

Gene expression analysis

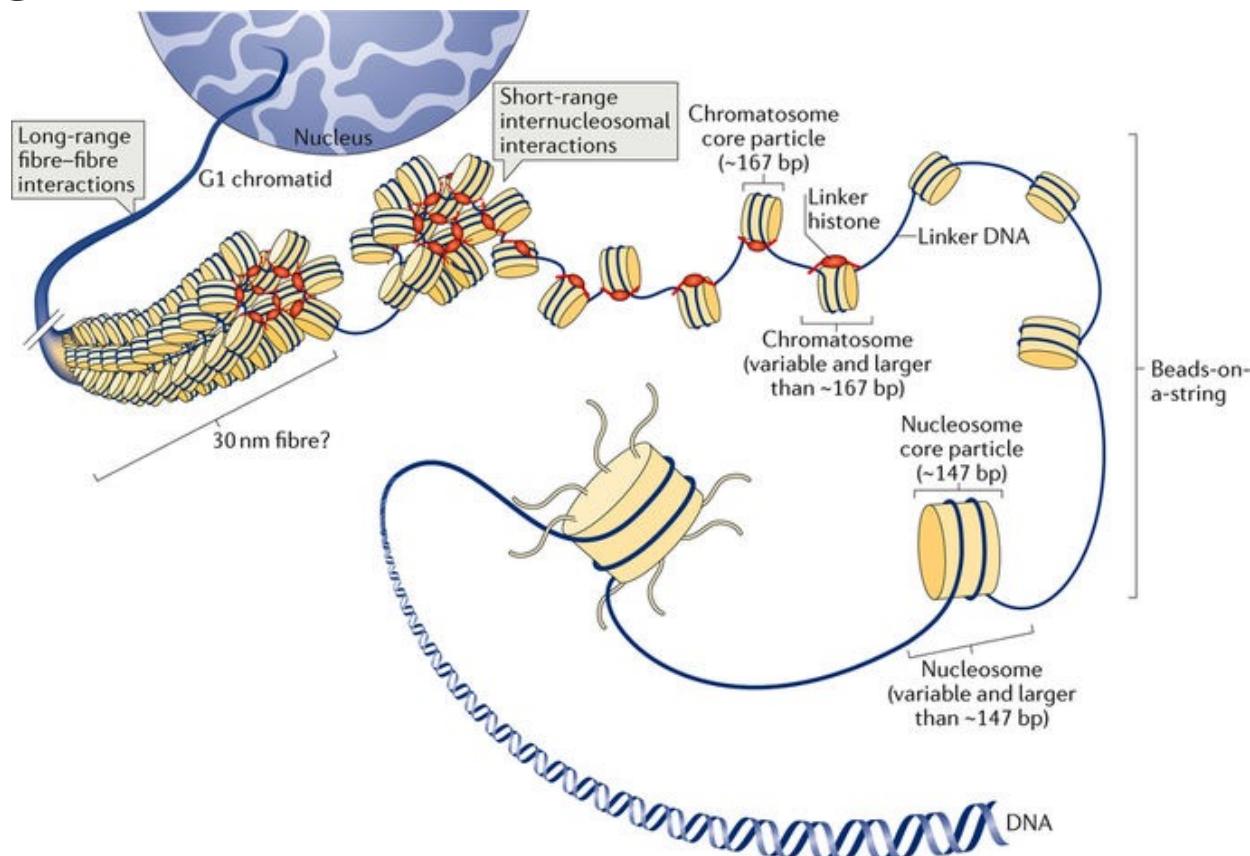
PSL374

Brian Cox, PhD

Instructor

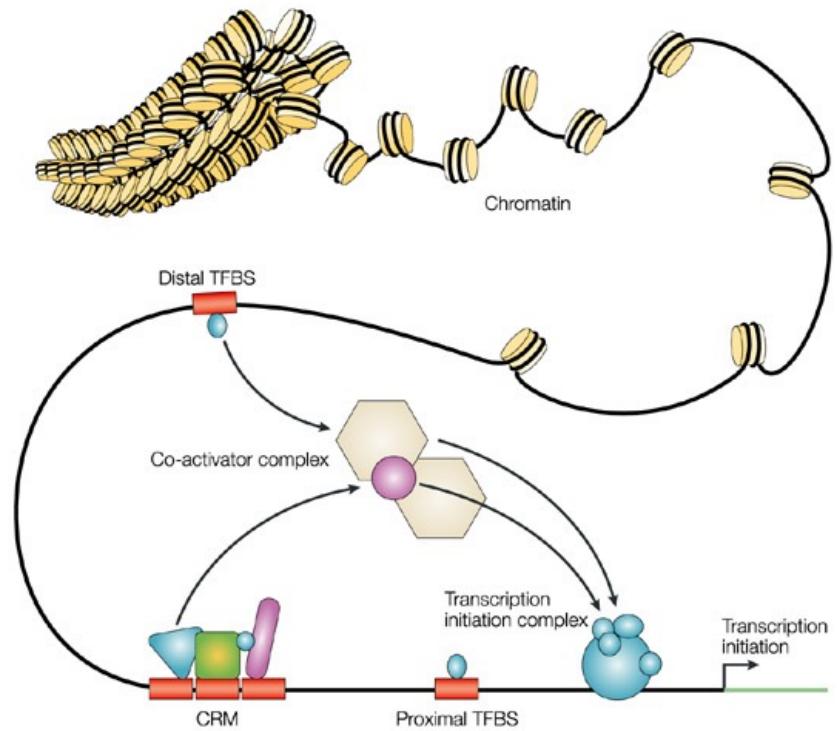


Genome



Gene expression is regulated

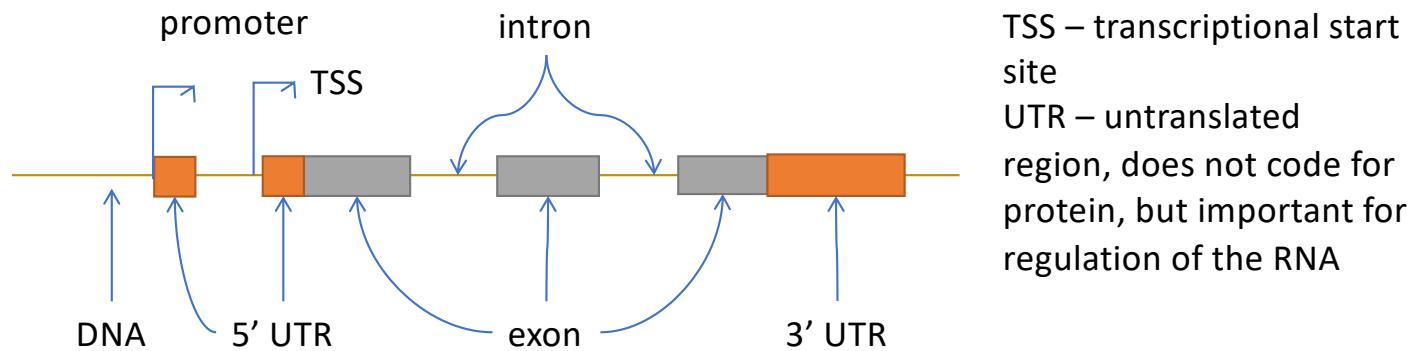
- Different regions of the genome
- Co-ordinate expression
- Activators, repressors
- Enhancers, promoter



Wyeth W. Wasserman & Albin Sandelin
Nature Reviews Genetics 5, 276-287 (2004)

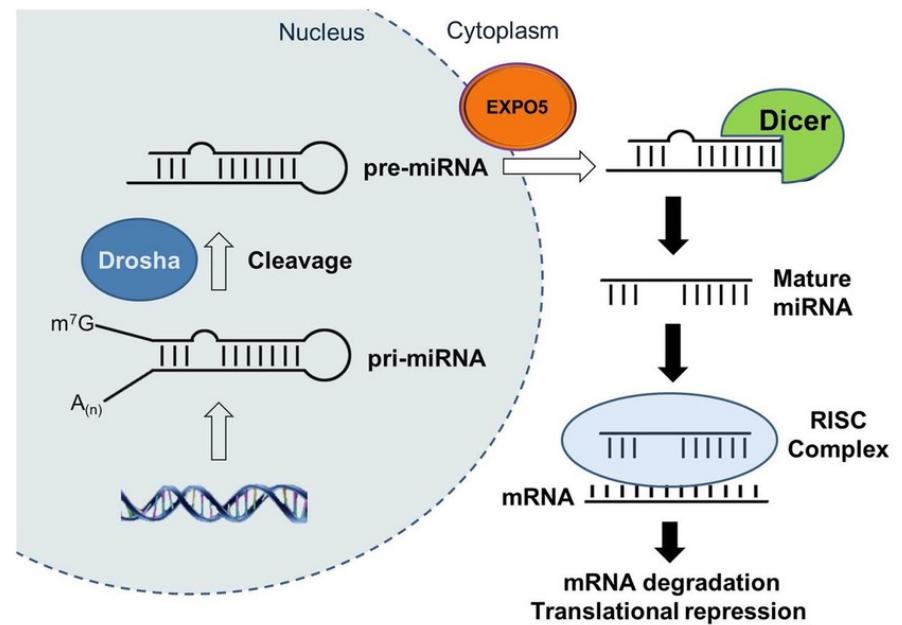
Transcriptome, the transcribed genome

- In the case of gene expression we are looking only at the transcribed regions (exons)
 - Total of all expressed RNA is the transcriptome
 - Transcribed but, spliced out are the introns
 - Introns can contain other used transcription products like small RNAs (microRNAs and piwi RNA)



