



# Placental Bioinformatics Course

Dr Russell S. Hamilton

Email: rsh46@cam.ac.uk  
Twitter: @drrshamilton

Dr Xiaohui Zhao

Email: xz289@cam.ac.uk

Dr Malwina Prater

Email: mn367@cam.ac.uk

Course Materials:

<https://github.com/darogan/2018-PlacentalBiologyCourse>



License:

Attribution-Non Commercial-Share Alike CC BY-NC-SA ( <https://creativecommons.org/licenses/by-nc-sa/> )

Attribution: You must give appropriate credit, provide a link to the license, and indicate if changes were made.

NonCommercial: You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

You may not use the material for commercial purposes.

ShareAlike: If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

Version 0.3: 20180705



# An introduction to RNA-Seq

Objective for the practical

To learn about how to process, analyse and interpret RNA-Seq data

Learning goals:

Experience using the Linux operating system (Ubuntu)

Learn how to run bioinformatics tools on the command line

Learn how to use R and R-studio

Visualise sequencing alignments in a genome browser

# Introduction

## RNA-Seq Differential Gene expression between the Placenta and Yolk Sac Mouse

### Data availability

mRNA-Seq datasets have been deposited in Gene Expression Omnibus (GEO) under accession numbers GSE66137 and GSE66138. NanoString datasets have been deposited in GEO under accession number GSE79767.

### Bioinformatics Top Tip:

Download fastq files directly from <http://www.ebi.ac.uk/ena>



### Warning:

*This is a demo with a reduced data set and parameters,  
so take any genes identified with caution*

### RESEARCH ARTICLE

## Deficiency of the placenta- and yolk sac-specific tristetraprolin family member ZFP36L3 identifies likely mRNA targets and an unexpected link to placental iron metabolism

Deborah J. Stumpo<sup>1</sup>, Carol S. Trempus<sup>2</sup>, Charles J. Tucker<sup>2</sup>, Weichuan Huang<sup>3</sup>, Leping Li<sup>4</sup>, Kimberly Kluckman<sup>5</sup>, Donna M. Bortner<sup>2</sup> and Perry J. Blackshear<sup>1,6,\*</sup>

### ABSTRACT

The ZFP36L3 protein is a rodent-specific, placenta- and yolk sac-specific member of the tristetraprolin (TTP) family of CCCH tandem zinc-finger proteins. These proteins bind to AU-rich elements in target mRNAs, and promote their deadenylation and decay. We addressed the hypothesis that the absence of ZFP36L3 would result in the accumulation of target transcripts in placenta and/or yolk sac, and that some of these would be important for female reproductive physiology and overall fecundity. Mice deficient in ZFP36L3 exhibited decreased neonatal survival rates, but no apparent morphological changes in the placenta or surviving offspring. We found Zfp36l3 to be paternally imprinted, with profound parent-of-origin effects on gene expression. The protein was highly expressed in the syncytiotrophoblast cells of the labyrinth layer of the placenta, and the epithelial cells of the yolk sac. RNA-Seq of placental mRNA from Zfp36l3 knockout (KO) mice revealed many significantly upregulated transcripts, whereas there were few changes in KO yolk sacs. Many of the upregulated placental transcripts exhibited decreased decay rates in differentiated trophoblast stem cells derived from KO mastocytes. Several dozen transcripts were deemed high probability targets of ZFP36L3; these include proteins known to be involved in implantation and placenta physiology. Type 1 transferrin receptor mRNA was unexpectedly decreased in KO placenta, despite an increase in its stability in KO stem cells. This receptor is crucial for placental iron uptake, and its decrease was accompanied by decreased iron stores in the KO fetus, suggesting that this intrauterine deficiency might have deleterious consequences in later life.

**KEY WORDS:** Deadenylation, mRNA binding proteins, mRNA decay, Placenta, Zinc finger proteins

### INTRODUCTION

The tristetraprolin (TTP) family of CCCH tandem zinc finger proteins consists of mRNA-binding proteins that are thought to regulate gene expression by binding to and promoting the decay of mRNAs containing specific types of AU-rich element-binding sites (Blackshear and Preys, 2014; Brooks and Blackshear, 2013).

<sup>1</sup>Laboratory of Signal Transduction, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA; <sup>2</sup>Laboratory of Clinical Pharmacology, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA; <sup>3</sup>Coastal Microscopy Core, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA; <sup>4</sup>Biostatistics Branch, National Institute of Environmental Health Sciences, Research Triangle Park, NC 27709, USA; <sup>5</sup>Theravance, Inc., Research Triangle Park, NC 27709, USA; <sup>6</sup>Department of Molecular and Biochemical Medicine, Duke University Medical Center, Durham, NC 27710, USA

\*Author for correspondence: ([dev130389@nremsa.nih.gov](mailto:dev130389@nremsa.nih.gov))

Received 1 September 2015; Accepted 23 February 2016

1424

These sites are generally located in the 3'-untranslated regions (3'UTRs) of mRNAs and have the optimal sequence of UUAUAUUUU (Lai et al., 2009); physiologically relevant confirmed target transcripts often contain multiple binding sites in close proximity to each other (Carballo et al., 1998, 2008). The mechanism of the induced mRNA instability is not entirely understood, but appears to involve initial stimulation of deadenylation, or removal of the poly(A) tail, thought to be the rate-limiting step in mRNA decay in all eukaryotes (Carballo et al., 2008; Fabian et al., 2013; Lai et al., 2003).

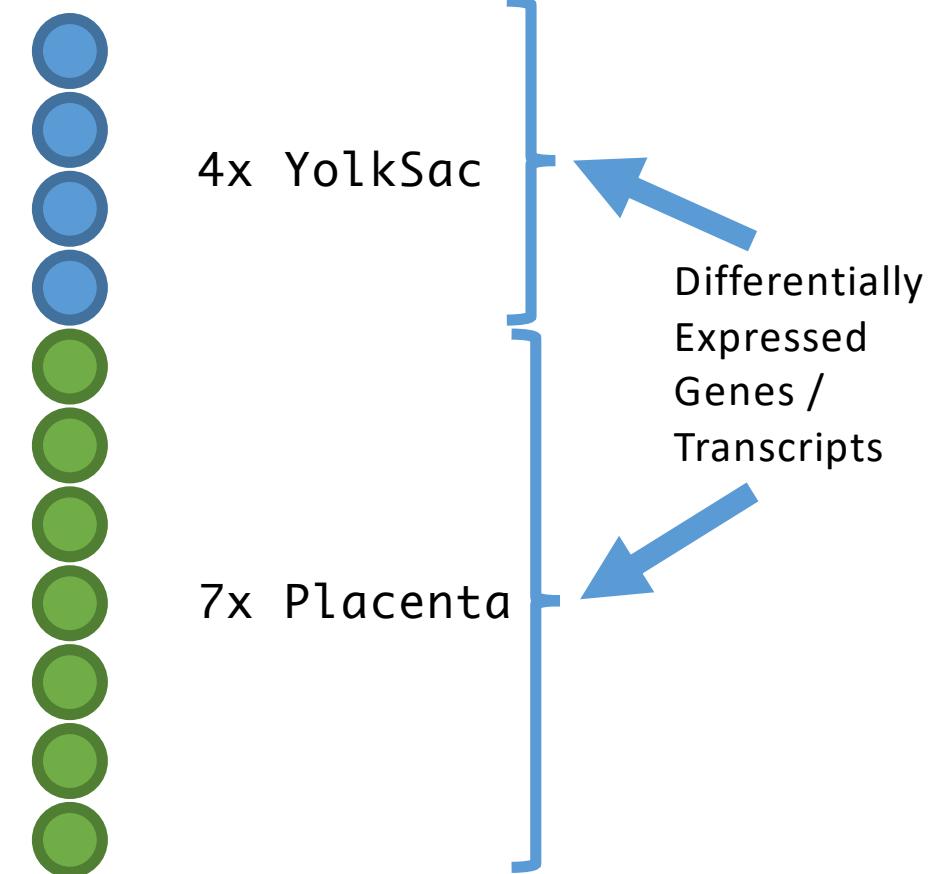
In most mammals, the TTP protein family consists of three members, which behave similarly in biochemical assays of mRNA binding and destabilization (Froderick et al., 2008; Lai et al., 2003). However, knockout (KO) mice for these three family members exhibit dramatic physiological specificity. For example, TTP KO mice (*Zfp36l3*<sup>-/-</sup>) develop a systemic inflammatory syndrome that is largely due to chronic elevation in tumor necrosis factor alpha (TNFα); the *Puf* mRNA was found to be a direct target of TTP binding and induced mRNA destabilization (Carballo and Blackshear, 2001; Carballo et al., 1998; Taylor et al., 1996). KO of a second family member, *Zfp36L1*, results in embryonic lethality, apparently due to failure of choriocarcinoma fusion, an essential step in the development of the umbilical circulation (Stumpo et al., 2006). The third family member, *Zfp36L2*, is crucial for definitive hematopoiesis (Stumpo et al., 2009).

A fourth mammalian gene in this family, *Zfp36l3*, is an X chromosome gene expressed only in the yolk sac and placenta of certain rodents, including mice and rats (Blackshear et al., 2005; Froderick et al., 2008; Gingrich et al., 2014). The ZFP36L3 protein differs from its other family members in several respects; for example, in contrast to the other family members, which are macrocytovironetting proteins (Phillips et al., 2002), the mouse and rat ZFP36L3 proteins contain long series of carboxy-terminal repeats that serve to maintain the protein in the cytoplasm (Froderick et al., 2008). We describe here the development of a *Zfp36l3* KO mouse, and its phenotype and biochemical characterization. Our findings highlight the importance of ZFP36L3 in mouse fertility, as well as its influence on post-transcriptional gene expression in placenta.

**RESULTS**  
**Expression of ZFP36L3 during normal development**  
In placenta, *Zfp36l3* mRNA was readily detected by embryonic day (E) 9.5, increased to near maximum levels by E14.5, and then remained elevated until E18.5. In the yolk sac, *Zfp36l3* mRNA was essentially undetectable at E10.5, and then began accumulating, reaching a peak by E18.5 (Fig. 1A). Transcripts for the other family members were readily detectable and largely constant in both yolk

# Introduction

sample	condition
SRR1811706	WT Yolk Sac
SRR1811707	WT Yolk Sac
SRR1811708	WT Yolk Sac
SRR1811709	WT Yolk Sac
SRR1823638	WT Placenta
SRR1823639	WT Placenta
SRR1823640	WT Placenta
SRR1823641	WT Placenta
SRR1823642	WT Placenta
SRR1823643	WT Placenta
SRR1823644	WT Placenta





# Bioinformatics Pipeline

Sequencing Files (FASTQ)

