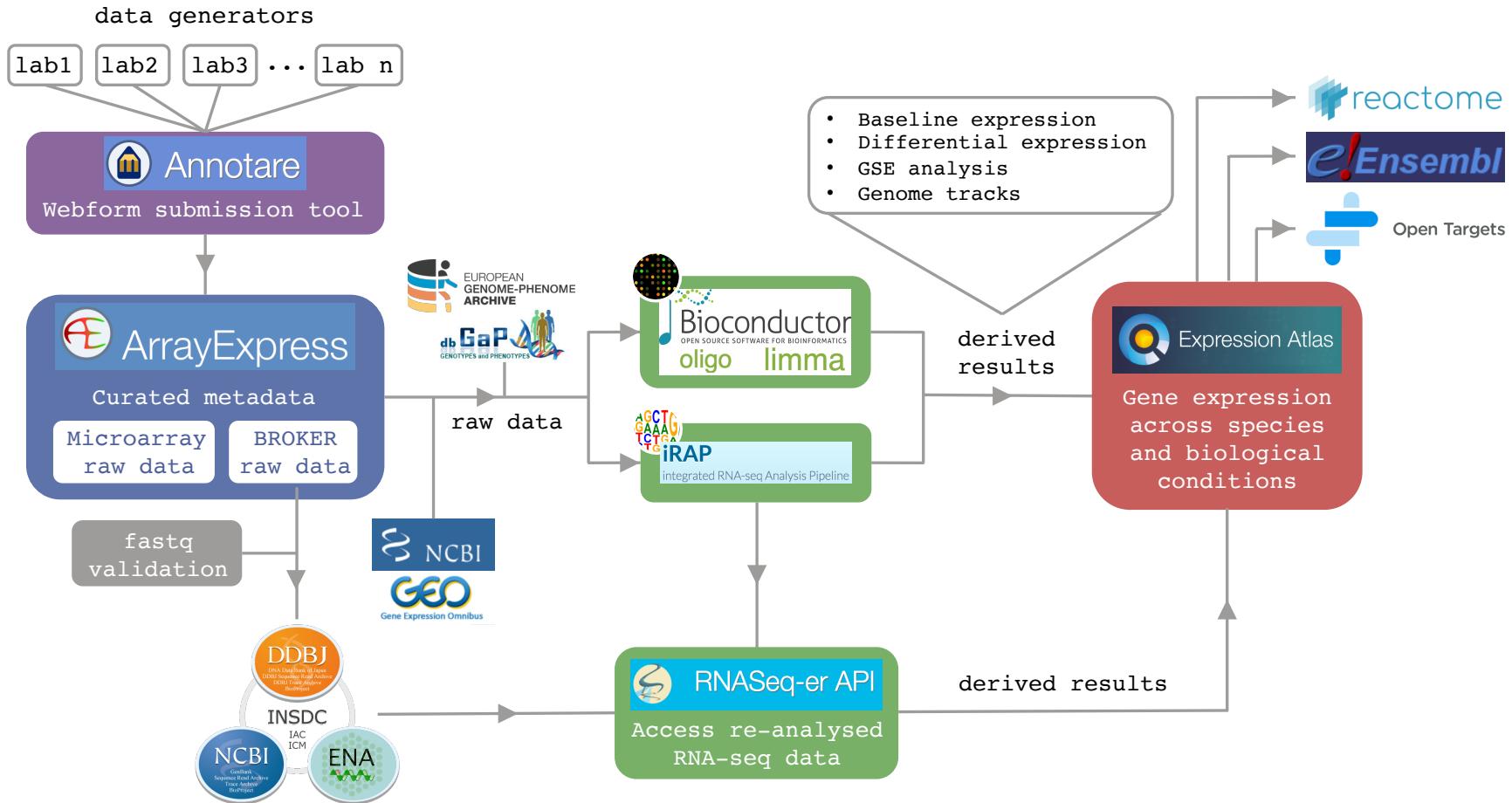
*Access gene expression quantification results of a RNA-seq study*