Transcriptome is more than protein coding

- RNA can be one of two kinds
 - Protein coding
 - Loads on to the ribosome for translation
 - Non-coding
 - Long non-coding RNA, 1000s of bases long
 - XIST (X inactivation in females), rRNA (functional parts of the ribosome)
 - Small non-coding RNAs, 100s of bases when transcribed but processed by nucleases into 19-14 base final product
 - microRNAs, Piwi, tRNA
 - Pseudogenes, variable 100s to 1000s of bases long
 - Lost reading frame, will not make a protein
 - Spliced and unspliced mRNAs



Gene expression analysis

- By sequencing the RNA what kinds of analyses can be performed?

Experimental design

A quality experiment involves 3 main things:

1) The Question



Questions on the unique characteristics and functions of a population, tissues, cells etc

2) The Sample Size
(power)



Do you have enough samples to detect a difference? Considers noise of samples and technique.

3) The Controls



How do you know the experiment worked/failed?
That the effect is biologically relevant?

Power analysis/ Sample size

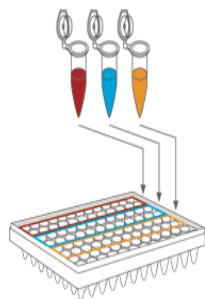
- Two issues, ***detection*** and ***quantification***
- **Detection**
 - Is the gene or transcript present/absent?
- **Quantification**
 - Is there more or less between groups
- Power to detect differences
 - This is proportional to the variance (μ)
 - High variance needs more samples to be certain of a difference
 - Low variance needs fewer samples
 - Often we start with a few sample to determine how many samples are needed

Comparison of techniques

qPCR



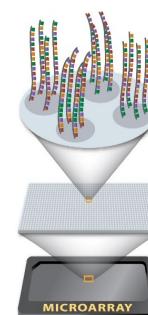
Targeted method
known



Microarray



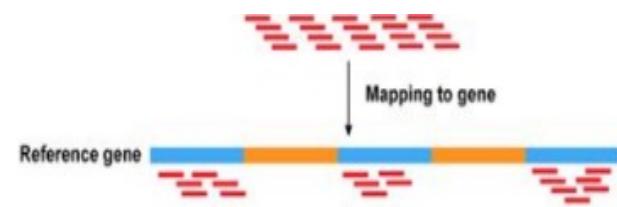
Genome wide
Known genes



RNA-
Sequencing



Genome wide
Known and
discovery

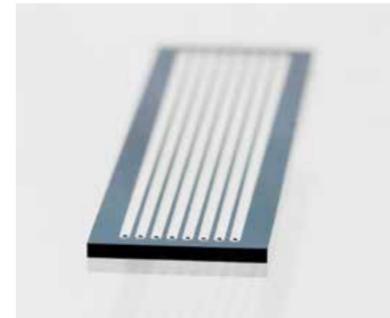


Sequencing, The next generation

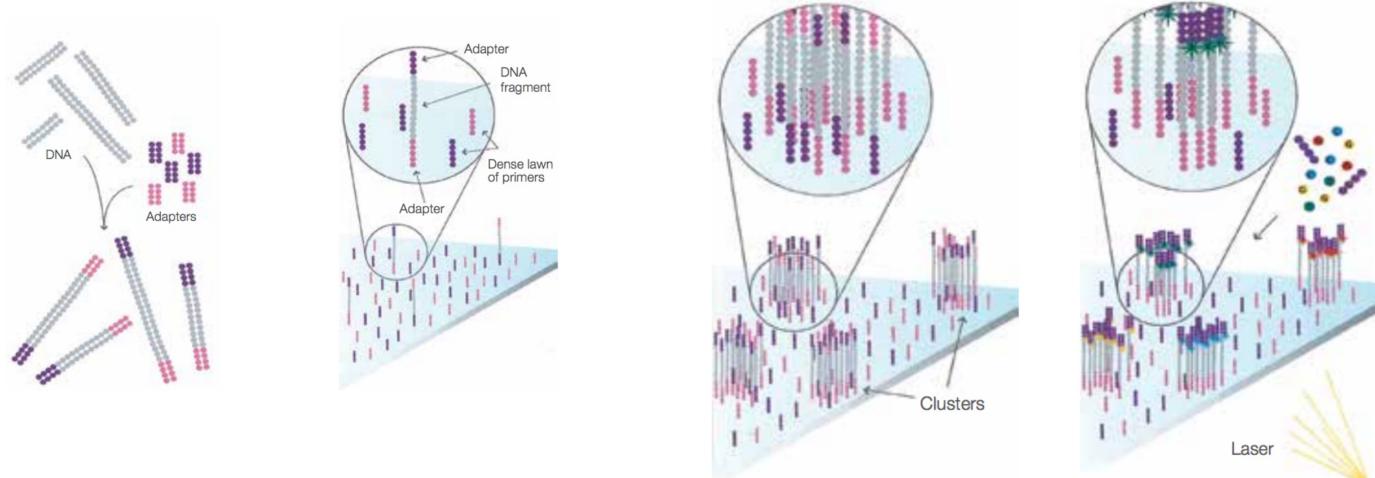
- Massively parallel, millions to billions of reads per analysis run
 - A read is a single stretch of sequence
 - 75-150 bases long
- 500 million reads for \$2500
 - 200,000 reads/dollar
- Typical read by Sanger method \$5 each
 - But covers 500-1000 bases

Sequencing, Illumina

- On board amplification
 - In a flow cell
- Uses PCR
- Fluorescence based sequencing
- Imaging to detect sequence

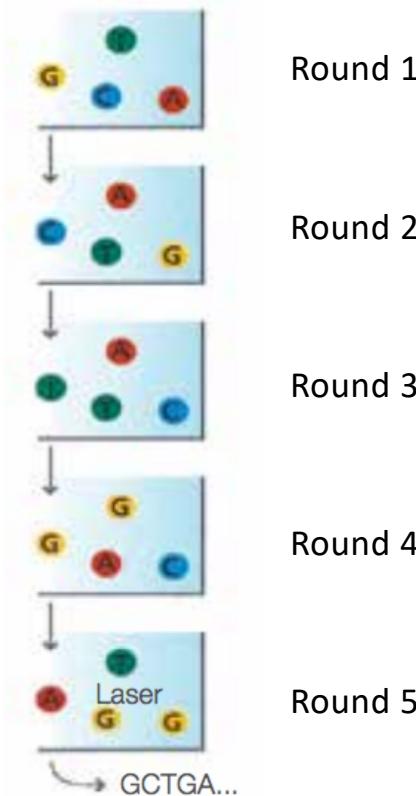


Sequencing, Illumina



- Sequences are randomly spaced on a flat surface
- Use PCR to build the sequence colony

Sequencing cycle



- Labeled nucleotides are added
- Polymerase incorporates
- Each colony should have one new nucleotide added
- The flow cell is imaged
- The dye is cleaved and washed
- Repeat the cycle
- The order of the images is the sequence

FASTQ

- Raw sequence data with quality scores

Identifier	• @SRR566546.970 HWUSI-EAS1673_11067_FC7070M:4:1:2299:1109 length=50
Sequence	• TTGCCTGCCTATCATTAGTGCCTGTGAGGTGGAGATGTGAGGATCAGT
'+' sign	• +
Quality scores	• hhhhhhhhhhhhhhhhhhhhhhhhhfffffe'ee['X]b[d[ed'[Y[^Y
Identifier	• @SRR566546.971 HWUSI-EAS1673_11067_FC7070M:4:1:2374:1108 length=50
Sequence	• GATTGTATGAAAGTATAACACTAAACTGCAGGTGGATCAGAGTAAGTC
'+' sign	• +
Quality scores	• hhhhgfhcghghggfcffdhfeyhhcehdchhdhaehffffde'bVd

Alignment

Shakespearomics

- **Reads**

```
ds, Romans, count  
ns, countrymen, le  
Friends, Rom  
send me your ears;  
crymen, lend me
```



- **Overlaps**

```
Friends, Rom  
ds, Romans, count  
ns, countrymen, le  
crymen, lend me  
send me your ears;
```

- **Majority rule**

```
Friends, Romans, countrymen, lend me your ears;
```

FASTQ
Format

Sequencing reads

Align reads to genome



Many Mapping tools have been developed to assist with mapping of transcriptome data. These trade speed and accuracy.