FastQC



Perform quality control (adapter contamination, base quality)

trim\_galore

Align reads to the genome/transcriptome

kallisto

Command Line  
&  
FireFox

Summarise QC and alignment metrics

MultiQC

Coffee break

Perform differential gene/transcript expression analysis

sleuth/DESeq2

R/R-Studio

Visualise sequencing alignments in a genome browser

IGV

IGV

Look at differentially enriched genes

ensEMBL/DAVID

Firefox



# Using the Bioinformatics Training Facility Computers

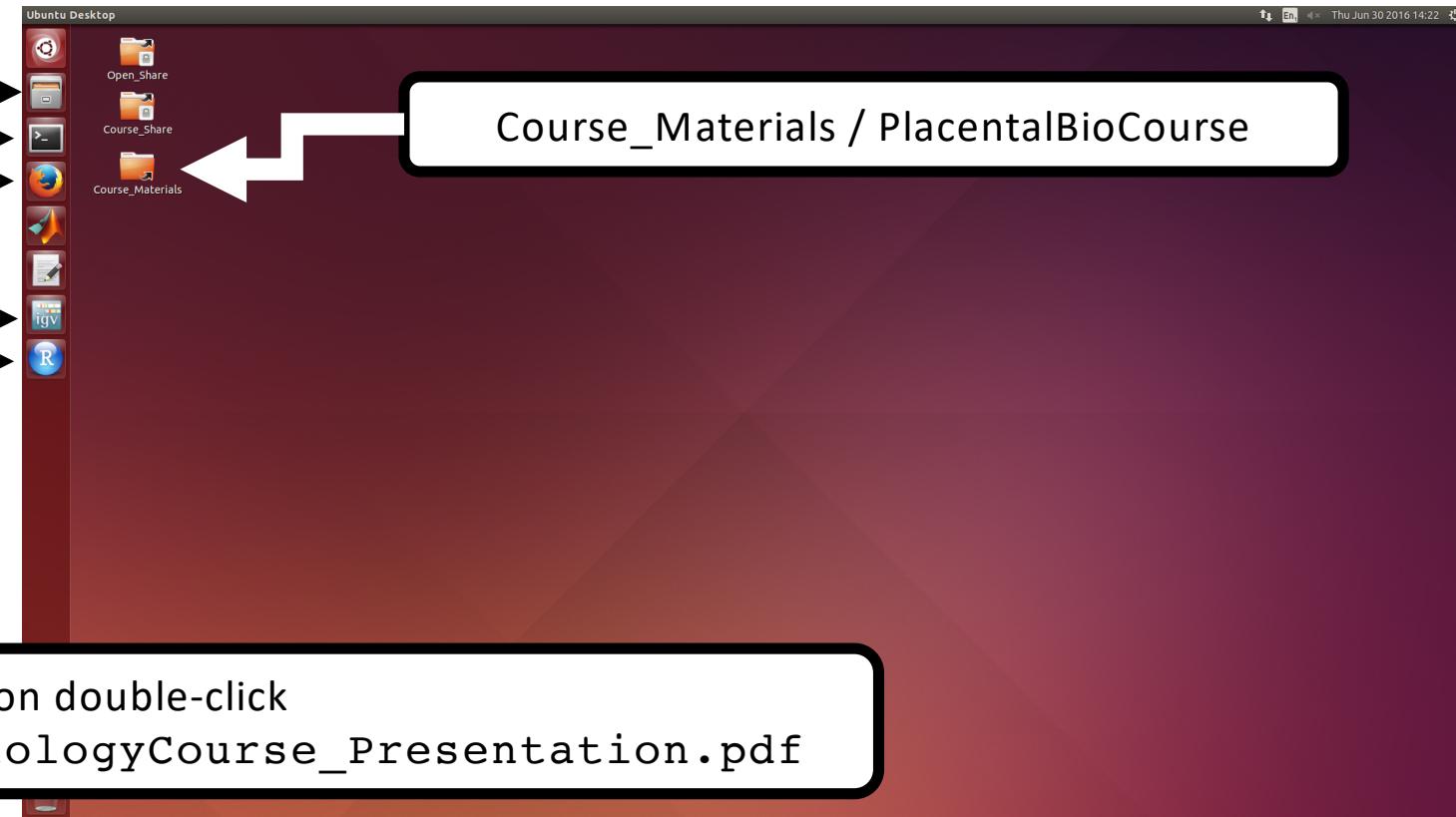
Finder / Windows Explorer

Terminal

Firefox

IGV

R-Studio



Linux :::: Ubuntu

## Course Materials

2018\_PlacentalBiologyCourse.pptx  
2018\_PlacentalBiologyCourse.pdf  
2018\_PlacentalBiologyCourse\_Sleuth.Rmd  
2018\_PlacentalBiologyCourse\_commands.txt

## Publication

stumpo\_2016\_Development.pdf

## Sample Data

SRR1811706\_ES610\_WT\_Yolk\_Sac/  
SRR1811707\_ES611\_WT\_Yolk\_Sac/  
SRR1811708\_ES612\_WT\_Yolk\_Sac/  
SRR1811709\_ES613\_WT\_Yolk\_Sac/  
SRR1823638\_ES51\_WT\_Placenta/  
SRR1823639\_ES51\_WT\_Placenta/  
SRR1823640\_ES52\_WT\_Placenta/  
SRR1823641\_ES52\_WT\_Placenta/  
SRR1823642\_ES53\_WT\_Placenta/  
SRR1823643\_ES54\_WT\_Placenta/  
SRR1823644\_ES55\_WT\_Placenta/

## Alignments (Premade)

IGV\_Alignments/  
ENST\_ENSG\_GeneName.GRCm38.kallisto.table  
Mus\_musculus.GRCm38.cdna.all.idx

## Reference Genome

SRR1823638\_sub\_1.fastq.gz  
SRR1823638\_sub\_2.fastq.gz

## SRR1823638 Sequencing Data

SRR1823638\_1.fastq.gz\_trimming\_report.txt  
SRR1823638\_1\_fastqc.html  
SRR1823638\_1\_fastqc.zip  
SRR1823638\_1\_val\_1.fq.gz\_kallisto\_output/  
SRR1823638\_1\_val\_1\_fastqc.html  
SRR1823638\_1\_val\_1\_fastqc.zip  
SRR1823638\_2.fastq.gz\_trimming\_report.txt  
SRR1823638\_2\_fastqc.html  
SRR1823638\_2\_fastqc.zip  
SRR1823638\_2\_val\_2\_fastqc.html  
SRR1823638\_2\_val\_2\_fastqc.zip

## Kallisto Output

abundance.h5  
abundance.tsv  
run\_info.json

Placenta\_SRR1823638\_kallisto.bam  
Placenta\_SRR1823638\_kallisto.bam.bai  
YolkSac\_SRR1811706\_kallisto.bam  
YolkSac\_SRR1811706\_kallisto.bam.bai

Subset of real data for this practical 1M reads only (~10%)



# Using the Bioinformatics Training Facility Computers

## Terminal

```
rhamilton — participant@winfields: ~/Course_Materials — ssh participant@131.111.146.51 -Y — 84x15
rsh46@login-sand3:~/scratch/CTR... participant@winfields: ~/Course_Materials...
participant@winfields:~$ cd Course_Materials
[participont@winfields:~/Course_Materials]
participant@winfields:~/Course_Materials$ cd PlacentalBiologyCourse
```

change directory      Directory name  
  
\$ cd Course\_Materials  
  
\$ cd PlacentalBioCourse

### Bioinformatics Top Tip: More Linux Commands

ls	list files in directory
cd ~	change back to home directory
tree	view files and directories in a hierarchical structure
history	view a list of the most recent commands used

Caution!

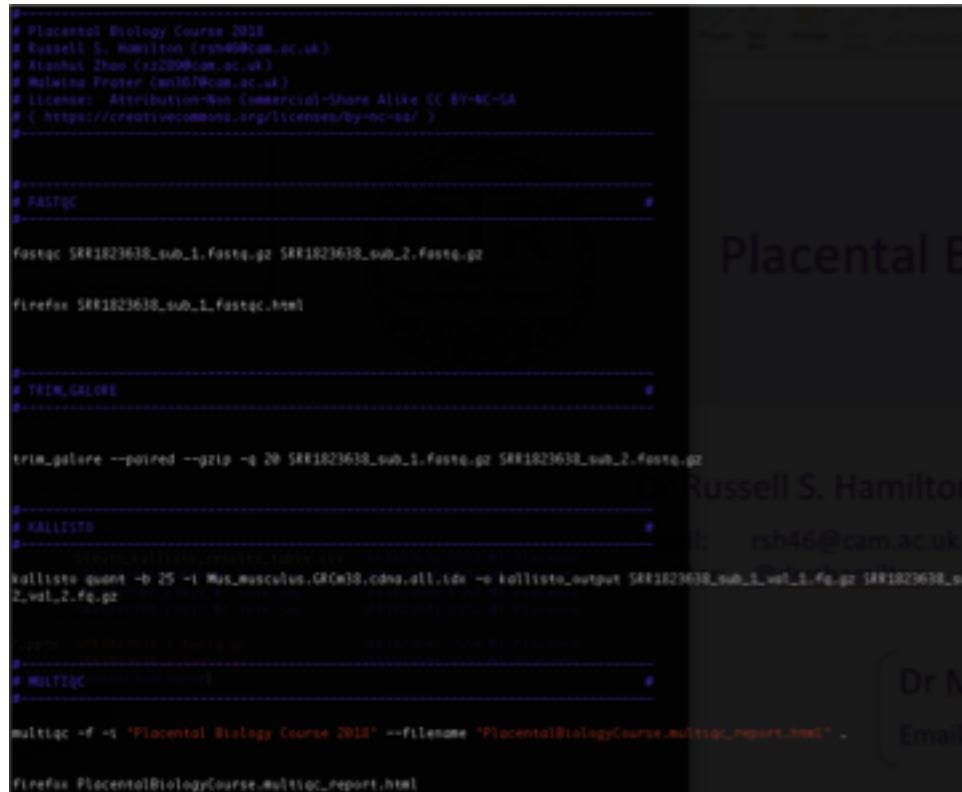
Commands are case sensitive

Take care to correctly specify spaces and flags (dashes)

# Command Line

Copy & Paste from PDF or Powerpoint to the command line doesn't always work as expected

2018\_PlacentalBiologyCourse\_Commands.txt



```
# Placental Biology Course 2018
# Russell S. Hamilton (rsh46@cam.ac.uk)
# Xiaohui Zhao (xz289@cam.ac.uk)
# Malwina Prater (mn367@cam.ac.uk)
# License: Attribution-Non Commercial-Share Alike CC BY-NC-SA
# < https://creativecommons.org/licenses/by-nc-sa/ >
# 

# FASTQC
# 

fastqc SRR1823638_sub_1.fastq.gz SRR1823638_sub_2.fastq.gz

firefox SRR1823638_sub_1.fastqc.html

# 

# TRIM_GALORE
# 

trim_galore --paired --gzip -q 20 SRR1823638_sub_1.fastq.gz SRR1823638_sub_2.fastq.gz

# 

# KALLISTO
# 

kallisto quant -b 25 -l Mus_musculus.GRCm38.cdna.all.tcr -o kallisto_output SRR1823638_sub_1.vol_1.fq.gz SRR1823638_sub_2.vol_2.fq.gz

# 

# MULTICQ
# 

multiqc -f -s "Placental Biology Course 2018" --filename "PlacentalBiologyCourse.multiqc_report.html"

firefox PlacentalBiologyCourse.multiqc_report.html
```

FastQC

Version

A quality control tool for high throughput sequence data

0.11.5

Download

<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Terminal:

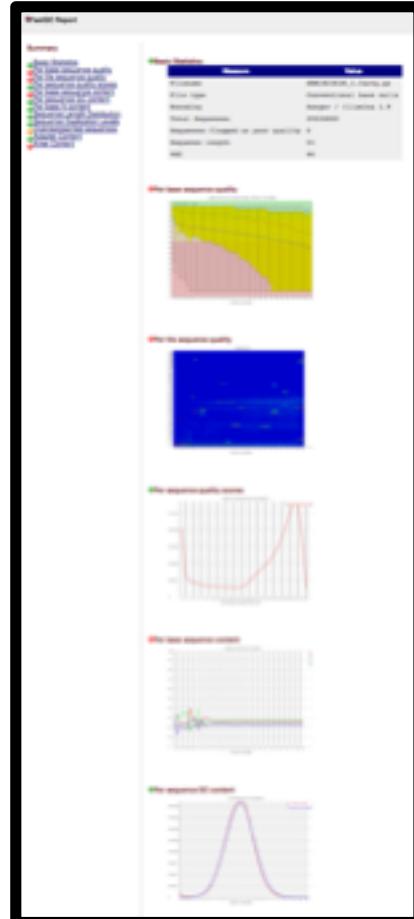
```
Read 1
```

```
Read 2
```

```
$ fastqc SRR1823638_sub_1.fastq.gz SRR1823638_sub_2.fastq.gz
```

```
$ firefox SRR1823638_sub_1_fastqc.html
```