<https://www.ebi.ac.uk/fg/rnaseq/api/tsv/getStudy/SRP007483>

Secure | https://www.ebi.ac.uk/fg/rnaseq/api/tsv/getStudy/SRP007483

STUDY_ID	ORGANISM	REFERENCE_ORGANISM	ASSEMBLY_USED	GTF_USED	STATUS	GENES_FPKM_COUNTS_FTP_LOCATION	GENES TPM COUNTS FTP LOCATION	GENES_RAW_COUNTS_FTP_LOCATION	EXONS_FPKM_COUNTS_FTP_LOCATION																																																																																																																																								
SRP007483	homo sapiens	homo sapiens	GRCh38	Homo sapiens.GRCh38.79.gtf.gz	Complete																																																																																																																																												
<pre>ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/studies/ena/SRP007483/homo_sapiens/genes.fpkm.htseq2.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/genes.tpm.htseq2.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/genes.raw.htseq2.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/exons.fpkm.dexseq.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/exons.tpm.dexseq.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/exons.raw.dexseq.tsv ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/studies/ena/SRP007483/homo_sapiens/irap VERSIONS.tsv</pre> Thu Jun 30 2016 21:43:57																																																																																																																																																	
<i>gene expression quantification matrices</i>																																																																																																																																																	
<a href="ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/studies/ena/SRP007483/homo_sapiens/genes.tpm.htseq2.tsv">ftp://ftp.ebi.ac.uk/pub/databases/arrayexpress/data/atlas/rnaseq/studies/ena/SRP007483/homo_sapiens/genes.tpm.htseq2.tsv</a>																																																																																																																																																	
<table><thead><tr><th>Gene ID</th><th>SRR309133</th><th>SRR309134</th><th>SRR309135</th><th>SRR309136</th><th>SRR309137</th><th>SRR309138</th><th>SRR309139</th></tr><tr><th>SRR309140</th><th>SRR309141</th><th>SRR309142</th><th>SRR309143</th><th>SRR309144</th><th></th><th></th><th></th></tr></thead><tbody><tr><td>ENSG00000000003</td><td>3.13</td><td>3.04</td><td>3.39</td><td>3.33</td><td>5.97</td><td>7.32</td><td>3.35</td><td>3.75</td><td>3.25</td></tr><tr><td>ENSG00000000005</td><td>0</td><td>0</td><td>0.06</td><td>0.12</td><td>0.25</td><td>0.18</td><td>0.43</td><td>0.29</td><td>0.27</td></tr><tr><td>ENSG00000000419</td><td>7.47</td><td>7.95</td><td>16.86</td><td>15.17</td><td>21.79</td><td>22.01</td><td>24.15</td><td>24.09</td><td>42.88</td></tr><tr><td>ENSG00000000457</td><td>0.67</td><td>0.73</td><td>0.55</td><td>0.48</td><td>2.7</td><td>3.24</td><td>3.2</td><td>2.15</td><td>3.97</td></tr><tr><td>ENSG00000000460</td><td>0.21</td><td>0.32</td><td>0.1</td><td>0.21</td><td>1.06</td><td>1.14</td><td>0.66</td><td>0.52</td><td>0.53</td></tr><tr><td>ENSG00000000938</td><td>0.78</td><td>0.55</td><td>1.44</td><td>1.73</td><td>8.71</td><td>9.22</td><td>3.48</td><td>3.74</td><td>2.5</td></tr><tr><td>ENSG00000000971</td><td>0.88</td><td>1.09</td><td>6.58</td><td>6.57</td><td>5.12</td><td>4.69</td><td>1.85</td><td>1.76</td><td>1.28</td></tr><tr><td>ENSG00000001036</td><td>4.93</td><td>5.5</td><td>5.5</td><td>5.31</td><td>14.87</td><td>15.48</td><td>9.32</td><td>8.59</td><td>9.99</td></tr><tr><td>ENSG00000001084</td><td>5.49</td><td>5.32</td><td>5.74</td><td>6.07</td><td>11.03</td><td>10.79</td><td>8.32</td><td>7.28</td><td>8.88</td></tr><tr><td>ENSG00000001167</td><td>5.31</td><td>5.69</td><td>6.83</td><td>7.37</td><td>11.83</td><td>12.78</td><td>8.28</td><td>8.19</td><td>8.08</td></tr><tr><td>ENSG00000001460</td><td>0.78</td><td>0.69</td><td>1.34</td><td>1.04</td><td>6.2</td><td>6.48</td><td>5.02</td><td>4.67</td><td>8.36</td></tr><tr><td>ENSG00000001461</td><td>14.64</td><td>14.75</td><td>17.54</td><td>17.8</td><td>47.25</td><td>46.31</td><td>18.4</td><td>17.97</td><td>39.94</td></tr></tbody></table>										