*BOWTIE2, BWA, TOPHAT,
SHRIMP, STAR, GSNAP,
HISAT2, SAILFISH*

Count table

countData

gene	ctrl_1	ctrl_2	exp_1	exp_1
geneA	10	11	56	45
geneB	0	0	128	54
geneC	42	41	59	41
geneD	103	122	1	23
geneE	10	23	14	56
geneF	0	1	2	0
...
...
...

colData

id	treatment	sex
ctrl_1	control	male
ctrl_2	control	female
exp_1	treatment	male
exp_2	treatment	female

Sample names:

ctrl_1, ctrl_2, exp_1, exp_2

countData is the count matrix

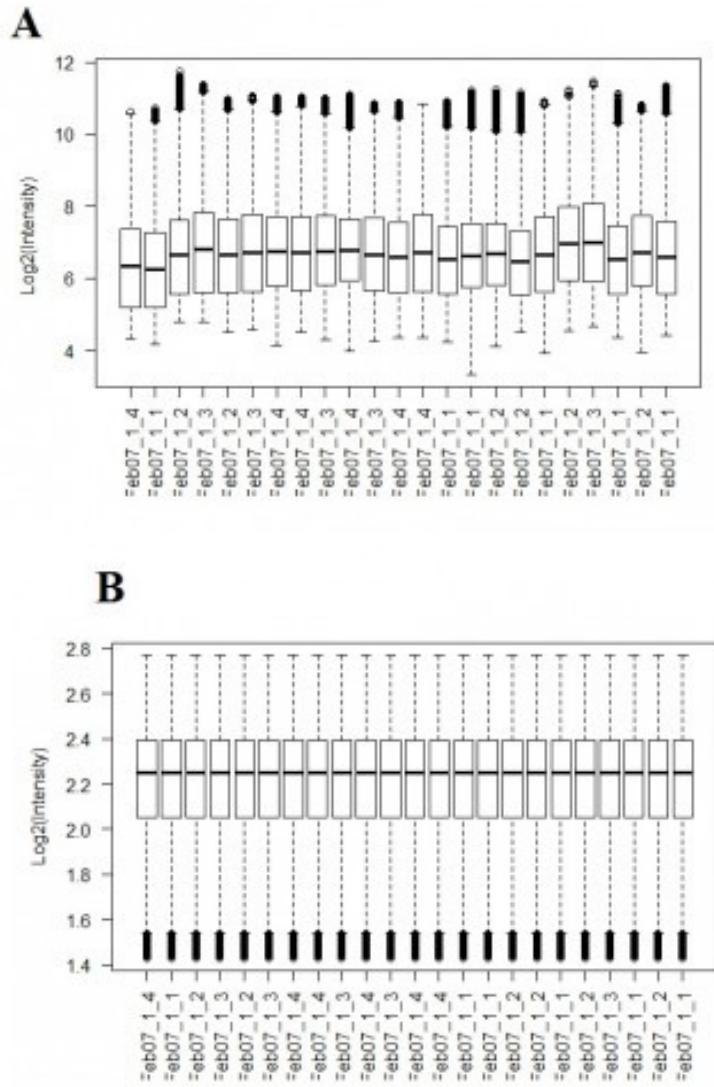
(number of reads mapping to each gene for each sample)

colData describes metadata about the *columns* of countData

First column of colData must match column names of countData (-1st)

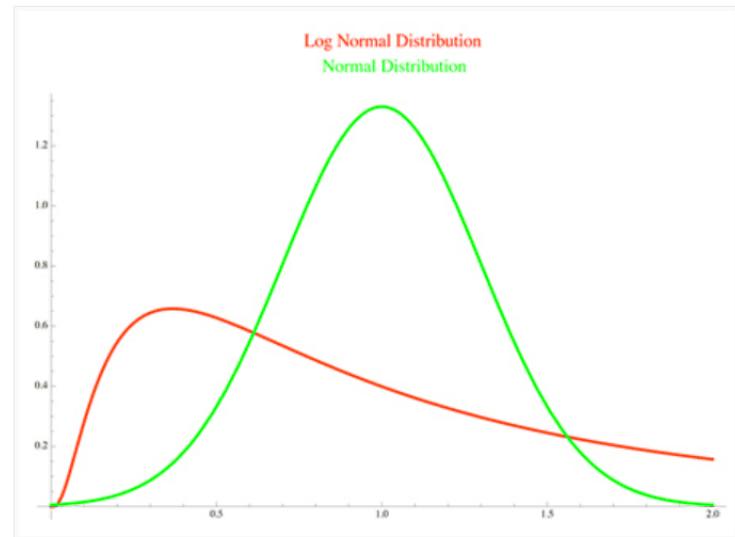
Gene expression concepts

- **Normalize or Standardize**
 - Data collected under different conditions
 - Range (min and max) and means may be different
 - Must normalize to make equivalent
 - **Normalcy**
 - Is the distribution normal
 - Transform, typical is \log_2
 - **Differential expression**
 - Many methods, linear models, t-test
 - Must correct or adjust the test statistic, FDR, Bonferroni, etc.
 - Too many tests, inflates the false positives



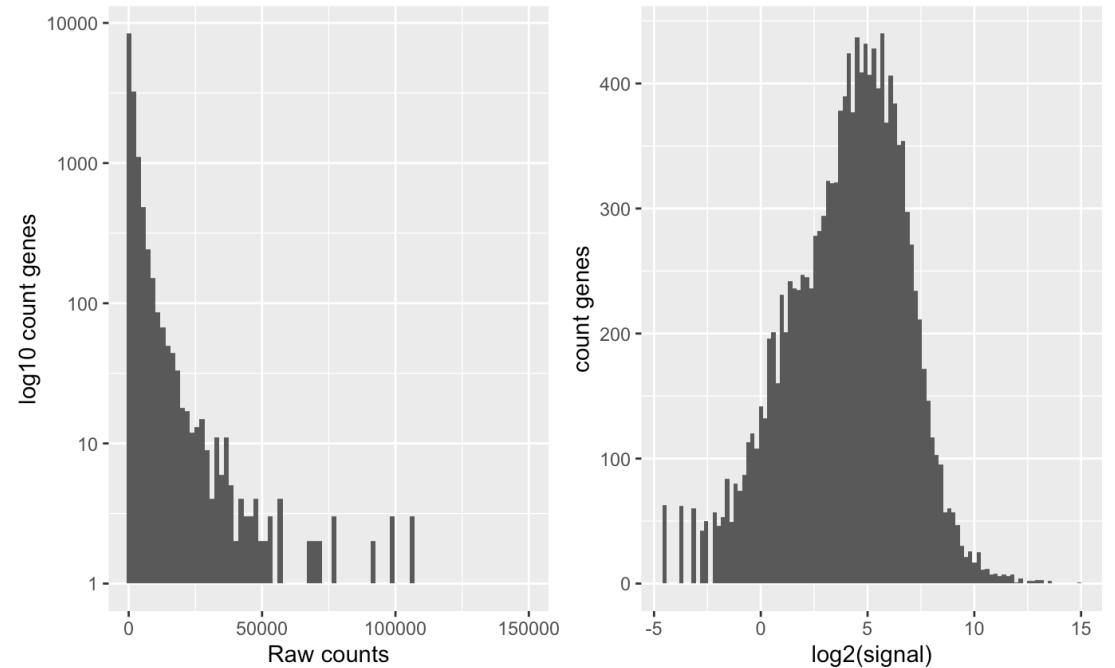
Gene expression concepts

- Normalize
 - Data collected under different conditions
 - Range (min and max) and means may be different
 - Must normalize to make equivalent
- **Normalcy**
 - Is the distribution normal
 - Transform, typical is \log_2
- Differential expression
 - Many methods, linear models, t-test
 - Must correct or adjust the test statistic, FDR, Bonferroni, etc.
 - Too many tests, inflates the false positives