Output:

*HTML Reports*SRR1823638\_sub\_1\_fastqc.html  
SRR1823638\_sub\_2\_fastqc.html*Archive of data/images*SRR1823638\_sub\_1\_fastqc.zip  
SRR1823638\_sub\_2\_fastqc.zip

*Bioinformatics Top Tip:* Simon Andrews' <https://sequencing.qcfail.com/>



# trim\_galore

trim\_galore A wrapper tool around Cutadapt to consistently apply quality and adapter trimming to FastQ files  
Version 0.4.4  
Download [http://www.bioinformatics.babraham.ac.uk/projects/trim\\_galore/](http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/)

Terminal:

```
$ trim_galore --paired --gzip -q 20 SRR1823638_sub_1.fastq.gz SRR1823638_sub_2.fastq.gz
```

Compress the output fastq files

Read 1

Read 2

--paired

--gzip

-q 20

SRR1823638\_sub\_1.fastq.gz

SRR1823638\_sub\_2.fastq.gz

Treat as paired-end

Quality score threshold (PHRED > 20)

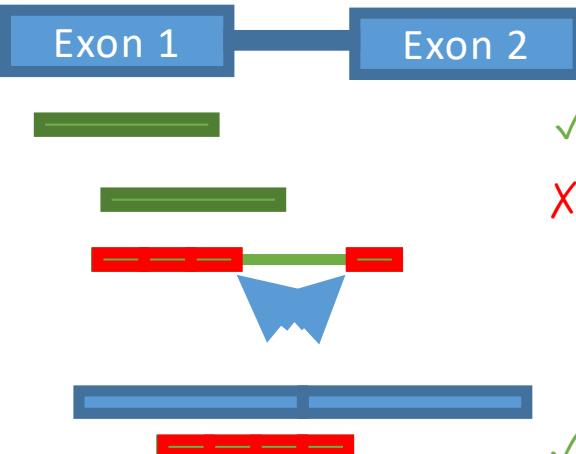
Output:

*Trimmed Fastq files*

SRR1823638\_sub\_1\_val\_1.fq.gz  
SRR1823638\_sub\_2\_val\_2.fq.gz

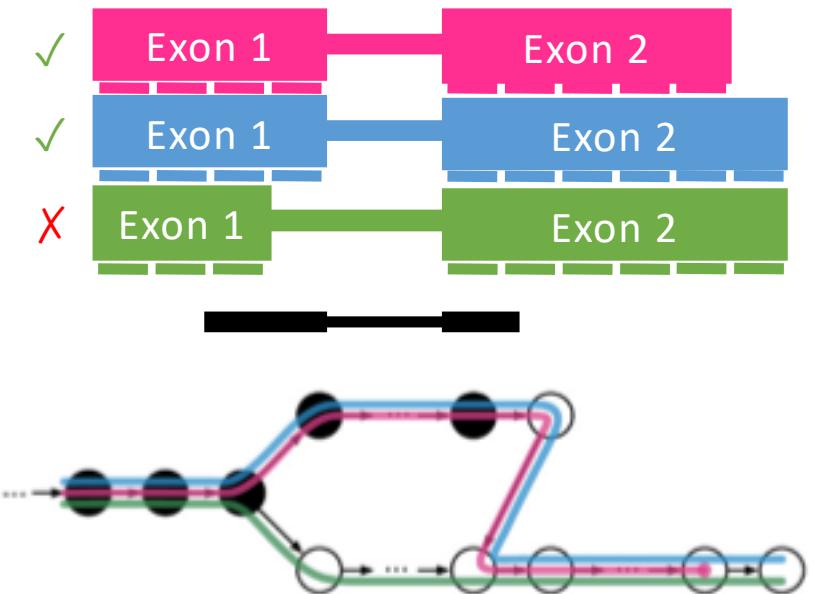
# Alignment: STAR Vs Kallisto

## STAR: Align to genome



- ✓ Single exon mapping
- ✗ Multi-exon reads  
Reads divided into segments  
splice site identified
- ✓ Segments aligned and  
assembled

## Kallisto: Align to transcriptome



	STAR	Kallisto
<i>Run time</i>	hours	minutes
<i>Hardware requirements</i>	Multi-core	Laptop
<i>Novel Splice Sites</i>	yes	no

kallisto      Program for quantifying abundances of transcripts from RNA-Seq data, without the need for alignment  
Version      0.43.1  
Download      <https://pachterlab.github.io/kallisto>

Terminal:

```
$ kallisto quant -b 25 -i Mus_musculus.GRCm38.cdna.all.idx  
      -o kallisto_output SRR1823638_sub_1_val_1.fq.gz SRR1823638_sub_2_val_2.fq.gz
```

Number of bootstraps      Indexed transcriptome  
Output directory      Trimmed Read 1      Trimmed Read 2

Output:

*Note command must be all on one single line*

```
kallisto_output/  
    abundance.h5  
    abundance.tsv  
    run_info.json
```

# RNA-Seq Mapping Metrics: Counts Vs FPKM Vs TPM

## Counts

The number of reads mapping to a transcript or gene  
Longer transcripts will generally have more mapped reads

None of these are for comparing across samples

## FPKM (Fragments Per Kilobase of transcript per Million mapped reads)

Normalises the counts for the length of the transcript

Sample normalisation required as performed by DESeq2 and Sleuth

## TPM (Transcripts Per Million)

Measurement of the proportion of transcripts in your pool of RNA

MultiQC	Aggregate results from bioinformatics analyses across many samples into a single report
Version	1.1dev
Download	<a href="http://multiqc.info/">http://multiqc.info/</a>

Terminal:

```
$ multiqc -f -i "Placental Biology Course 2018"  
    --filename "PlacentalBiologyCourse.multiqc_report.html" .  
$ firefox PlacentalBiologyCourse.multiqc_report.html
```

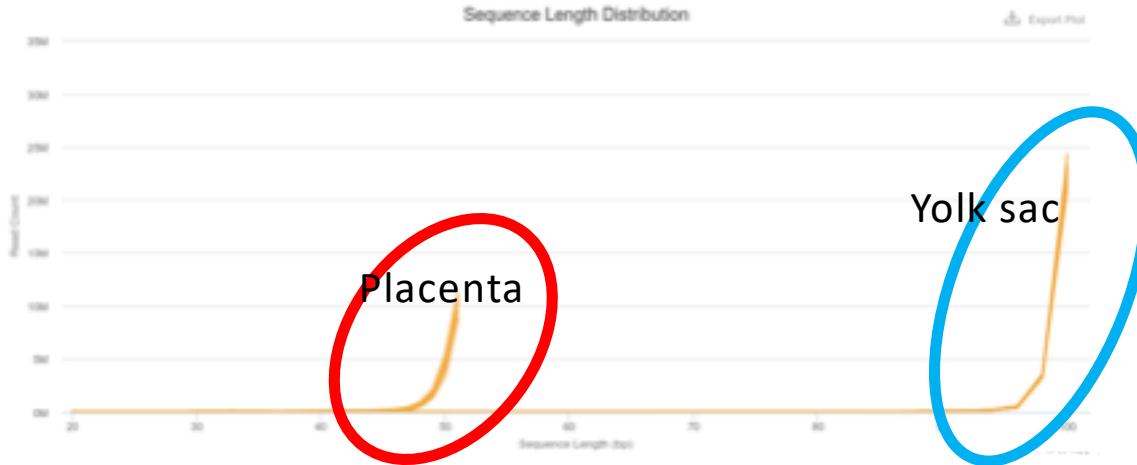
Overwrite existing report      A title for your report  
Output filename      ." Is a special Linux symbol which means the current directory

Output:

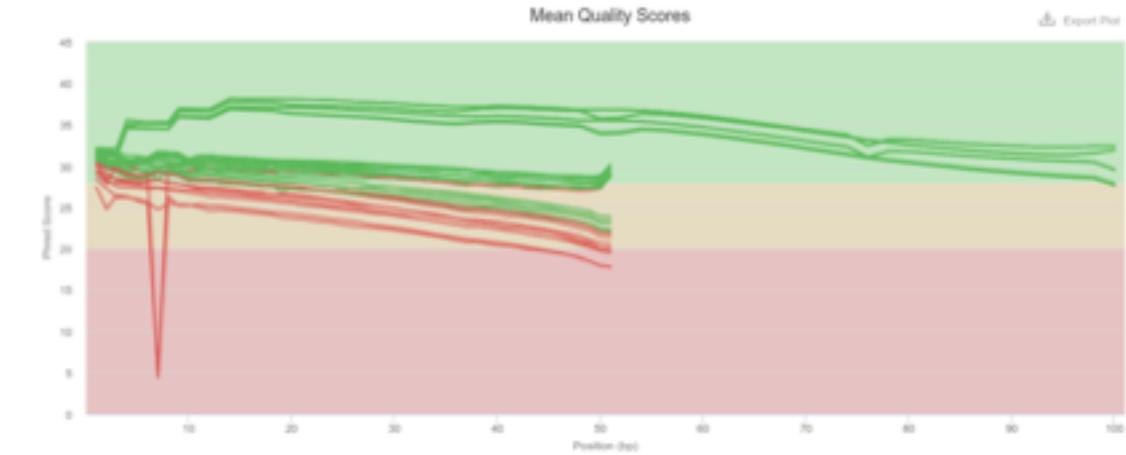
*HTML Report*  
PlacentalBiologyCourse.multiqc\_report.html  
PlacentalBiologyCourse.multiqc\_report\_data