Gene ID	SRR309133	SRR309134	SRR309135	SRR309136	SRR309137	SRR309138	SRR309139	SRR309140	SRR309141	SRR309142	SRR309143	SRR309144				ENSG00000000003	3.13	3.04	3.39	3.33	5.97	7.32	3.35	3.75	3.25	ENSG00000000005	0	0	0.06	0.12	0.25	0.18	0.43	0.29	0.27	ENSG00000000419	7.47	7.95	16.86	15.17	21.79	22.01	24.15	24.09	42.88	ENSG00000000457	0.67	0.73	0.55	0.48	2.7	3.24	3.2	2.15	3.97	ENSG00000000460	0.21	0.32	0.1	0.21	1.06	1.14	0.66	0.52	0.53	ENSG00000000938	0.78	0.55	1.44	1.73	8.71	9.22	3.48	3.74	2.5	ENSG00000000971	0.88	1.09	6.58	6.57	5.12	4.69	1.85	1.76	1.28	ENSG00000001036	4.93	5.5	5.5	5.31	14.87	15.48	9.32	8.59	9.99	ENSG00000001084	5.49	5.32	5.74	6.07	11.03	10.79	8.32	7.28	8.88	ENSG00000001167	5.31	5.69	6.83	7.37	11.83	12.78	8.28	8.19	8.08	ENSG00000001460	0.78	0.69	1.34	1.04	6.2	6.48	5.02	4.67	8.36	ENSG00000001461	14.64	14.75	17.54	17.8	47.25	46.31	18.4	17.97	39.94
Gene ID	SRR309133	SRR309134	SRR309135	SRR309136	SRR309137	SRR309138	SRR309139																																																																																																																																										
SRR309140	SRR309141	SRR309142	SRR309143	SRR309144																																																																																																																																													
ENSG00000000003	3.13	3.04	3.39	3.33	5.97	7.32	3.35	3.75	3.25																																																																																																																																								
ENSG00000000005	0	0	0.06	0.12	0.25	0.18	0.43	0.29	0.27																																																																																																																																								
ENSG00000000419	7.47	7.95	16.86	15.17	21.79	22.01	24.15	24.09	42.88																																																																																																																																								
ENSG00000000457	0.67	0.73	0.55	0.48	2.7	3.24	3.2	2.15	3.97																																																																																																																																								
ENSG00000000460	0.21	0.32	0.1	0.21	1.06	1.14	0.66	0.52	0.53																																																																																																																																								
ENSG00000000938	0.78	0.55	1.44	1.73	8.71	9.22	3.48	3.74	2.5																																																																																																																																								
ENSG00000000971	0.88	1.09	6.58	6.57	5.12	4.69	1.85	1.76	1.28																																																																																																																																								
ENSG00000001036	4.93	5.5	5.5	5.31	14.87	15.48	9.32	8.59	9.99																																																																																																																																								
ENSG00000001084	5.49	5.32	5.74	6.07	11.03	10.79	8.32	7.28	8.88																																																																																																																																								
ENSG00000001167	5.31	5.69	6.83	7.37	11.83	12.78	8.28	8.19	8.08																																																																																																																																								
ENSG00000001460	0.78	0.69	1.34	1.04	6.2	6.48	5.02	4.67	8.36																																																																																																																																								
ENSG00000001461	14.64	14.75	17.54	17.8	47.25	46.31	18.4	17.97	39.94																																																																																																																																								