Quantification

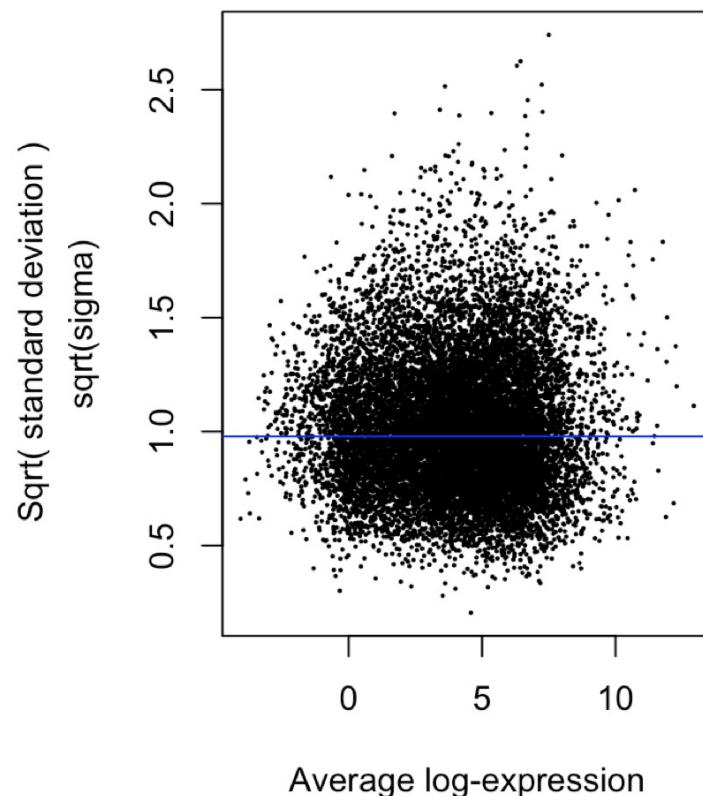
- RNA-sequencing
 - Counting, integer
 - heteroskedasticity, variance is dependent on the mean



Quantification

- RNA-sequencing
 - Counting, integer
 - heteroskedasticity, variance is dependent on the mean

Final model: Mean-variance trend



Batch effect

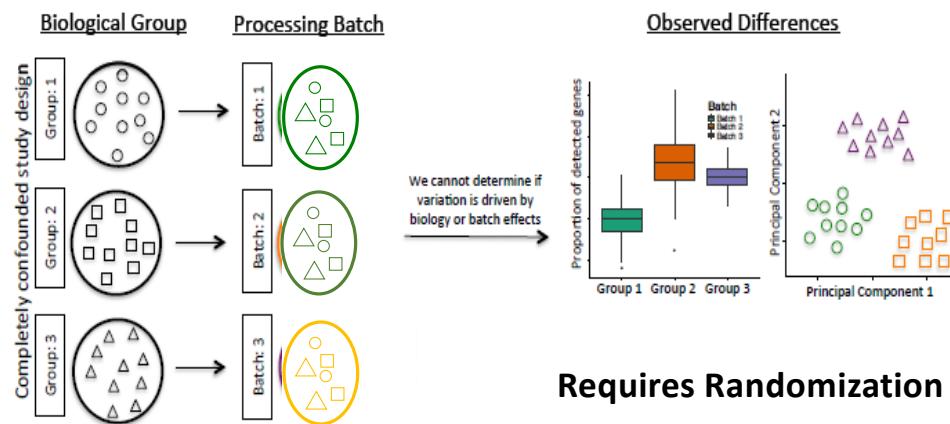
- After normalization can still have an issue
- Batch
 - A grouping of samples that were analyzed at the same time
- The batch effect
 - Presence of multiple batches
 - Can affect the measurement of some genes in a batch
 - Each batch can have its own
 - Possible sources are
 - Different day, Lot of reagents, Person performing experiment, Different instrument (same model)

Experimental design

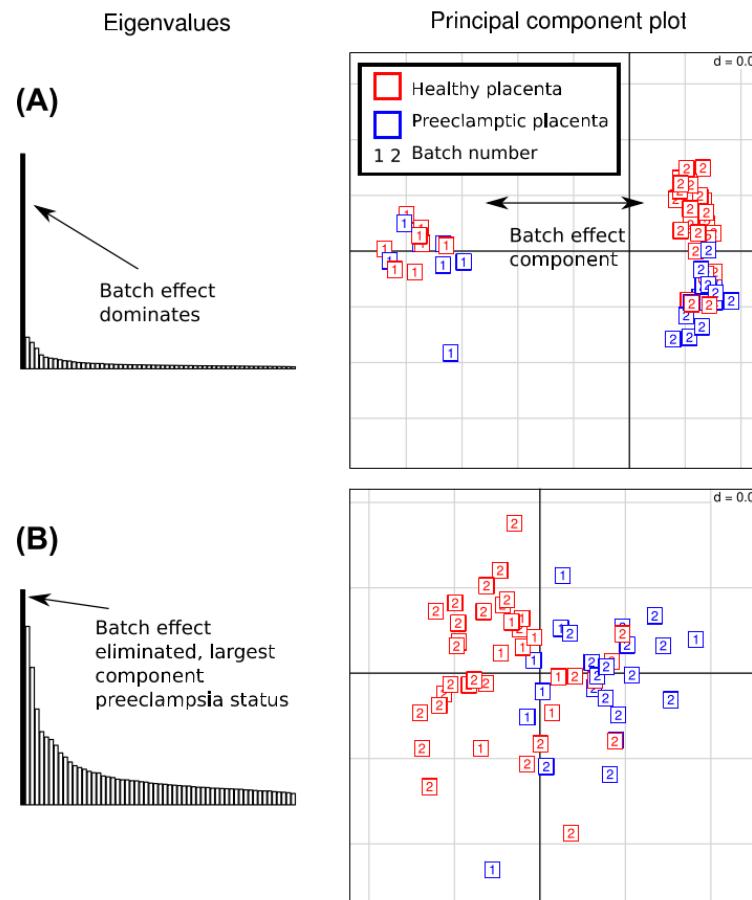
- Experimental design can be used to prevent or compensate for the batch effect
- Each batch should have all phenotypes represented
- If each batch is a unique sample type, controls, treatment) batch effect can not be fixed
 - Both batch and treatment are the same, **Co-variant**

Batch Effect:

Technical variation in the processing sample (Batch)



A batch effect



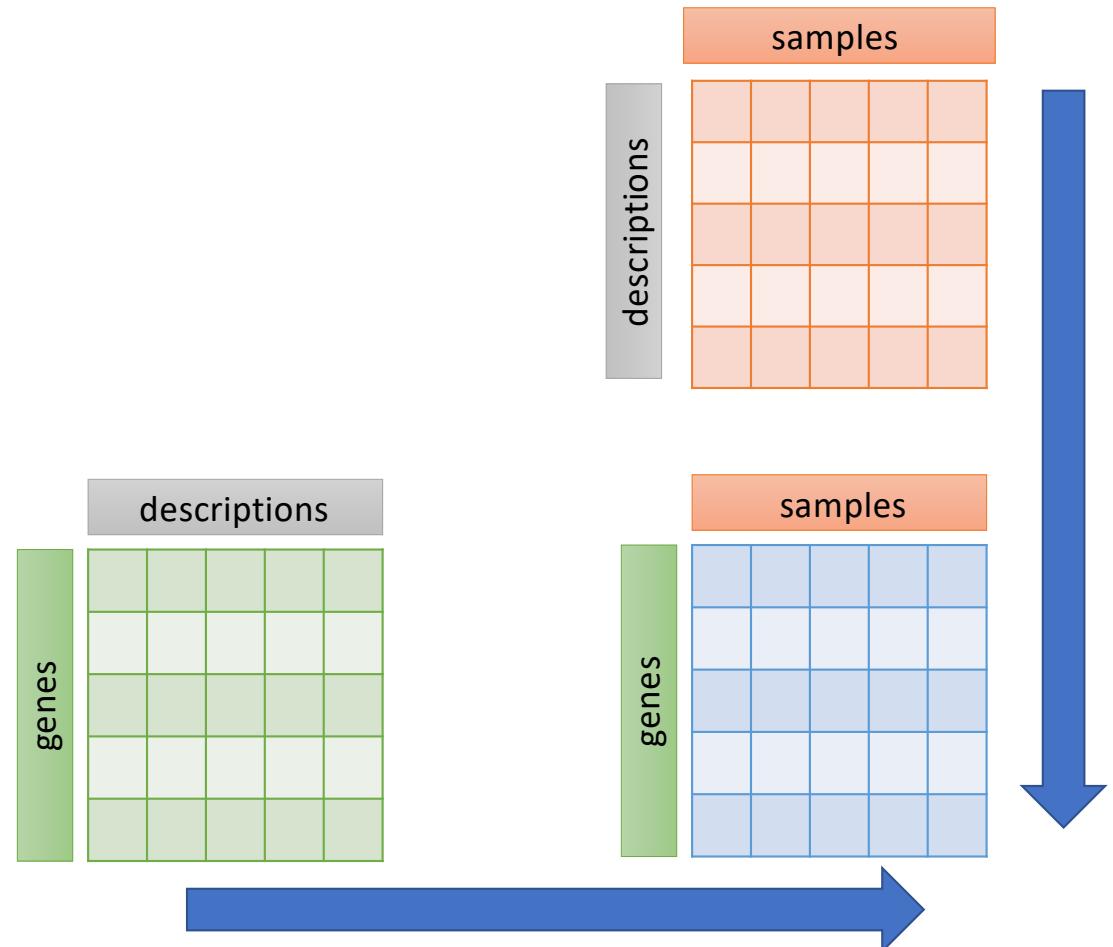
- The histogram is the PC % variance
- The scatter plot are the samples
 - Coloured by phenotype
 - Numbered by batch
- After correction the batches intermingle and the dominate component is the phenotype
- Can be removed as each batch has both phenotypes represented