## Batch effects

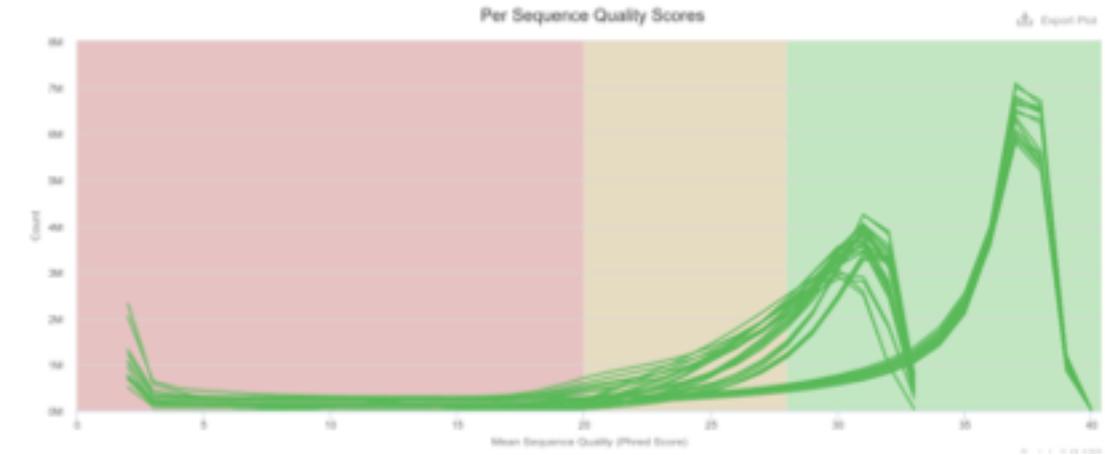
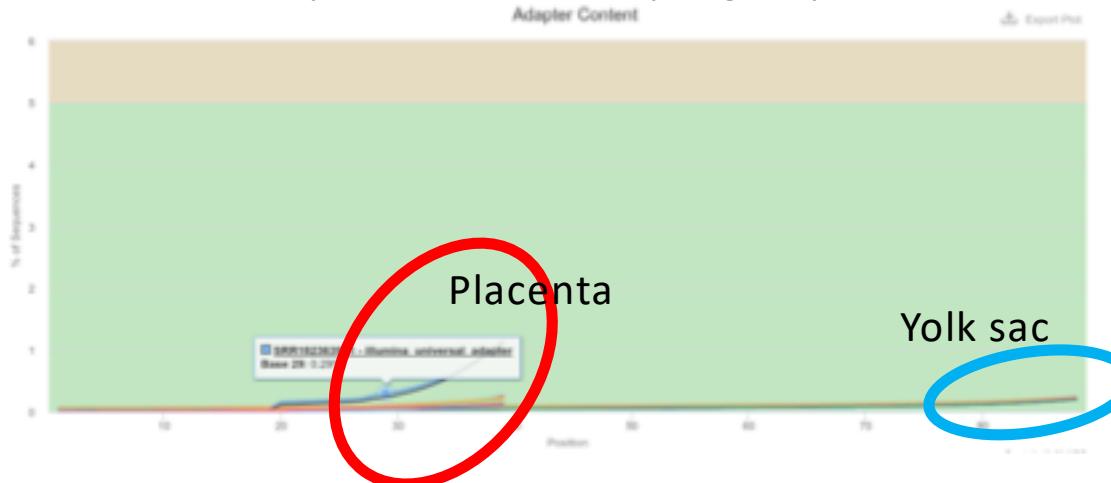
Sample groups have different read lengths



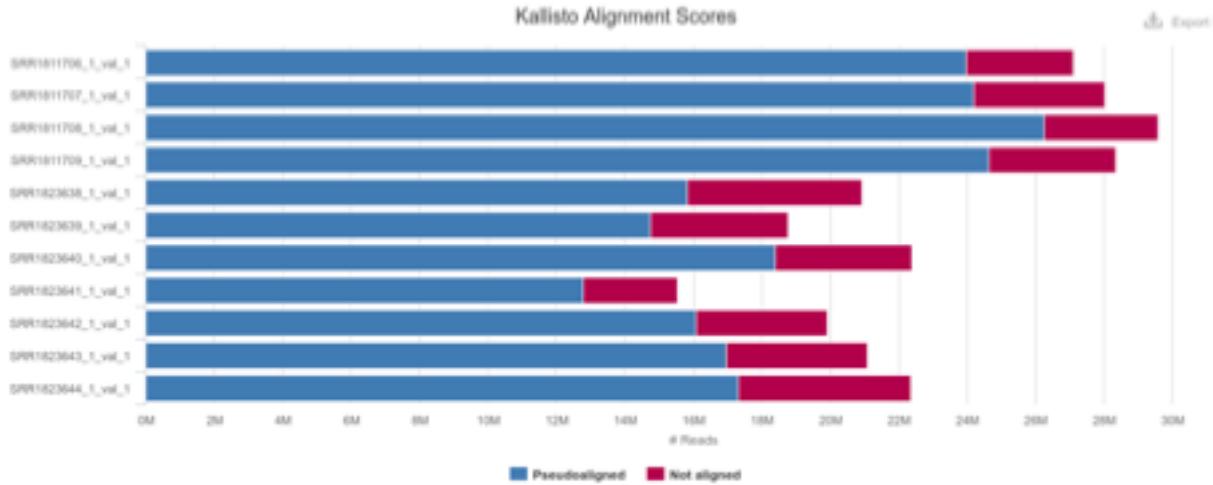
Some Placenta samples have low quality scores



There are adapters in both sample groups



# QC Alignments

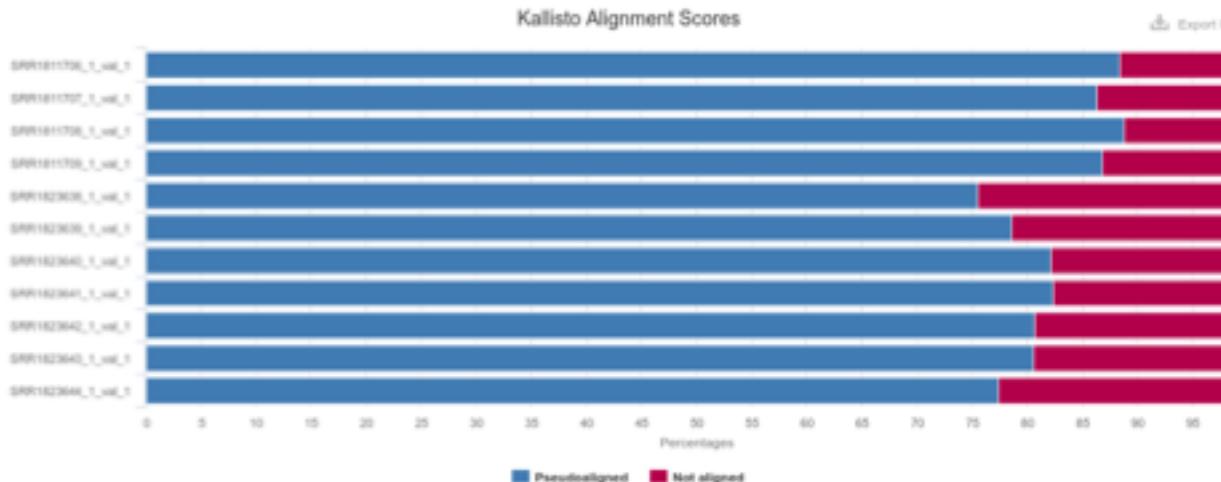


Yolk Sac

Placenta

*Why do you never see 100% alignment?*

- Incomplete reference genomes / transcriptomes
- Repetitive reads hard to map uniquely
- Sample: Structural Variants Copy Number Variants



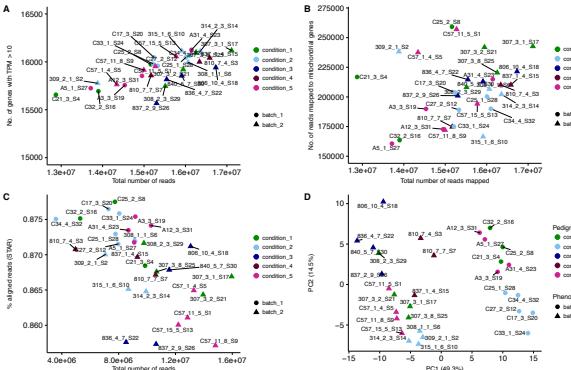
Yolk Sac

Placenta

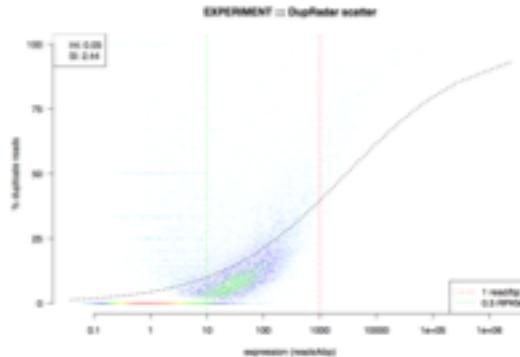
Harsher  
trimming,  
more reads  
removed /  
trimmed

# Further QC

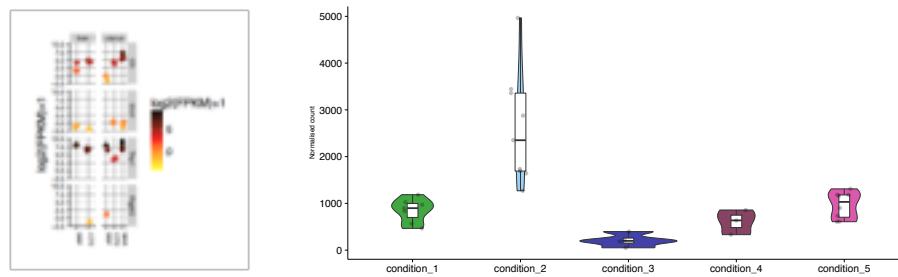
## rRNA and mt content



## dupRadar



## Custom Plots



## Batch effect correction

e.g. Two sequencing experiments performed on different days by two different people