# Functional genomics resources at EMBL-EBI





# Expression Atlas data in other resources



Open Targets

About Us

Projects

People

News

Jobs

Contact

Open Targets Platform ➔

## Welcome to Open Targets

For biomedical researchers who need to identify a biological target for a  
of therapeutic targets based on genome-scale experiments and analysis  
human diseases, and is committed to sharing its data openly with the sc

About Open Targets ➔



Copyright (c) 2016 Open Targets. All rights reserved.

Site maintained by the Web Development team at EMBL-EBI | Terms of Use | Privacy | Co



Open Targets Platform

## Find new targets for drug discovery

CYP1A2



### CYP1A2

cytochrome P450 family 1 subfamily A member 2

156 diseases associated

Target

Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics. Most active in catalyzing 2-hydrox...



# Expression Atlas data in other resources

156 diseases associated with CYP1A2

[View CYP1A2 profile](#)

Filter by

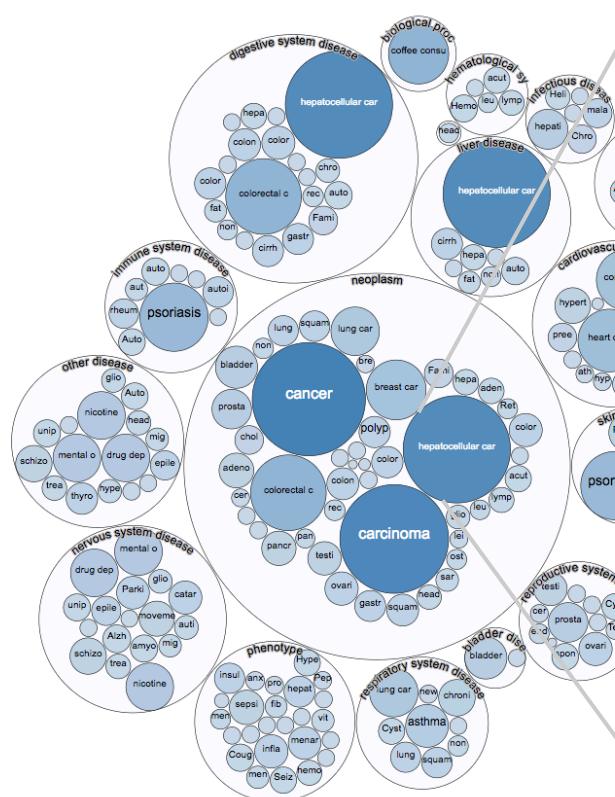
Bubbles Table Tree

Data types

- Genetic associations (4)
- Somatic mutations (0)
- Drugs (0)
- Affected pathways (0)
- RNA expression (13)
- Text mining (152)
- Animal models (0)

Therapeutic area

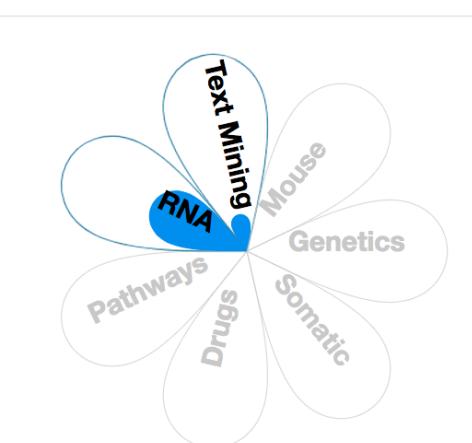
- Neoplasm (46)
- Phenotype (25)
- Digestive system disease (20)
- Nervous system disease (17)
- Endocrine system disease (16)
- Other (16)
- Metabolic disease (15)
- Cardiovascular disease (13)
- Reproductive system disease (12)
- Genetic disorder (10)
- Immune system disease (9)
- Respiratory system disease (9)
- Liver disease (8)
- Infectious disease (6)
- Hematological system disease (5)
- Skin disease (4)
- Skeletal system disease (4)
- Eye disease (3)
- Bladder disease (2)
- Head disease (1)
- Kidney disease (1)
- Biological process (1)