General object structure

- RNA-seq uses the DGEList (differential gene expression list)
 - Used by edgeR, limma
 - There are other object types in other packages
- Consists of three main elements/tables
 - Expression data matrix, columns of samples rows of genes
 - Phenotype data frame, columns of attributes rows of samples
 - Feature data frame, columns of attributes, rows of genes

DGEList object

- The tables are held dynamically
- Processing the samples table automatically processes the gene expression table
- Similar for the feature table and expression table



DGEList – Differential Gene Expression List

- Stores similar information
 - Uses a list style object
 - names() accesses the positions
 - Samples is equivalent to phenotype table
 - Counts is same as assay data
 - Genes is same as feature data
- > names(x)
[1] "samples" "counts" "genes"

DGEList

- Multiple list slots
- The samples slot is updated by different functions
- lib.size is calculated automatically
- Norm.factors are calculated by a function

```
An object of class "DGEList"
$samples
  files group lib.size norm.factors lane
10_6_5_11 GSM1545535_10_6_5_11.txt    LP 32832881  0.8957309 L004
9_6_5_11  GSM1545536_9_6_5_11.txt    ML 35296067  1.0349196 L004
purep53   GSM1545538_purep53.txt Basal 57089259  1.0439552 L004
JMS8-2    GSM1545539_JMS8-2.txt Basal 51304051  1.0405040 L006
JMS8-3    GSM1545540_JMS8-3.txt  ML 75724862  1.0323599 L006
JMS8-4    GSM1545541_JMS8-4.txt  LP 60457787  0.9223424 L006
JMS8-5    GSM1545542_JMS8-5.txt Basal 55014064  0.9836603 L006
JMS9-P7c  GSM1545544_JMS9-P7c.txt  ML 21282257  1.0827381 L008
JMS9-P8c  GSM1545545_JMS9-P8c.txt  LP 19938942  0.9792607 L008

$counts
  Samples
Tags   10_6_5_11 9_6_5_11 purep53 JMS8-2 JMS8-3 JMS8-4 JMS8-5 JMS9-P7c JMS9-P8c
497097      1       2     342     526      3      3    535      2      0
27395       431     771    1368    1268    1564    769    818    468    342
18777       768    1722    2517    1923    3865   1888   1830   1246    693
21399       810     977    2472    1870    2251   1716   1932    756    619
58175       452     358      17     14    622    571     12   203    224
14160 more rows ...

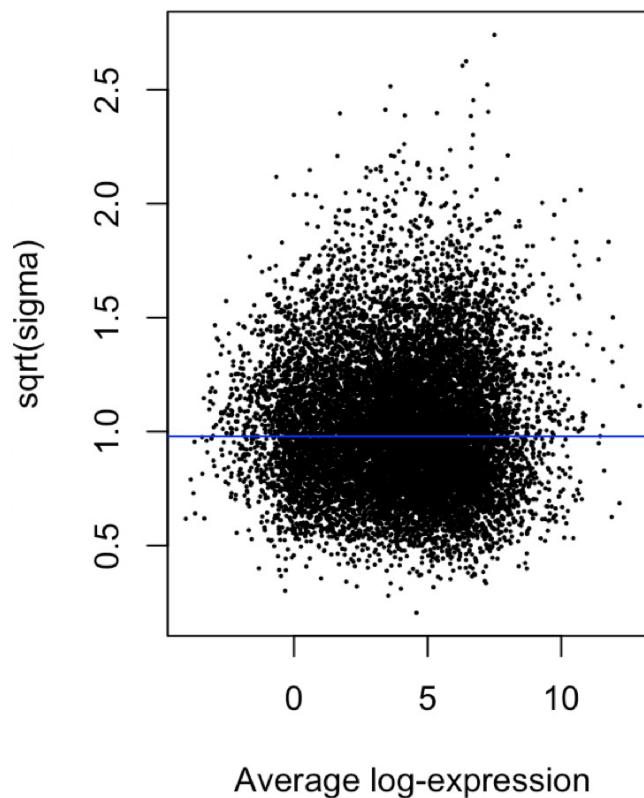
$genes
  ENTREZID SYMBOL TXCHROM
1 497097  Xkr4  chr1
2 100503874 Gm19938 <NA>
3 100038431 Gm10568 <NA>
4 19888    Rp1   chr1
5 20671    Sox17 chr1
27174 more rows ...
```

Limma

The voom object

- Voom is similar to the DGEList object
- An EList (Expression List)
- Genes is the genes table,
- Targets is the samples table
- E is the expression matrix
- Weights are to transform the data
- Design is the experiment design, samples batches etc.

Final model: Mean-variance trend



Design matrix

- Limma uses linear models
 - $y=mx+b$
- Sample annotation used
- Categorical or continuous variables
- Solving for
 - Differentiation expression
 - E.g., Healthy vs Sick
 - Adjusting (for a technical or confounding variable, batch, weight, age etc)

```
```{r}
the experiment design to be tested, 0 is the intercept,
ch1.2 is the pathology and ch1.5 is the batch
design <- model.matrix(~0 + gset$characteristics_ch1.2 , gset)
#name the columns
colnames(design) <- c("Control", "Preeclamptic")

#check out the design object
design|```
```

```

| | Control | Preeclamptic |
|-----------|---------|--------------|
| GSM635904 | 0 | 1 |
| GSM635905 | 0 | 1 |
| GSM635906 | 0 | 1 |
| GSM635907 | 0 | 1 |
| GSM635908 | 0 | 1 |
| GSM635909 | 1 | 0 |
| GSM635910 | 1 | 0 |
| GSM635911 | 1 | 0 |
| GSM635912 | 0 | 1 |
| GSM635913 | 1 | 0 |
| GSM635914 | 0 | 1 |
| GSM635915 | 0 | 1 |
| GSM635916 | 1 | 0 |
| GSM635917 | 1 | 0 |

Contrast matrix

- The question you are asking
- Data is log transformed
- $\text{Log}(a/b)$
- $\text{Log}(a) - \text{Log}(b)$
- All values in the design matrix are modelled
 - If not part of the contrast they are used as adjustments/corrections
- Order affects direction of change
- What is up or down?

```
```{r}
cont.matrix <- makeContrasts(Preeclamptic-Control, levels=design)
cont.matrix
````
```

| Levels | Contrasts | |
|--------------|--------------|-----------|
| | Preeclamptic | - Control |
| Control | | -1 |
| Preeclamptic | | 1 |

Differential expression

- Once groups/clusters are defined need to understand the difference in expression
- Challenge of sequencing data
 - Counting
 - Not continuous Gaussian data
 - Discrete
 - More related to Poisson or negative binomial
 - Cannot do an ordinary T-test

Differential expression

- Methods
 - DESeq2, EdgeR, Voom + others
- Each uses a different method to model the data and separate technical variation from biological variation
- Differences have a statistical test
 - This should use a correction for multiple testing
- Different count levels have different variance, noise in their measurement
 - Referred to as over-dispersion
 - Need to find the relationship between mean of the expression and the variance

Differential expression

- Lots of tests
- Need to reconsider errors
 - Type I, False positives
 - Type II, False negatives
- If you test 1 comparison, a p-value of 0.05 means
 - A 5% chance that this could be discovered by chance given the data's distributions.
 - Note reducing type I error (lower p values) increases type II error
- Family wise error rate
- Adjust the p-value to account for all of these tests

| | | Reality | |
|---------------|----------|---|---|
| | | Positive | Negative |
| Study Finding | Positive | True Positive
(Power)
$(1-\beta)$ | False Positive
Type I Error
(α) |
| | Negative | False Negative
Type II Error
(β) | True Negative |