Batch effect removal:  
Limma / Combat

Use batch in comparison in  
DESeq2 design  
~ batch + condition



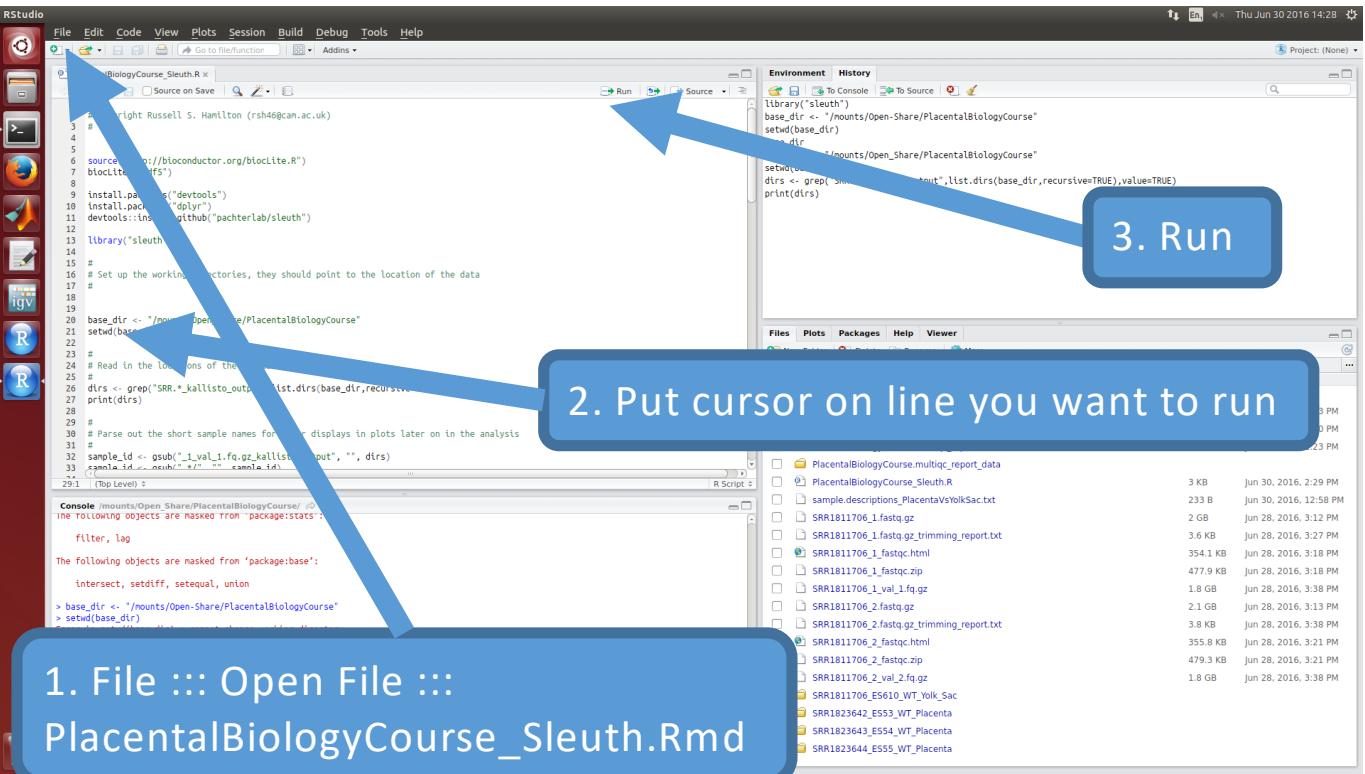
## Coffee Break

Part 1. Quality control, alignment and transcript quantification

Part 2. Differential Expression Analysis

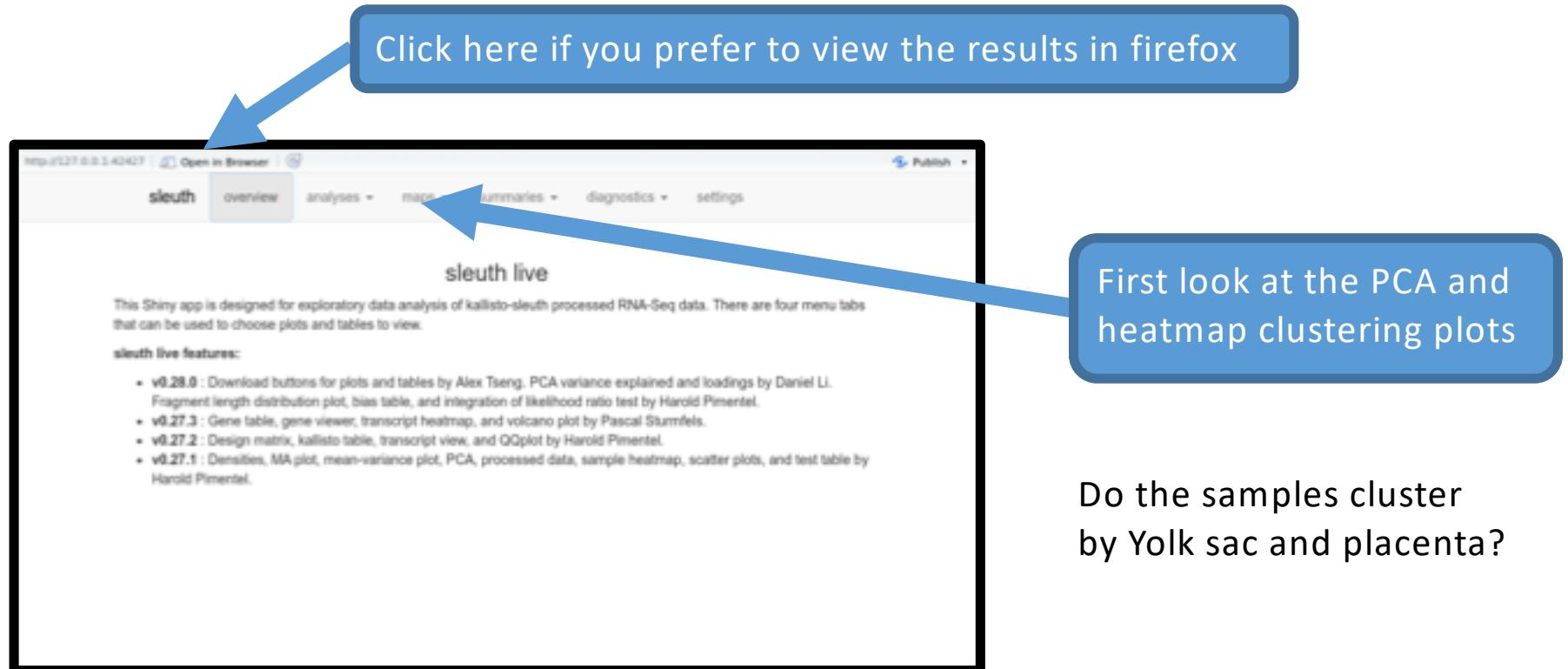
sleuth  
Version  
Download

Analysis of RNA-Seq experiments for which transcript abundances have been quantified with kallisto  
0.28.1  
<http://pachterlab.github.io/sleuth/> R



- A statistical programming language
  - R-Studio, a graphical environment for using R
  - # denotes a comment

Click here if you prefer to view the results in firefox



sleuth live

This Shiny app is designed for exploratory data analysis of kallisto-sleuth processed RNA-Seq data. There are four menu tabs that can be used to choose plots and tables to view.

sleuth live features:

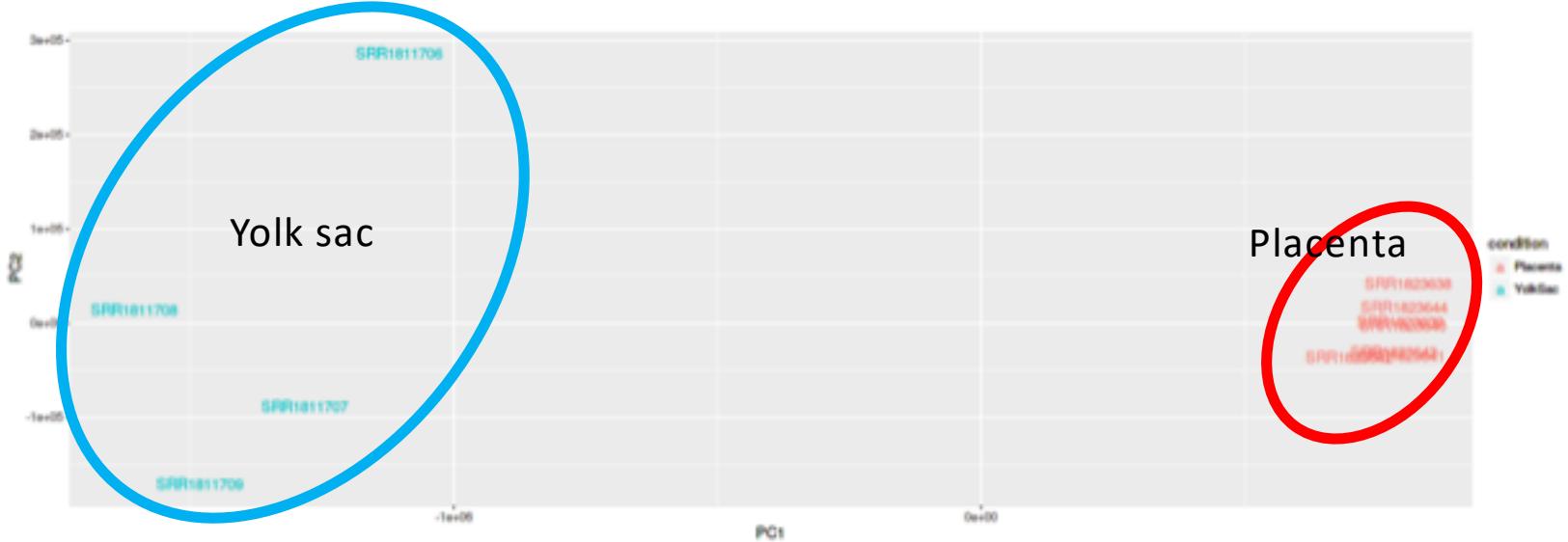
- v0.28.0 : Download buttons for plots and tables by Alex Tseng. PCA variance explained and loadings by Daniel Li. Fragment length distribution plot, bias table, and integration of likelihood ratio test by Harold Pimentel.
- v0.27.3 : Gene table, gene viewer, transcript heatmap, and volcano plot by Pascal Sturmels.
- v0.27.2 : Design matrix, kallisto table, transcript view, and QQplot by Harold Pimentel.
- v0.27.1 : Densities, MA plot, mean-variance plot, PCA, processed data, sample heatmap, scatter plots, and test table by Harold Pimentel.

First look at the PCA and heatmap clustering plots

Do the samples cluster by Yolk sac and placenta?

# Sample Clustering

PCA Plot



Heat Map



# Volcano Plot

## volcano plot

Plot of beta value (regression) versus log of significance. Select a set of transcripts to explore their variance across samples.



# Differentially Expressed Genes

Paste an external gene  
shortcode here

gene view

(Boxplots of abundances of transcript mapping to a given gene, and their technical variation. This step can take a while, especially with many plots.)

gene:

ENSGALUS000000022150.16

color by:

condition

units:

tpm

genes from:

ens\_gene

# of plots (max 15):

3

view

ENSMUST00000160134.7

Yolk sac

Placenta

sample

ENSMUST00000080880.11

Yolk sac

Placenta

sample

Select TPM  
Transcripts Per Million

Select external\_  
gene\_ shortcode

Yolk sac

Placenta

sample

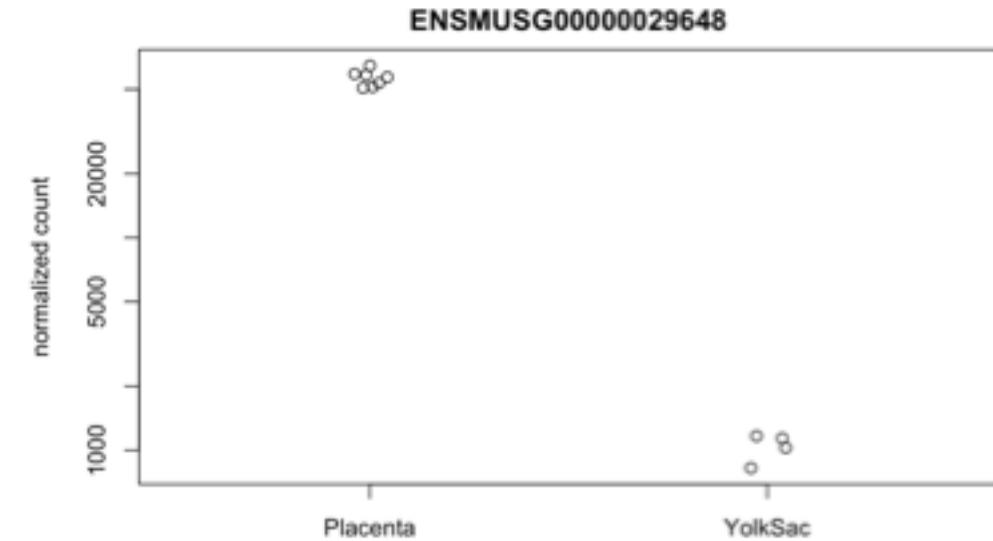
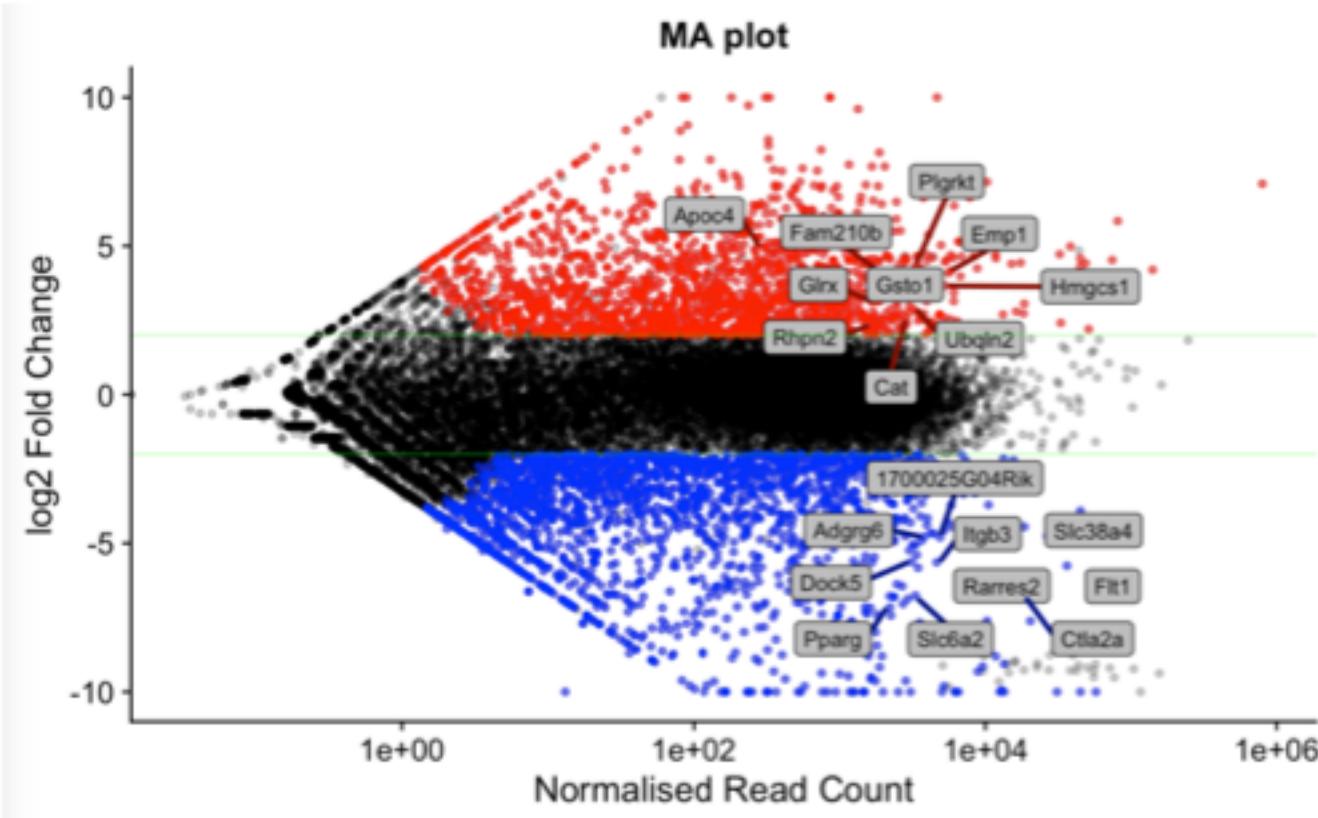
ENSMUST00000080880.11