**hepatocellular carcinoma**  
**(Association Score: 0.57)**



[Profile | Associations](#)



[View evidence details](#)

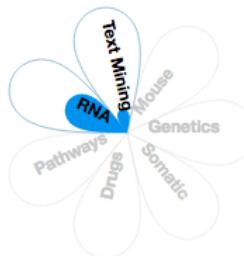
EMBL-EBI





# Expression Atlas data in other resources

Evidence for CYP1A2 in hepatocellular carcinoma



## CYP1A2

cytochrome P450 family 1 subfamily A member 2

Synonyms: P3-450, CP12

Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds.

## hepatocellular carcinoma

Synonyms: Cancer, Liver, Hepatocellular Carcinomas, Liver Cancer, Adult, NEOPL HEPATIC, Hepatomas, Primary Car...

Tumors or cancer of the LIVER.. A primary malignant neoplasm of epithelial liver cells. It ranges from a well-differentiated tumor with EPITHELIAL CELLS indistinguishable from normal HEPATOCYTES to a ...

Genetic associations

Somatic mutations

Drugs

Affected pathways

RNA expression

Text mining

Animal models



# Expression Atlas data in other resources

Evidence for CYP1A2 in hepatocellular carcinoma

Source: Expression Atlas

Showing 1 to 4 of 4 entries

Search:



Disease	Gene-disease evidence							Experiment overview and data	Publications
	Comparison	Activity	Tissue/cell	Evidence source	Log <sub>2</sub> fold change	P-value	Percentile rank		
hepatocellular carcinoma	'liver tumor' vs 'adjacent non-tumor tissue'	decreased ↘	liver	differential gene expression evidence from RNA-seq experiment	-10.1	3.32e-63	100	Transcriptome sequencing of human hepatocellular carcinoma ↗	
hepatocellular carcinoma	'tumor' vs 'paratumoral tissue'	decreased ↘	liver	differential gene expression evidence from RNA-seq experiment	-3.3	3.67e-6	98	Transcriptomic characterization of Hepatocellular Carcinoma with CTNNB1 mutation ↗	
hepatocellular carcinoma	'minor tumor' vs 'normal' in 'IL28B rs8099917 TG/GG'	decreased ↘	liver	differential gene expression evidence from microarray experiment	-3.3	4.93e-3	99	Hepatic gene expression of HCV related Hepatocellular carcinoma and non-cancerous tissue with IL28B rs8099917 TT genotype and TG/GG genotype ↗	

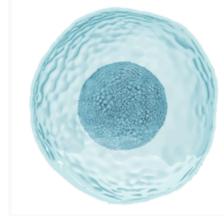


# New project on single-cell...

 Single Cell Expression Atlas

Single-cell gene expression across species

[Home](#) [Browse experiments](#) [Download](#) [Help](#) [About](#)



## Analysing gene expression for single cells experiments

The cell is a natural unit of biology, whose type and state can vary according to external influences or to internal processes. In multicellular organisms, all cells are derived from a single zygote which, through regulated programmes of proliferation and differentiation, generates all of the diverse cell types that populate the organism. Dysregulation of these programmes in single 'renegade' cells can lead to diseases such as cancers, neurological disorders and developmental disorders.