Family wise error rate

$$1-(.95)^n$$

N is the number of tests

2 tests is a 0.0975 frequency of false discovery

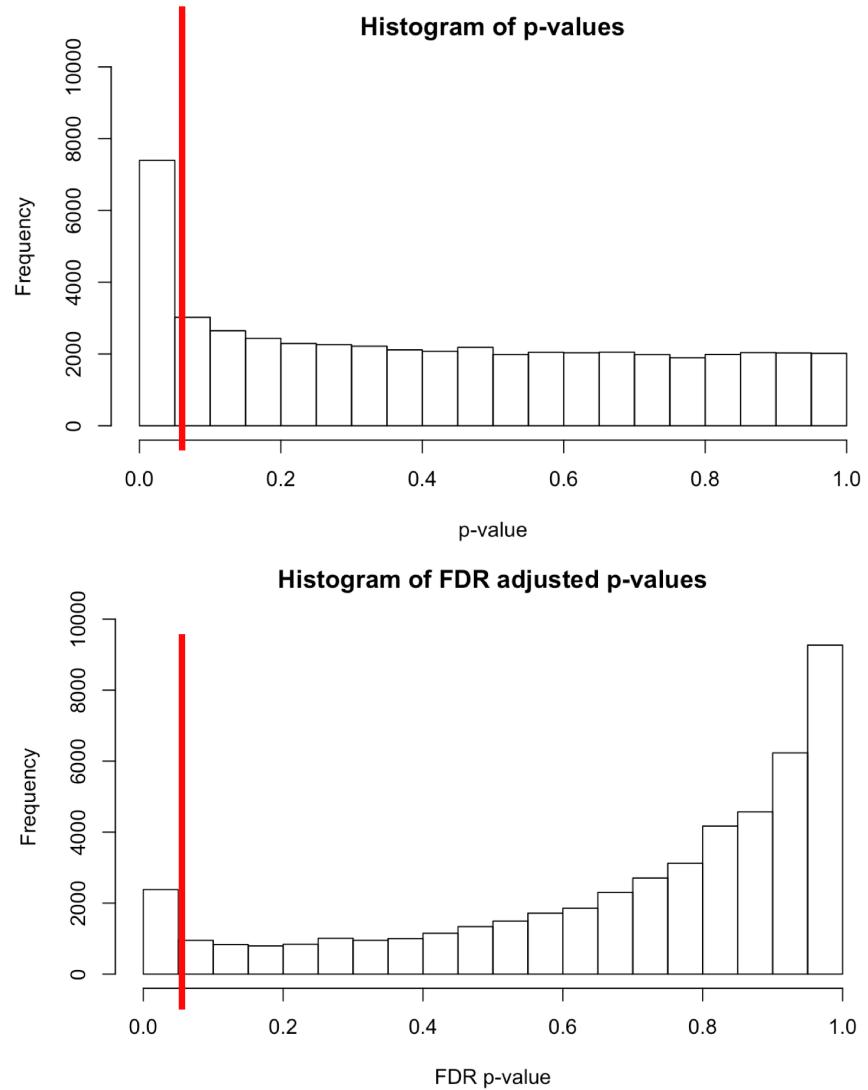
Differential expression

- Log of a fold change
- Statistical values, key is the adjusted p-value
 - Correction for multiple testing

| Symbol | logFC | AveExpr | t | P.Value | adj.P.Val | B |
|---------------|-------|---------|-------|----------|-----------|-------|
| DKFZp586I1420 | 0.36 | 6.69 | 8.33 | 1.21E-11 | 5.89E-07 | 16.02 |
| TRIM24 | 0.56 | 8.70 | 7.76 | 1.17E-10 | 2.86E-06 | 13.91 |
| SIAE | 0.45 | 6.30 | 7.60 | 2.23E-10 | 3.12E-06 | 13.31 |
| TUBA1A | -0.55 | 10.25 | -7.56 | 2.60E-10 | 3.12E-06 | 13.17 |
| MPHOSPH1 | 0.30 | 6.28 | 7.51 | 3.21E-10 | 3.12E-06 | 12.97 |
| PPIG | 0.49 | 6.85 | 7.27 | 8.15E-10 | 5.77E-06 | 12.10 |
| HSP90B1 | 0.58 | 11.29 | 7.27 | 8.29E-10 | 5.77E-06 | 12.09 |
| ENG | 0.92 | 10.66 | 7.08 | 1.73E-09 | 9.65E-06 | 11.40 |
| ERCC5 | 0.44 | 8.14 | 7.05 | 1.95E-09 | 9.65E-06 | 11.29 |
| ZNF33B | 0.49 | 7.08 | 7.05 | 1.98E-09 | 9.65E-06 | 11.28 |

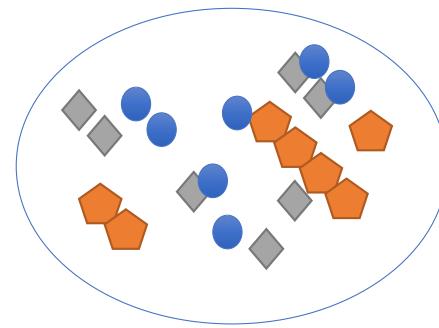
A tail of two p-values

- Adjusting can be as simple as dividing
- $0.05/20,000 = 2.5 \times 10^{-6}$, new threshold
- But we generally apply other methods like
- FDR, Bonferroni, Holm
- These will give different results
- Considered conservative or liberal



Ontologies

- Genes can be grouped in to sets
- The Gene Ontology
 - Molecular Function
 - Transcription factors, kinases
 - Biological process
 - Involved in spleen development, apoptosis
 - Cellular Component
 - Nuclear or mitochondrial localized
- Genes can be in more than one group



All the differential expressed genes



apoptosis



Transcription factor



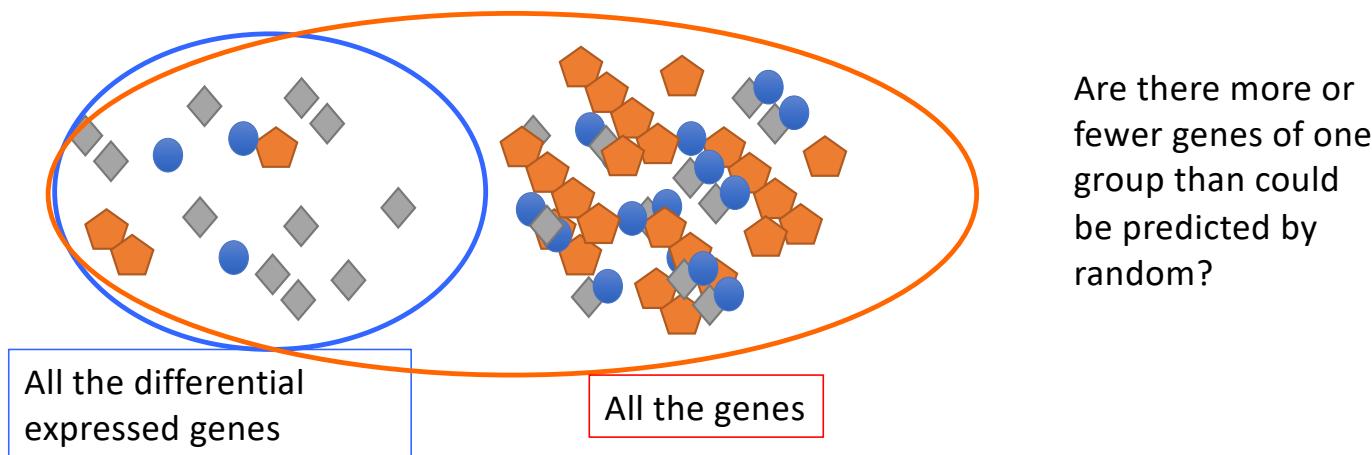
kinase

Gene sets and statistical tests

- Overrepresentation analysis
 - Test for enrichment is a select set or group of genes
 - Possible bias, selection of genes of interest
 - Arbitrary setting of thresholds, fold change and p-values
- Gene set enrichment
 - Test genes collectively as a set with common annotation
 - Individually they may not be enriched (not statistically significant), group statistic
 - Advantage, uses all the data, no selection

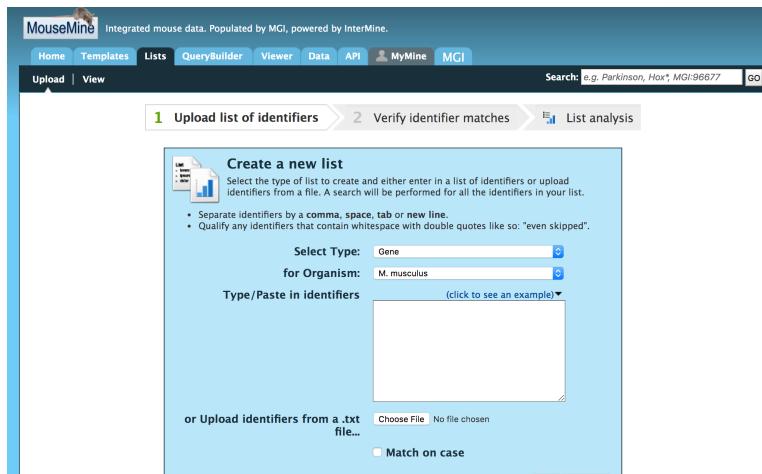
Ontology over-representation

- Of the genes differentially expressed are there more than expected, based on a random sample/draw?
 - Governed by the hypergeometric and binomial probabilities



Over representation

- Mouse mine is a good resource
- Add a list of selected genes
- For example increased after treatment with a drug



Over representation

Mammalian Phenotype Ontology Enrichment
MP terms enriched for items in this list.