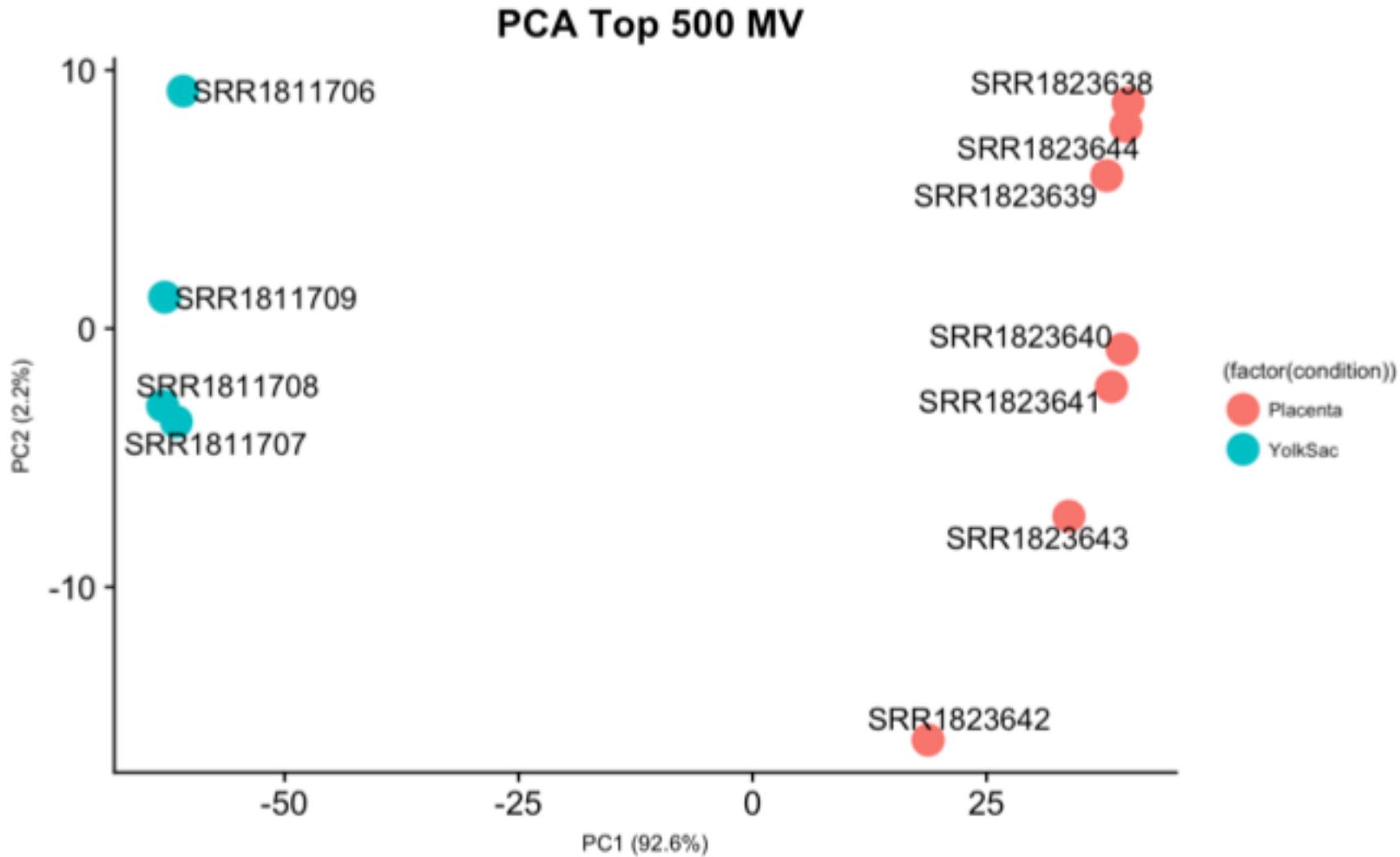
Yolk sac

Placenta

sample

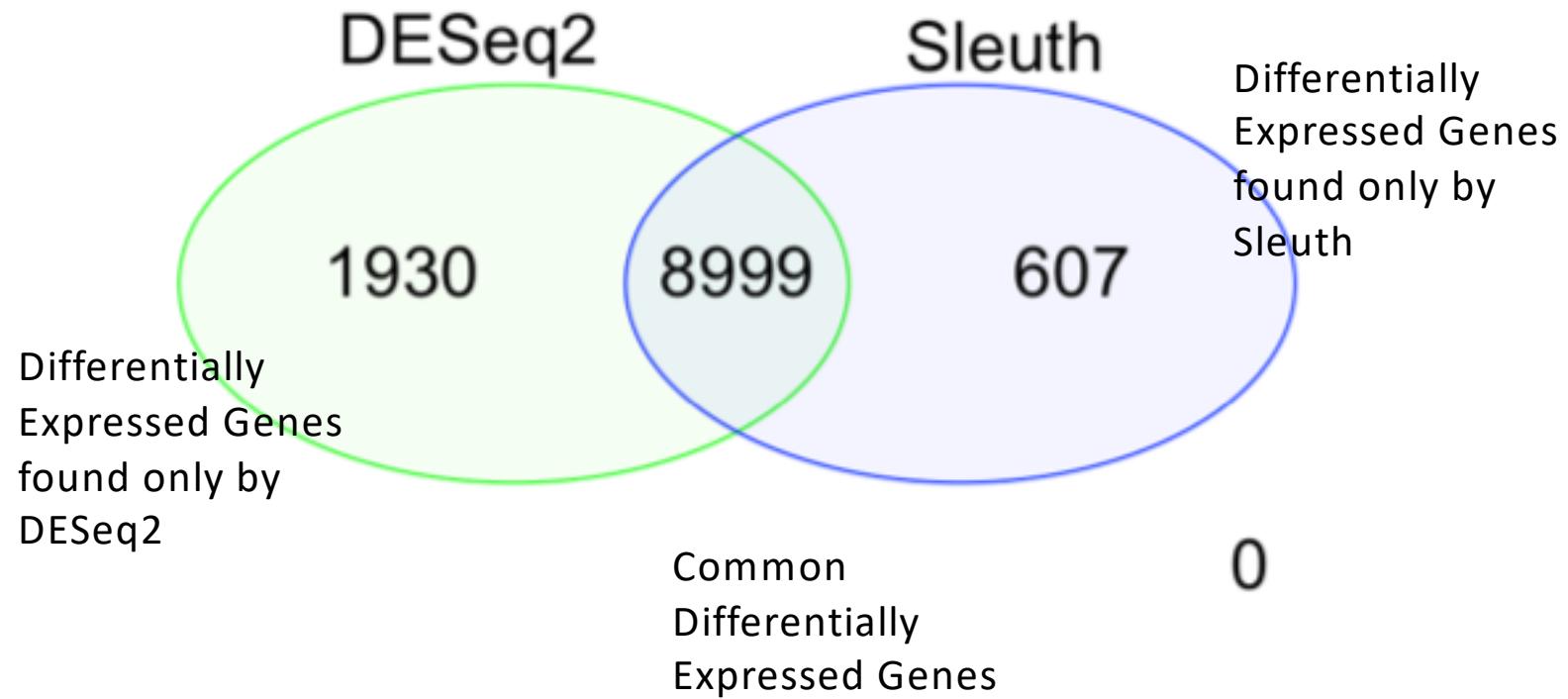
## From global picture to individual genes





Similar to the Sleuth version, but you can customise number of genes to use, colour, labels etc much more easily using the DESeq2/ggplot custom version

# Differentially Expressed Genes



- DESeq2 is considered to perform better at a gene level
- Sleuth performs better on a transcript level
  - Better for identifying alternatively spliced isoform usage

# Looking at the Alignments

File

Load From File...

