



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/CTR-BFX/2019-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.

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.



Version 0.1: 20190619



Program

Lecture 1: What is RNA-Seq?

- Introduction to RNA-Seq and sequencer options
- From Sequencer to Quality Control and Aligning reads

Practical 1:

- From FASTQ to BAM

Lecture 2: Gene Counts to hypothesis testing

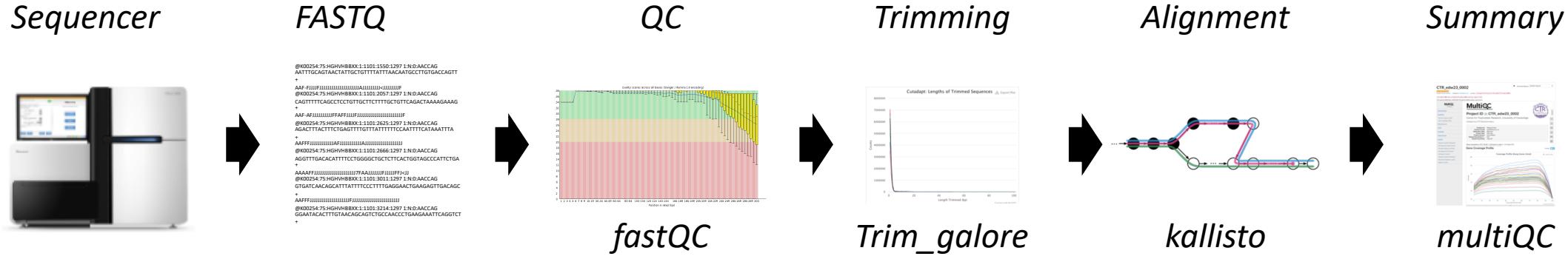
- Experimental Design
- Gene quantification
- Differential Gene Expression Analysis

Practical 2:

- BAM to DEGs

Lecture 2: Gene Counts to hypothesis testing

- Experimental Design
- Gene quantification
- Differential Gene Expression Analysis



Differential Gene Expression Analysis

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



4x YolkSac

7x Placenta

Identify genes differentially expressed between the two groups



How Many Replicates?

How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use?

NICHOLAS J. SCHURCH,^{1,6} PIETÁ SCHOFIELD,^{1,2,6} MAREK GIERLIŃSKI,^{1,2,6} CHRISTIAN COLE,^{1,6}
ALEXANDER SHERSTNEV,^{1,6} VIJENDER SINGH,² NICOLA WROBEL,³ KARIM GHARBI,³
GORDON G. SIMPSON,⁴ TOM OWEN-HUGHES,² MARK BLAXTER,³ and GEOFFREY J. BARTON^{1,2,5}

¹Division of Computational Biology, College of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

²Division of Gene Regulation and Expression, College of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

³Edinburgh Genomics, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

⁴Division of Plant Sciences, College of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

⁵Division of Biological Chemistry and Drug Discovery, College of Life Sciences, University of Dundee, Dundee DD1 5EH, United Kingdom

<https://doi.org/10.1261/rna.053959.115>

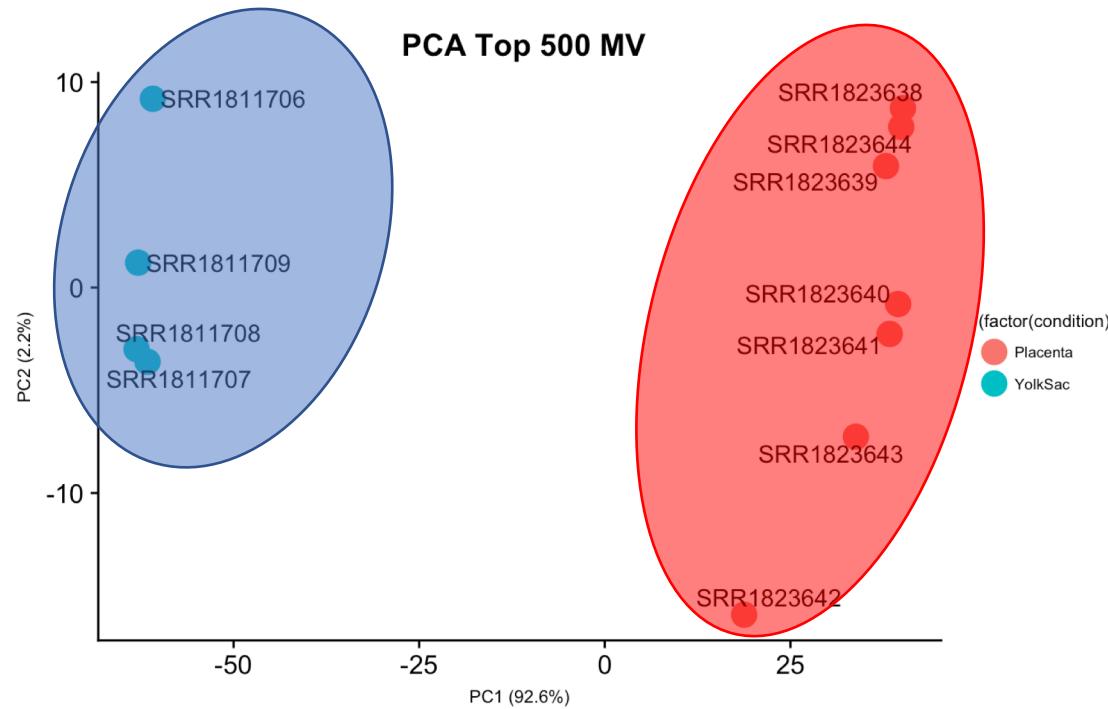
Started with 48 replicates, removed replicated and investigated the power of statistical analysis