Number of Genes in this list not analysed in this widget: 1

Test Correction Max p-value Background population
Holm-Bonferroni 0.05 Default Change

[View](#) [Download](#)

| MP Term | p-Value <i>i</i> | Matches |
|--|------------------|---------|
| abnormal extraembryonic tissue morphology [MP:0002086] | 6.795267e-15 | 16 |
| abnormal trophoblast layer morphology [MP:0005031] | 4.200447e-13 | 11 |
| embryo phenotype [MP:0005380] | 8.889541e-13 | 19 |
| abnormal trophoblast giant cell morphology [MP:0005033] | 4.064482e-12 | 9 |
| abnormal mural trophectoderm morphology [MP:0012057] | 4.481981e-12 | 9 |
| embryonic lethality during organogenesis, complete penetrance [MP:0011098] | 1.004714e-11 | 14 |
| embryonic lethality [MP:0008762] | 2.283471e-11 | 18 |

Gene Ontology Enrichment
GO terms enriched for items in this list.

All items in your list have been analysed.

Test Correction Max p-value Ontology
Holm-Bonferroni 0.05 biological_process

Background population
Default Change

[View](#) [Download](#)

| GO Term | p-Value <i>i</i> | Matches |
|--|------------------|---------|
| cell fate commitment [GO:0045165] | 1.693280e-16 | 14 |
| embryo development [GO:0009790] | 8.372479e-14 | 18 |
| positive regulation of gene expression [GO:0010628] | 1.056936e-11 | 19 |
| embryonic morphogenesis [GO:0048598] | 1.679798e-11 | 14 |
| positive regulation of transcription by RNA polymerase II [GO:0045944] | 1.456743e-10 | 16 |
| positive regulation of transcription, DNA-templated [GO:0045893] | 1.483348e-10 | 17 |
| positive regulation of nucleic acid-templated | 3.009109e-10 | 17 |

Gene set enrichment

- The camera() function in the limma library
- Rank genes in order of expression difference
- Test is on the distribution of the genes in the set
- Are the genes mostly higher or lower than expected?

| Gene | Fold change |
|------|-------------|
| ABC | 4.5 |
| DBC | 3.9 |
| NGF | 3.8 |
| | |
| XCD | -3.3 |
| AAD | -3.5 |
| | |

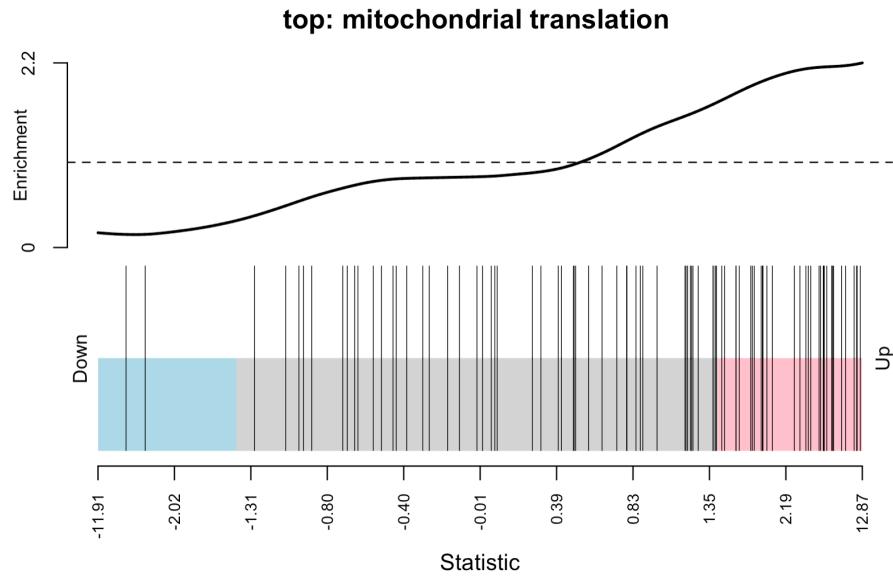
Enrichment Table

| | NGenes
dbl | Direction
chr | PValue
dbl | FDR
dbl |
|---|----------------------|-------------------------|----------------------|-------------------|
| lipoprotein particle%GOcc%GO:1990777 | 17 | Down | 2.859611e-07 | 0.002378921 |
| plasma lipoprotein particle%GOcc%GO:0034358 | 17 | Down | 2.859611e-07 | 0.002378921 |
| hemoglobin complex%GOcc%GO:0005833 | 4 | Up | 3.924802e-07 | 0.002378921 |
| haptoglobin binding%GOfm%GO:0031720 | 4 | Up | 3.924802e-07 | 0.002378921 |
| high-density lipoprotein particle%GOcc%GO:0034364 | 15 | Down | 1.226594e-06 | 0.005947756 |
| FORMATION OF TUBULIN FOLDING INTERMEDIATES BY CCT TRIC%... | 16 | Up | 2.816908e-06 | 0.010768109 |
| very-low-density lipoprotein particle%GOcc%GO:0034361 | 13 | Down | 3.553099e-06 | 0.010768109 |
| triglyceride-rich plasma lipoprotein particle%GOcc%GO:0034385 | 13 | Down | 3.553099e-06 | 0.010768109 |
| succinyltransferase activity%GOfm%GO:0016748 | 3 | Up | 7.134972e-06 | 0.019220823 |
| MITOCHONDRIAL TRANSLATION%REACTOME DATABASE ID RELEA... | 86 | Up | 1.018332e-05 | 0.021271797 |

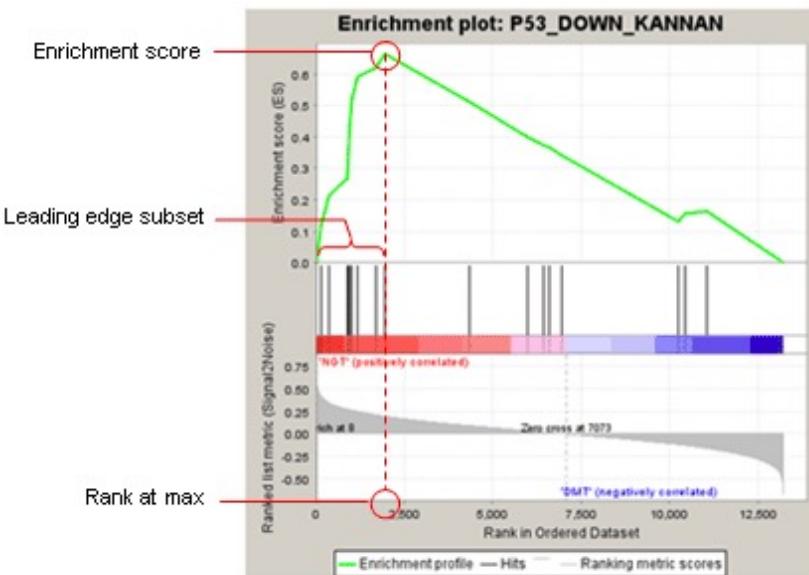
- Important to read the enriched terms and consider the meaning, 1000's of members to fewer than 10
- How specific is the term?
- Is “response to hormone” meaningful?
- “regulation of amine transport” binding?

Set enrichment graph

Barcode plot in R



Enrichment plot from GSEA



Summary

- Different gene expression techniques geared toward each situation/question/hypothesis
- Assay of the RNA in cells/tissues
- Avoid or track and correct for batches
- Adjust for possible confounding variables
- Adjust the p-value
- Genome wide techniques are for assessing sets of genes