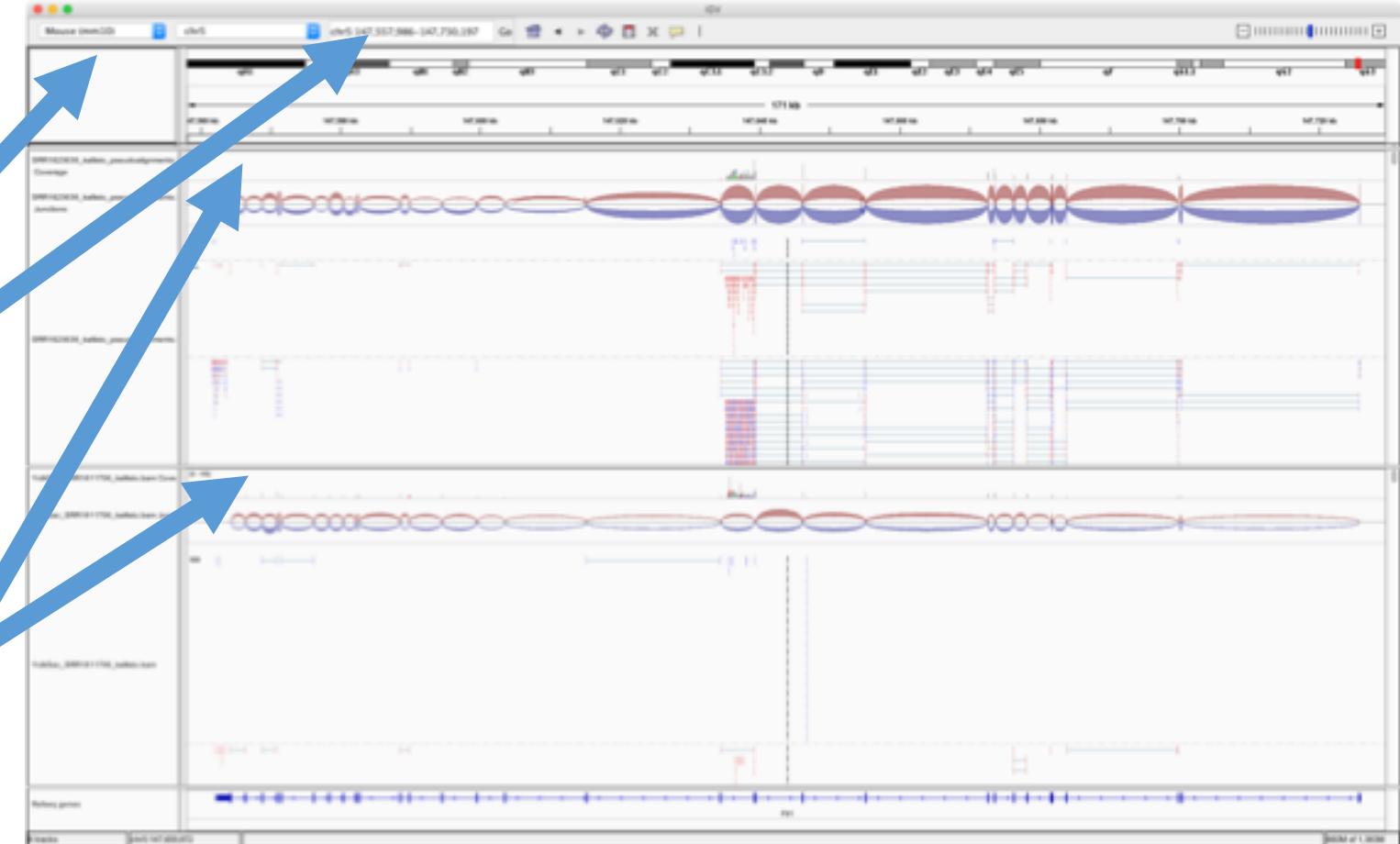
IGV\_Alignments/

Placenta\_SRR1823638\_kallisto.bam  
~~Placenta\_SRR1823638\_kallisto.bam.bai~~  
YolkSac\_SRR1811706\_kallisto.bam  
~~YolkSac\_SRR1811706\_kallisto.bam.bai~~

Select: Mouse(mm10)

Search for Flt1

Can you see the difference in coverage?



[http://www.ensembl.org/Mus\\_musculus/Location/Genome](http://www.ensembl.org/Mus_musculus/Location/Genome)

Enter gene to search here  
e.g. Trf



The Ensembl genome browser for *Mus musculus* (GRCm38) - [Gene ID: 114](#) - [Search](#)

Gene-based analysis

- Summary
- Update variants
- Transcript comparison
- Protein comparison
- Gene details
- Sequence
- Secondary Structure
- Expression summaries
- Evolution
- Ontologies
  - GO Biological process
  - GO Cellular component
  - GO Function
- Comparative Genomics
  - Genomic alignments
  - Conservation scores
  - Gene phylogenetic tree
  - Orthologues
  - Proteins
  - Protein families
- Genetic Variation
  - Variant lists
  - Variant tracks
  - SNP variants
- External data
  - Gene expression
  - Proteins
  - ChIP-seq
- Configure file input
- Add your data
- Export data
- Show this page
- Bookmark this page

Gene: Trf [ensembl.geneview](#)

Description: transforming factor beta 1, isoform 1B

Synonyms: Trf, CtnnB, Ctnnb1, Ctnnb1ip1, Ctnnb1ip2

Location: Chromosome 2, 125,000,000-125,200,000 (mouse.m38), GRCm38.Chr2:125,000,000-125,200,000

About this gene: This gene has 11 transcripts (splice variants, 10 protein-coding, 1 non-coding). It is a member of the [Ensembl protein family](#) and is associated with [Lphenotrans](#).

Transcripts: [View transcript details](#)

Genebody columns (1 hidden)	View																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
Name	Transcript ID	Sp.	Protein	Biotype	ENST00000000000	ENST00000000001	ENST00000000002	ENST00000000003	ENST00000000004	ENST00000000005	ENST00000000006	ENST00000000007	ENST00000000008	ENST00000000009	ENST00000000010	ENST00000000011	ENST00000000012	ENST00000000013	ENST00000000014	ENST00000000015	ENST00000000016	ENST00000000017	ENST00000000018	ENST00000000019	ENST00000000020	ENST00000000021	ENST00000000022	ENST00000000023	ENST00000000024	ENST00000000025	ENST00000000026	ENST00000000027	ENST00000000028	ENST00000000029	ENST00000000030	ENST00000000031	ENST00000000032	ENST00000000033	ENST00000000034	ENST00000000035	ENST00000000036	ENST00000000037	ENST00000000038	ENST00000000039	ENST00000000040	ENST00000000041	ENST00000000042	ENST00000000043	ENST00000000044	ENST00000000045	ENST00000000046	ENST00000000047	ENST00000000048	ENST00000000049	ENST00000000050	ENST00000000051	ENST00000000052	ENST00000000053	ENST00000000054	ENST00000000055	ENST00000000056	ENST00000000057	ENST00000000058	ENST00000000059	ENST00000000060	ENST00000000061	ENST00000000062	ENST00000000063	ENST00000000064	ENST00000000065	ENST00000000066	ENST00000000067	ENST00000000068	ENST00000000069	ENST00000000070	ENST00000000071	ENST00000000072	ENST00000000073	ENST00000000074	ENST00000000075	ENST00000000076	ENST00000000077	ENST00000000078	ENST00000000079	ENST00000000080	ENST00000000081	ENST00000000082	ENST00000000083	ENST00000000084	ENST00000000085	ENST00000000086	ENST00000000087	ENST00000000088	ENST00000000089	ENST00000000090	ENST00000000091	ENST00000000092	ENST00000000093	ENST00000000094	ENST00000000095	ENST00000000096	ENST00000000097	ENST00000000098	ENST00000000099	ENST000000000100	ENST000000000101	ENST000000000102	ENST000000000103	ENST000000000104	ENST000000000105	ENST000000000106	ENST000000000107	ENST000000000108	ENST000000000109	ENST000000000110	ENST000000000111	ENST000000000112	ENST000000000113	ENST000000000114	ENST000000000115	ENST000000000116	ENST000000000117	ENST000000000118	ENST000000000119	ENST000000000120	ENST000000000121	ENST000000000122	ENST000000000123	ENST000000000124	ENST000000000125	ENST000000000126	ENST000000000127	ENST000000000128	ENST000000000129	ENST000000000130	ENST000000000131	ENST000000000132	ENST000000000133	ENST000000000134	ENST000000000135	ENST000000000136	ENST000000000137	ENST000000000138	ENST000000000139	ENST000000000140	ENST000000000141	ENST000000000142	ENST000000000143	ENST000000000144	ENST000000000145	ENST000000000146	ENST000000000147	ENST000000000148	ENST000000000149	ENST000000000150	ENST000000000151	ENST000000000152	ENST000000000153	ENST000000000154	ENST000000000155	ENST000000000156	ENST000000000157	ENST000000000158	ENST000000000159	ENST000000000160	ENST000000000161	ENST000000000162	ENST000000000163	ENST000000000164	ENST000000000165	ENST000000000166	ENST000000000167	ENST000000000168	ENST000000000169	ENST000000000170	ENST000000000171	ENST000000000172	ENST000000000173	ENST000000000174	ENST000000000175	ENST000000000176	ENST000000000177	ENST000000000178	ENST000000000179	ENST000000000180	ENST000000000181	ENST000000000182	ENST000000000183	ENST000000000184	ENST000000000185	ENST000000000186	ENST000000000187	ENST000000000188	ENST000000000189	ENST000000000190	ENST000000000191	ENST000000000192	ENST000000000193	ENST000000000194	ENST000000000195	ENST000000000196	ENST000000000197	ENST000000000198	ENST000000000199	ENST000000000200	ENST000000000201	ENST000000000202	ENST000000000203	ENST000000000204	ENST000000000205	ENST000000000206	ENST000000000207	ENST000000000208	ENST000000000209	ENST000000000210	ENST000000000211	ENST000000000212	ENST000000000213	ENST000000000214	ENST000000000215	ENST000000000216	ENST000000000217	ENST000000000218	ENST000000000219	ENST000000000220	ENST000000000221	ENST000000000222	ENST000000000223	ENST000000000224	ENST000000000225	ENST000000000226	ENST000000000227	ENST000000000228	ENST000000000229	ENST000000000230	ENST000000000231	ENST000000000232	ENST000000000233	ENST000000000234	ENST000000000235	ENST000000000236	ENST000000000237	ENST000000000238	ENST000000000239	ENST000000000240	ENST000000000241	ENST000000000242	ENST000000000243	ENST000000000244	ENST000000000245	ENST000000000246	ENST000000000247	ENST000000000248	ENST000000000249	ENST000000000250	ENST000000000251	ENST000000000252	ENST000000000253	ENST000000000254	ENST000000000255	ENST000000000256	ENST000000000257	ENST000000000258	ENST000000000259	ENST000000000260	ENST000000000261	ENST000000000262	ENST000000000263	ENST000000000264	ENST000000000265	ENST000000000266	ENST000000000267	ENST000000000268	ENST000000000269	ENST000000000270	ENST000000000271	ENST000000000272	ENST000000000273	ENST000000000274	ENST000000000275	ENST000000000276	ENST000000000277	ENST000000000278	ENST000000000279	ENST000000000280	ENST000000000281	ENST000000000282	ENST000000000283	ENST000000000284	ENST000000000285	ENST000000000286	ENST000000000287	ENST000000000288	ENST000000000289	ENST000000000290	ENST000000000291	ENST000000000292	ENST000000000293	ENST000000000294	ENST000000000295	ENST000000000296	ENST000000000297	ENST000000000298	ENST000000000299	ENST000000000300	ENST000000000301	ENST000000000302	ENST000000000303	ENST000000000304	ENST000000000305	ENST000000000306	ENST000000000307	ENST000000000308	ENST000000000309	ENST000000000310	ENST000000000311	ENST000000000312	ENST000000000313	ENST000000000314	ENST000000000315	ENST000000000316	ENST000000000317	ENST000000000318	ENST000000000319	ENST000000000320	ENST000000000321	ENST000000000322	ENST000000000323	ENST000000000324	ENST000000000325	ENST000000000326	ENST000000000327	ENST000000000328	ENST000000000329	ENST000000000330	ENST000000000331	ENST000000000332	ENST000000000333	ENST000000000334	ENST000000000335	ENST000000000336	ENST000000000337	ENST000000000338	ENST000000000339	ENST000000000340	ENST000000000341	ENST000000000342	ENST000000000343	ENST000000000344	ENST000000000345	ENST000000000346	ENST000000000347	ENST000000000348	ENST000000000349	ENST000000000350	ENST000000000351	ENST000000000352	ENST000000000353	ENST000000000354	ENST000000000355	ENST000000000356	ENST000000000357	ENST000000000358	ENST000000000359	ENST000000000360	ENST000000000361	ENST000000000362	ENST000000000363	ENST000000000364	ENST000000000365	ENST000000000366	ENST000000000367	ENST000000000368	ENST000000000369	ENST000000000370	ENST000000000371	ENST000000000372	ENST000000000373	ENST000000000374	ENST000000000375	ENST000000000376	ENST000000000377	ENST000000000378	ENST000000000379	ENST000000000380	ENST000000000381	ENST000000000382	ENST000000000383	ENST000000000384	ENST000000000385	ENST000000000386	ENST000000000387	ENST000000000388	ENST000000000389	ENST000000000390	ENST000000000391	ENST000000000392	ENST000000000393	ENST000000000394	ENST000000000395	ENST000000000396	ENST000000000397	ENST000000000398	ENST000000000399	ENST000000000400	ENST000000000401	ENST000000000402	ENST000000000403	ENST000000000404	ENST000000000405	ENST000000000406	ENST000000000407	ENST000000000408	ENST000000000409	ENST000000000410	ENST000000000411	ENST000000000412	ENST000000000413	ENST000000000414	ENST000000000415	ENST000000000416	ENST000000000417	ENST000000000418	ENST000000000419	ENST000000000420	ENST000000000421	ENST000000000422	ENST000000000423	ENST000000000424	ENST000000000425	ENST000000000426	ENST000000000427	ENST000000000428	ENST000000000429	ENST000000000430	ENST000000000431	ENST000000000432	ENST000000000433	ENST000000000434	ENST000000000435	ENST000000000436	ENST000000000437	ENST000000000438	ENST000000000439	ENST000000000440	ENST000000000441	ENST000000000442	ENST000000000443	ENST000000000444	ENST000000000445	ENST000000000446	ENST000000000447	ENST000000000448	ENST000000000449	ENST000000000450	ENST000000000451	ENST000000000452	ENST000000000453	ENST000000000454	ENST000000000455	ENST000000000456	ENST000000000457	ENST000000000458	ENST000000000459	ENST000000000460	ENST000000000461	ENST000000000462	ENST000000000463	ENST000000000464	ENST000000000465	ENST000000000466	ENST000000000467	ENST000000000468	ENST000000000469	ENST000000000470	ENST000000000471	ENST000000000472	ENST000000000473	ENST000000000474	ENST000000000475	ENST000000000476	ENST000000000477	ENST000000000478	ENST000000000479	ENST000000000480	ENST000000000481	ENST000000000482	ENST000000000483	ENST000000000484	ENST000000000485	ENST000000000486	ENST000000000487	ENST000000000488	ENST000000000489	ENST000000000490	ENST000000000491	ENST000000000492	ENST000000000493	ENST000000000494	ENST000000000495	ENST000000000496	ENST000000000497	ENST000000000498	ENST000000000499	ENST000000000500	ENST000000000501	ENST000000000502	ENST000000000503	ENST000000000504	ENST000000000505	ENST000000000506	ENST000000000507	ENST000000000508	ENST000000000509	ENST000000000510	ENST000000000511	ENST000000000512	ENST000000000513	ENST000000000514	ENST000000000515	ENST000000000516	ENST000000000517	ENST000000000518	ENST000000000519	ENST000000000520	ENST000000000521	ENST000000000522	ENST000000000523	ENST000000000524	ENST000000000525	ENST000000000526	ENST000000000527	ENST000000000528	ENST000000000529	ENST000000000530	ENST000000000531	ENST000000000532	ENST000000000533	ENST000000000534	ENST000000000535	ENST000000000536	ENST000000000537	ENST000000000538	ENST000000000539	ENST000000000540	ENST000000000541	ENST000000000542	ENST000000000543	ENST000000000544	ENST000000000545	ENST000000000546	ENST000000000547	ENST000000000548	ENST000000000549	ENST000000000550	ENST000000000551	ENST000000000552	ENST000000000553	ENST000000000554	ENST000000000555	ENST000000000556	ENST000000000557	ENST000000000558	ENST000000000559	ENST000000000560	ENST000000000561	ENST000000000562	ENST000000000563	ENST000000000564	ENST000000000565	ENST000000000566	ENST000000000567	ENST000000000568	ENST000000000569	ENST000000000570	ENST000000000571	ENST000000000572	ENST000000000573	ENST000000000574	ENST000000000575	ENST000000000576	ENST000000000577	ENST000000000578	ENST000000000579	ENST000000000580	ENST000000000581	ENST000000000582	ENST000000000583	ENST000000000584	ENST000000000585	ENST000000000586	ENST000000000587	ENST000000000588	ENST000000000589	ENST000000000590	ENST000000000591	ENST000000000592	ENST000000000593	ENST000000000594	ENST000000000595	ENST000000000596	ENST000000000597	ENST000000000598	ENST000000000599	ENST000000000600	ENST000000000601	ENST000000000602	ENST000000000603	ENST000000000604	ENST000000000605	ENST000000000606	ENST000000000607	ENST000000000608	ENST000000000609	ENST000000000610	ENST000000000611	ENST000000000612	ENST000000000613	ENST000000000614	ENST000000000615	ENST000000000616	ENST000000000617	ENST000000000618	ENST000000000619	ENST000000000620	ENST000000000621	ENST000000000622	ENST00