Conclusion/Recommendations:

- Ideally at least six biological replicates should be used. Commonly three are used.
- Rising to at least 12 when it is important to identify SDE genes for all fold changes.
- If fewer than 12 replicates are used:
edgeR and DESeq2 the leading tools for DGE

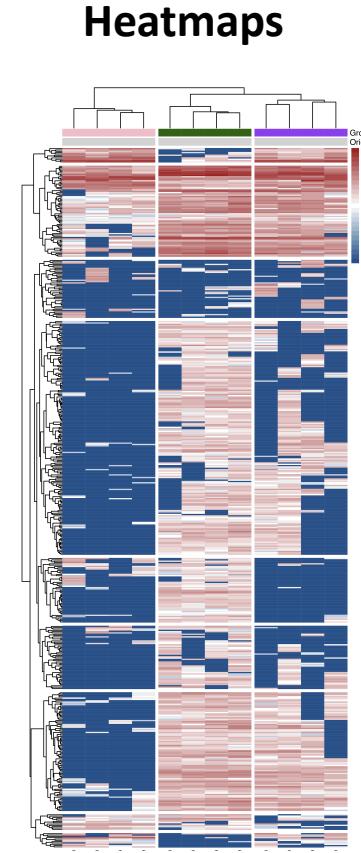
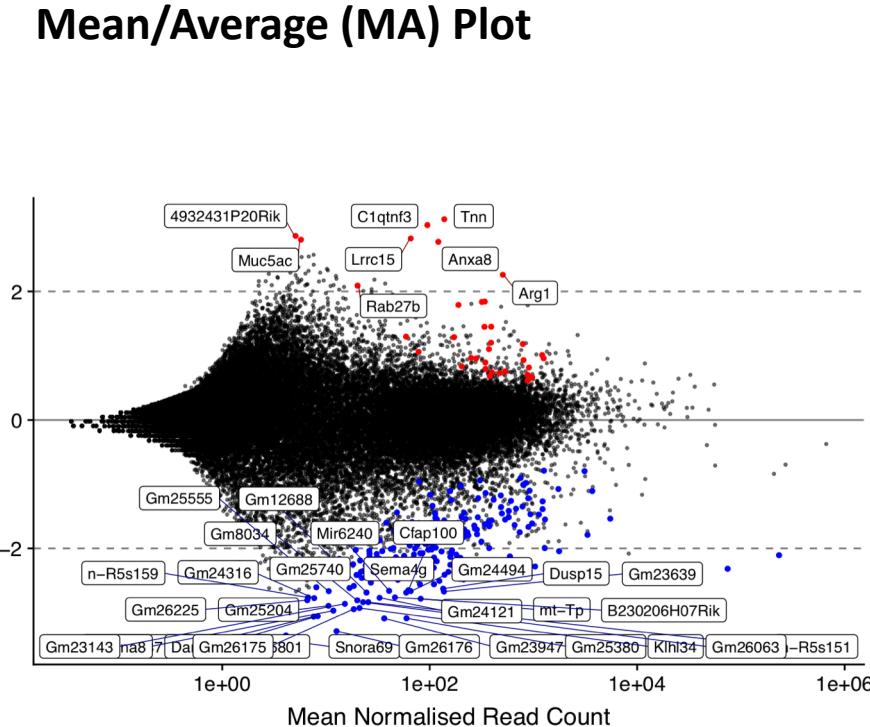
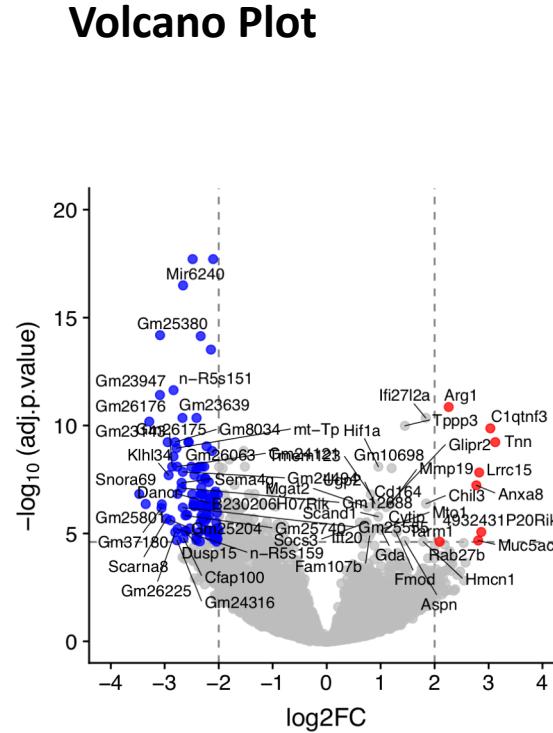
Differential Gene Expression Analysis