List of single-cell experiments

Total : 3 experiments

 06-02-2017 Single-cell RNA-seq analysis of human pancreas from healthy individuals and type 2 diabetes patients <i>Homo sapiens</i>	3514	<a href="#">Analysis result</a>	
 06-02-2017 Single-cell RNA-sequencing resolves a CD4+ T cell fate bifurcation <i>Mus musculus</i>	752	<a href="#">Analysis result</a>	
 06-02-2017 Single-cell RNA-sequencing of <i>Platynereis dumerilii</i> larval brain cells <i>Platynereis dumerilii</i>	285	<a href="#">Analysis result</a>	

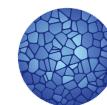
## Marker genes search

Gene ID, gene name or gene feature

Species

Any

Examples: REG1B, zinc finger, lung, leaf, valproic acid, cancer



HUMAN  
CELL  
ATLAS



# Expression Atlas: Who we are?

Robert  
Petryszak  
Team Leader



Irene  
Papatheodorou  
Co-ordinator



## Data curation



Laura Huerta

## Data analysis



Suhail Mohammed



Nuno Fonseca

## User Interface



Alfonso Fuentes



Elisabet Barrera



Nancy George

Anja Fullgrabe



Wojtek Bazant



Haider Iqbal

# Getting help and sending feedback

The screenshot shows the Expression Atlas homepage. At the top left is a blue circular logo with a yellow and white 'Q'. To its right is the text 'Expression Atlas' and 'Gene expression across species and biological conditions'. Below this is a horizontal navigation bar with six items: 'Home', 'Download', 'Release notes', 'FAQ' (which is highlighted in green), 'Help', 'Licence', and 'About' (which is also highlighted in green). The main content area has a teal background. It features a large heading 'About Expression Atlas', followed by 'What is Expression Atlas?'. Below this is a paragraph about the project's mission to provide freely available information on gene and protein expression across various biological conditions. Another paragraph explains the project's aim to answer questions like 'where is my favourite gene expressed?' and 'how its expression changes in a disease?'. A final paragraph describes the project's objective of involving data curation, analysis, and development of a web application. To the right of the text is a diagram consisting of four teal boxes labeled 'Data curation', 'Data analysis', 'Data search', and 'Data visualisation'. Below these are four smaller teal boxes: 'High quality data' with a document icon, 'High quality results' with a bar chart icon, 'Accessible results' with a sunburst chart icon, and '... easy to interpret' with a grid icon. A grey arrow points from the bottom of the first column to the bottom of the second column.

## About Expression Atlas

### What is Expression Atlas?

Expression Atlas is an open science resource that gives users a powerful way to find information about gene and protein expression. Our mission is to provide the scientific community with freely available information on the abundance and localisation of RNA (and proteins) across species and biological conditions such as different tissues, cell types, developmental stages and diseases among others.

Expression Atlas aims to help answering questions such as 'where is my favourite gene expressed?' or 'how its expression changes in a disease?'