What is the function of Trf?

# DAVID Gene Ontology Analysis

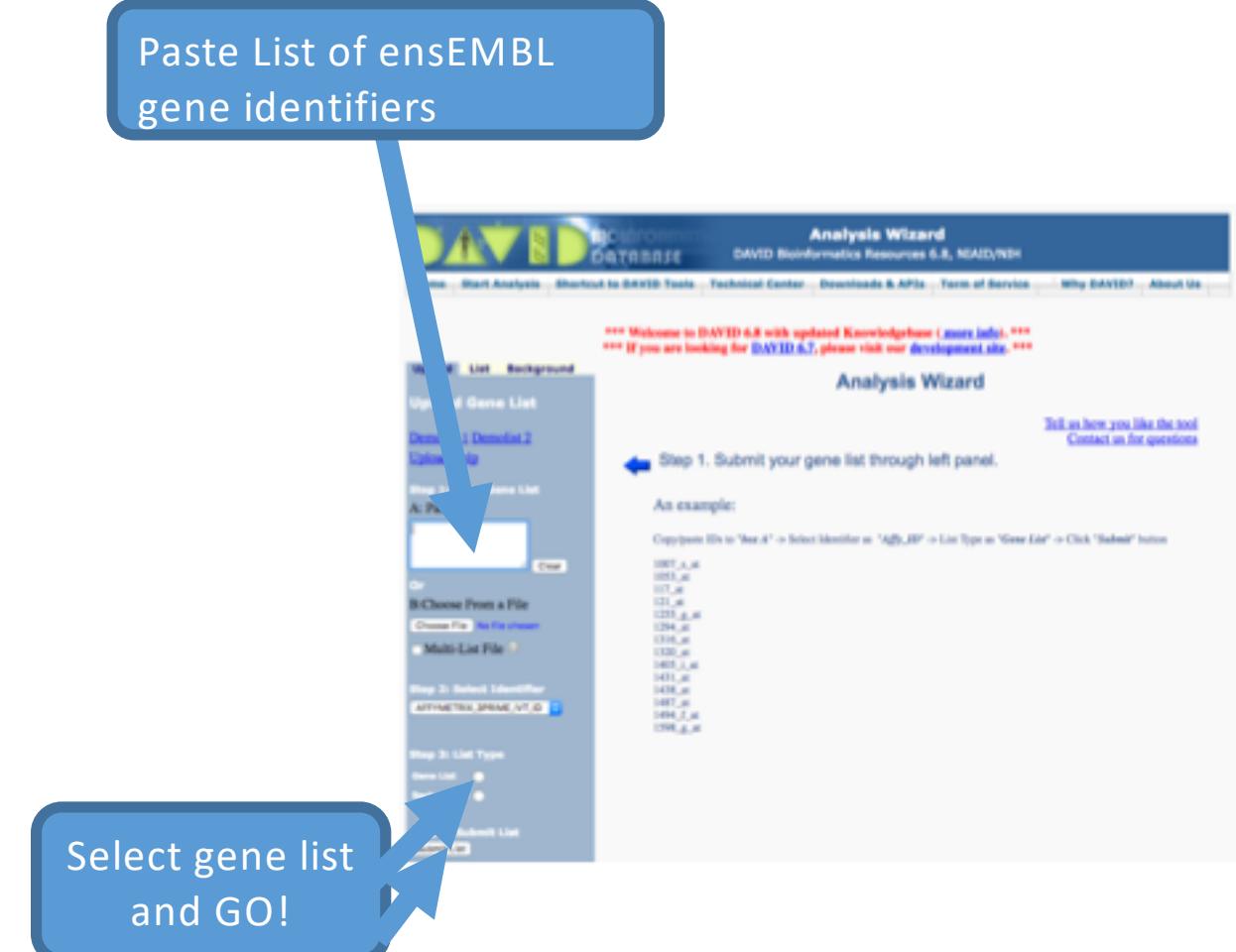
<https://david.ncifcrf.gov/>

Start Analysis



The screenshot shows the DAVID Bioinformatics Resources 6.8 homepage. It features a navigation bar at the top with links like Home, Start Analysis, Shortcut to DAVID Tools, Technical Center, Downloads & APIs, Terms of Service, Why DAVID?, and About Us. Below the navigation is a banner with a green arrow pointing right and the text "Welcome to DAVID 6.8 with updated Knowledgebase (more info)." and "If you are looking for DAVID 6.7, please visit our development site." A main content area titled "Welcome to DAVID 6.8" includes a search bar, a "2863 - 28E?" section, and a "What's Important in DAVID?" section with a bulleted list. To the right is a "Statistics of DAVID" section with a bar chart showing citations from 2000 to 2013. At the bottom are three small screenshots labeled "Screen Shot 1", "Screen Shot 2", and "Screen Shot 3".

Paste List of ensEMBL gene identifiers



The screenshot shows the DAVID Analysis Wizard interface. It has a header "Analysis Wizard" and a sub-header "DAVID Bioinformatics Resources 6.8, NIAID/NIH". The main area is titled "Analysis Wizard" and contains three steps: "Step 1: Submit your gene list through left panel.", "Step 2: Choose from a file", and "Step 3: Select Identifier". Step 1 is highlighted with a blue arrow. Step 2 shows a file selection dialog with "Multi-List File" selected. Step 3 shows a dropdown menu with "AFFYMETRIX\_SPNAME/VT\_ID" selected. A large blue arrow points from the "Paste List of ensEMBL gene identifiers" callout to the "Step 1" section.

Select gene list  
and GO!



# Reproducible Bioinformatics

## Versioning

If you write code or scripts use a versioning system (a bit like track changes in Word)  
Make it publicly available so people can comment and submit bug reports  
e.g. <http://www.github.com>

## Pipelines

Track program version numbers, consistent processing and reporting  
Avoid manual input of data or settings  
e.g. <http://clusterflow.io> or SnakeMake

## Data Repositories

Upload your published data to GEO, ENA, SRA etc



Dr Russell S. Hamilton

Email: rsh46@cam.ac.uk

Dr Xiaohui Zhao

Email: xz289@cam.ac.uk

Dr Malwina Prater

Email: mn367@cam.ac.uk

## Course Materials:

<https://github.com/darogan/2018-PlacentalBiologyCourse>



License: Attribution-Non Commercial-Share Alike CC BY-NC-SA ( <https://creativecommons.org/licenses/by-nc-sa/> )

Attribution: You must give appropriate credit, provide a link to the license, and indicate if changes were made.

NonCommercial: You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial: You may not use the material for commercial purposes.

ShareAlike: If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.