Principal Component Analysis

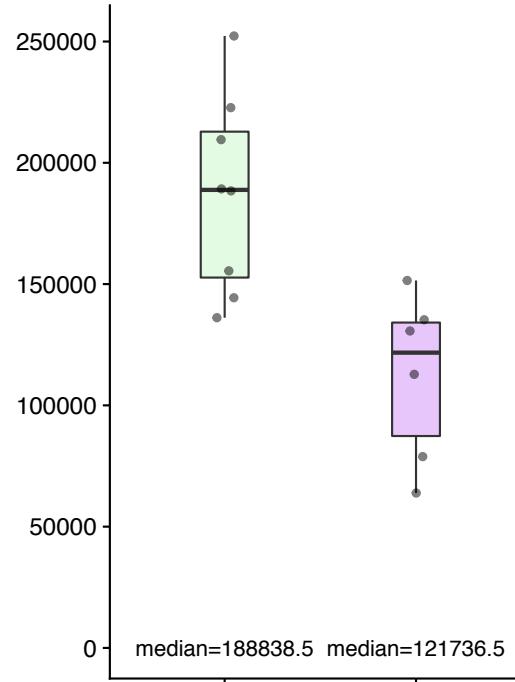


Do the samples cluster as expected?
Does the comparison make sense?
Is there a separation between groups?

Visualising Differentially Expressed Genes



Visualising Differentially Expressed Genes



Good to check the normalized counts of a gene identified to be differentially expressed

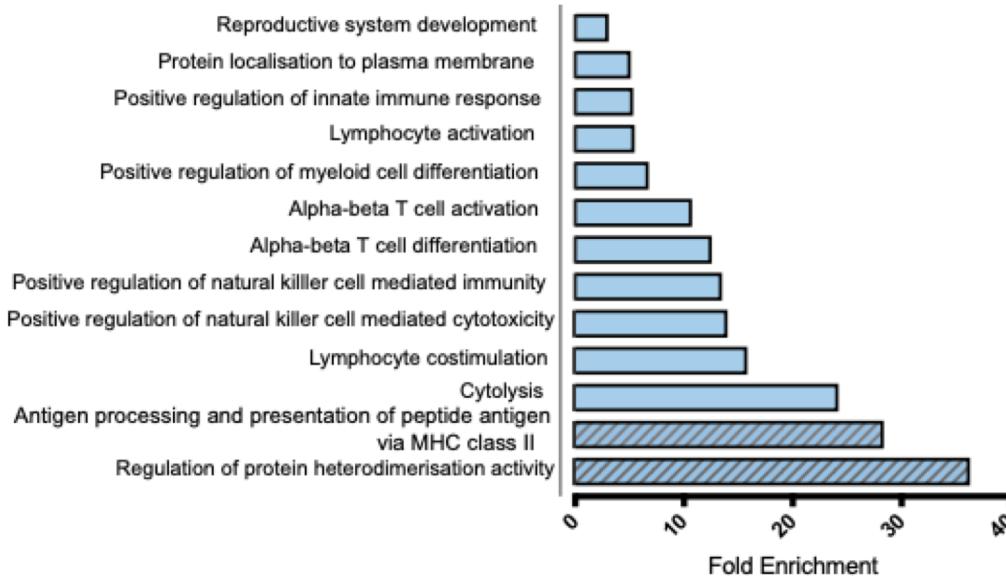
Gene Enrichment

Do the identified differentially expressed genes have anything in common?

Pathway enrichment

Protein Interactions

Compartment / Location



Filipovic, I, et al (2018) Molecular definition of group 1 innate lymphoid cells in the mouse uterus. *Nature Communications*, **9:1**, 4492



Reproducible Bioinformatics

Versioning

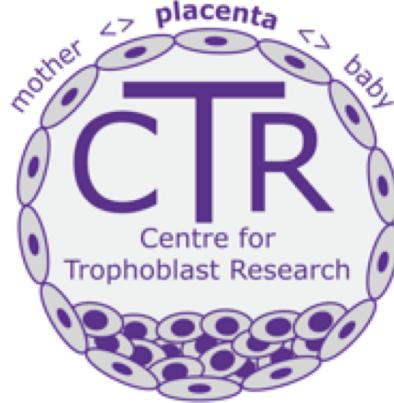
If you write code or scripts use a versioning system (a bit like track changes in Word)
Make it publicly available so people can comment and submit bug reports
e.g. <http://www.github.com>

Pipelines

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

Data Repositories

Upload your published data to GEO, ENA, SRA etc



Dr Russell S. Hamilton

Email: rsh46@cam.ac.uk

Dr Malwina Prater

Email: mn367@cam.ac.uk

Dr Xiaohui Zhao

Email: xz289@cam.ac.uk

Course Materials:

<https://github.com/CTR-BFX/2019-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.

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.