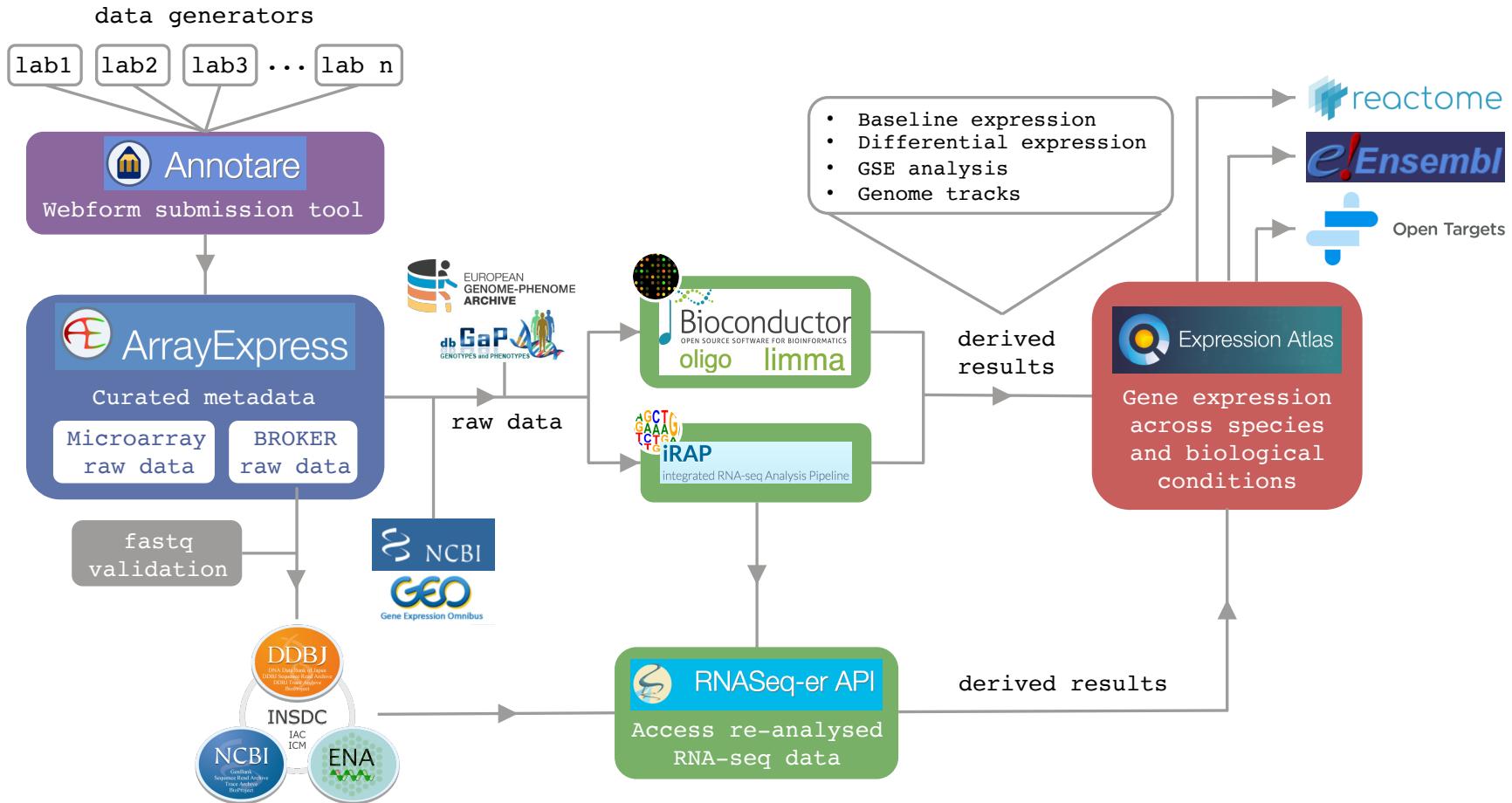
To achieve this objective, our project involves data curation, data analysis and the development of a web application to access and visualise publicly available data.

[atlas-feedback@ebi.ac.uk](mailto:atlas-feedback@ebi.ac.uk)



@ExpressionAtlas

# Functional genomics resources at EMBL-EBI



# Please fill in the survey before you go!

[https://www.surveymonkey.co.uk/r/PIWet17\\_genomics](https://www.surveymonkey.co.uk/r/PIWet17_genomics)



EMBL-EBI resources and tools for genomics, PIWet/NVRI, Puławy, Dec 2017

Workshop Evaluation Form

EMBL-EBI collects feedback from every course and workshop we run. The survey is a way for you to inform us about the workshop you have participated in, what you enjoyed, what you found useful and how we can make improvements. This information is also used to inform the development of new courses and workshops.

Next >>

# EMBL-EBI Resources and tools for genomics and proteomics

## Expression Atlas: gene expression results across species and conditions

Laura Huerta, PhD

Senior Scientific Curator

[lauhuema@ebi.ac.uk](mailto:lauhuema@ebi.ac.uk)

6 December 2017