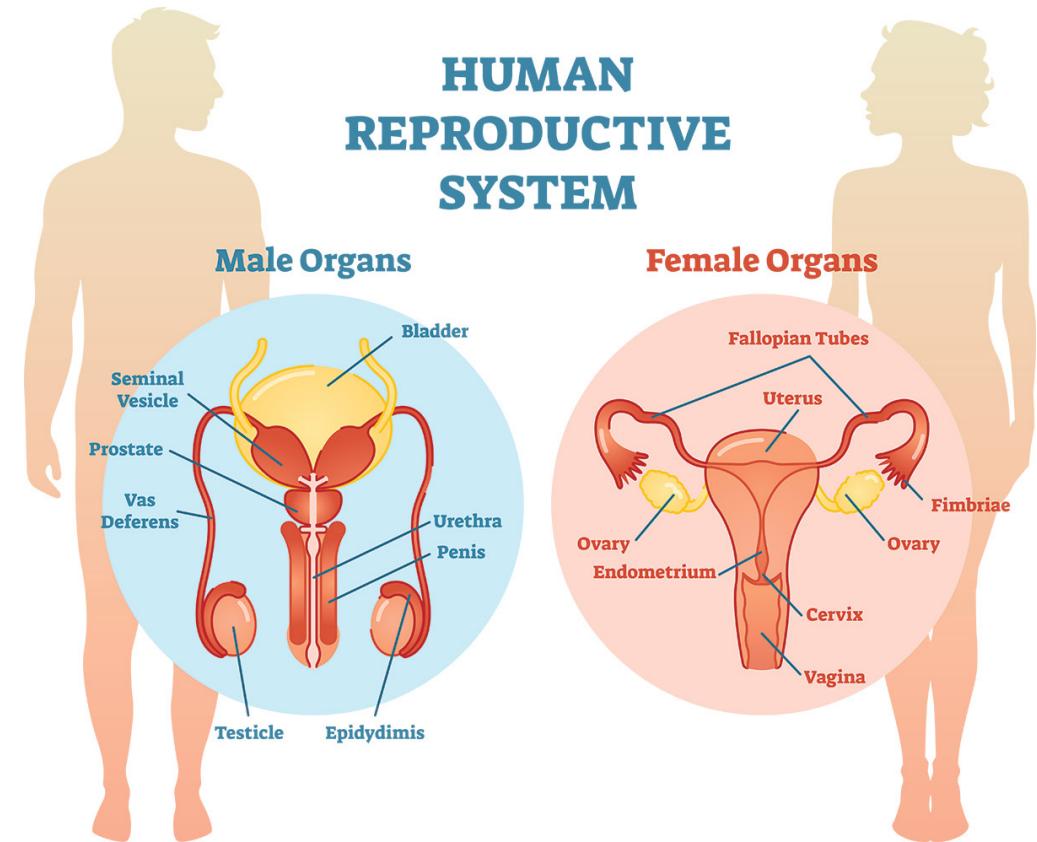
Pregnancy development and adaptations

PSL374

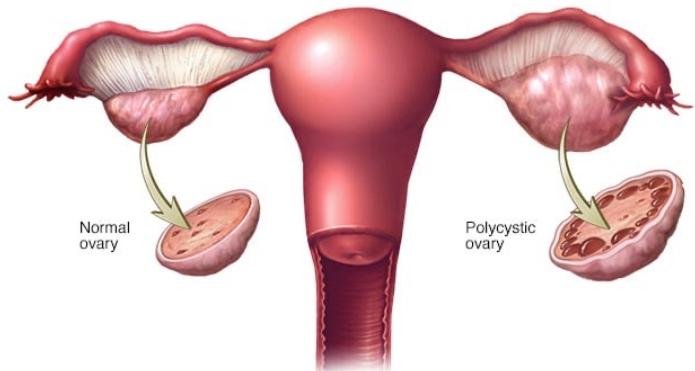
Brian Cox, PhD

More than making babies

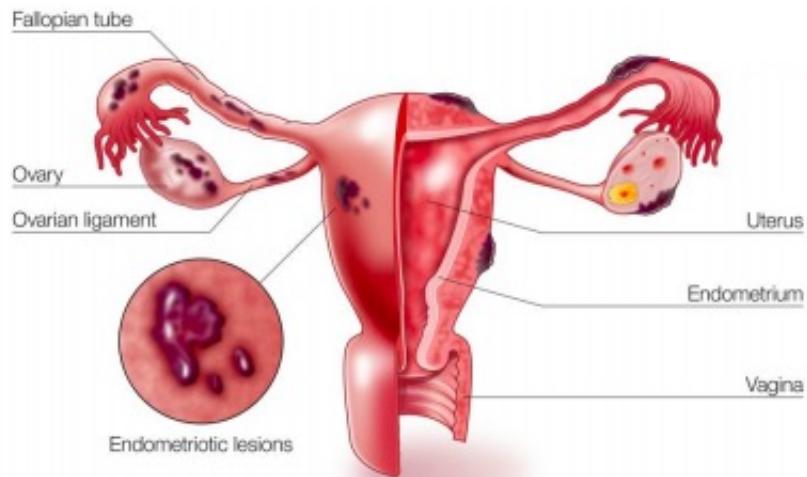
- Connects
 - Development
 - Metabolism
 - Cardiovascular
 - Skeletal, muscular
- In females
 - Regulation of being pregnant
 - Placenta



Female reproductive syndromes

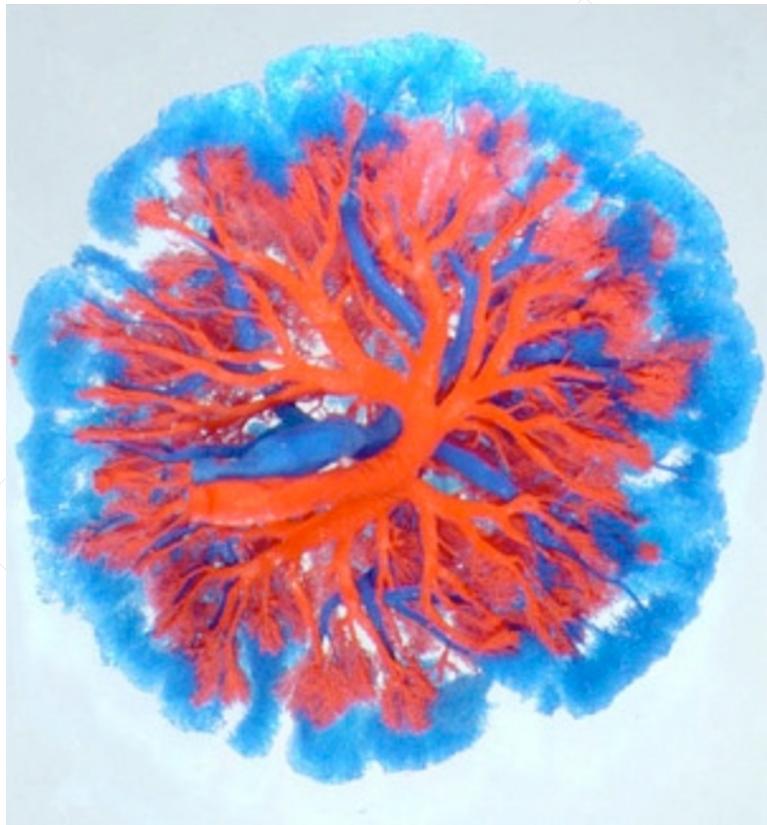


© MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. ALL RIGHTS RESERVED.



- Polycystic ovarian syndrome
- 8% of reproductive age females
- High risk for cardiovascular disease (CVD), diabetes, and metabolic syndrome
- Infertility
- Endometriosis
- 10% of reproductive age females
- Symptomatic,
 - Pelvic pain, and intense dysmenorrhea—painful menstruation
- Asymptomatic,
 - Damage and dysfunction of surrounding organs

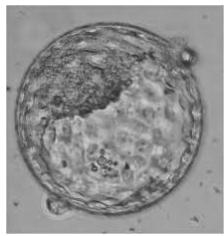
Pregnancy



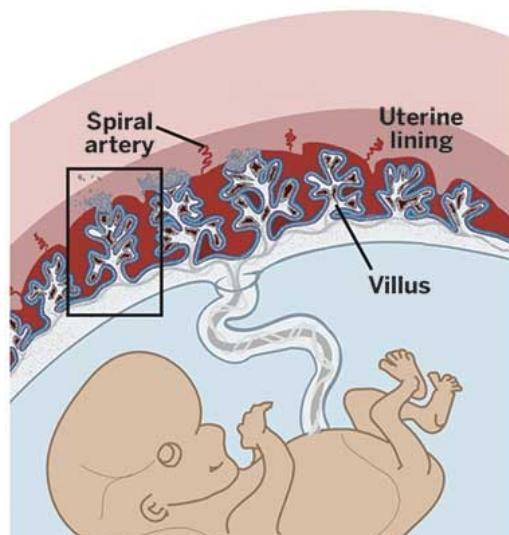
- Immune response
- Angiogenesis
- Endocrine function
- Transfer, nutrients/wastes
- *Must concurrently function and develop*

Organ development

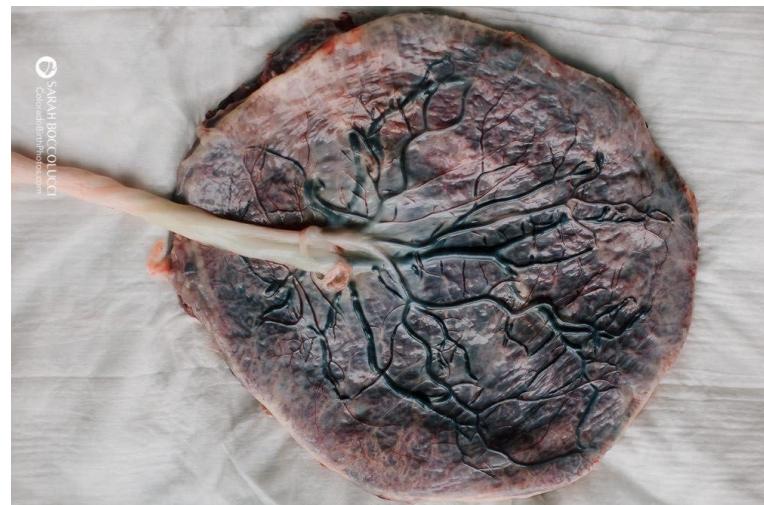
- Trophoblast forms the placenta
- The first lineage to develop
- Forms from an interaction of fetal and maternal tissue



7 days

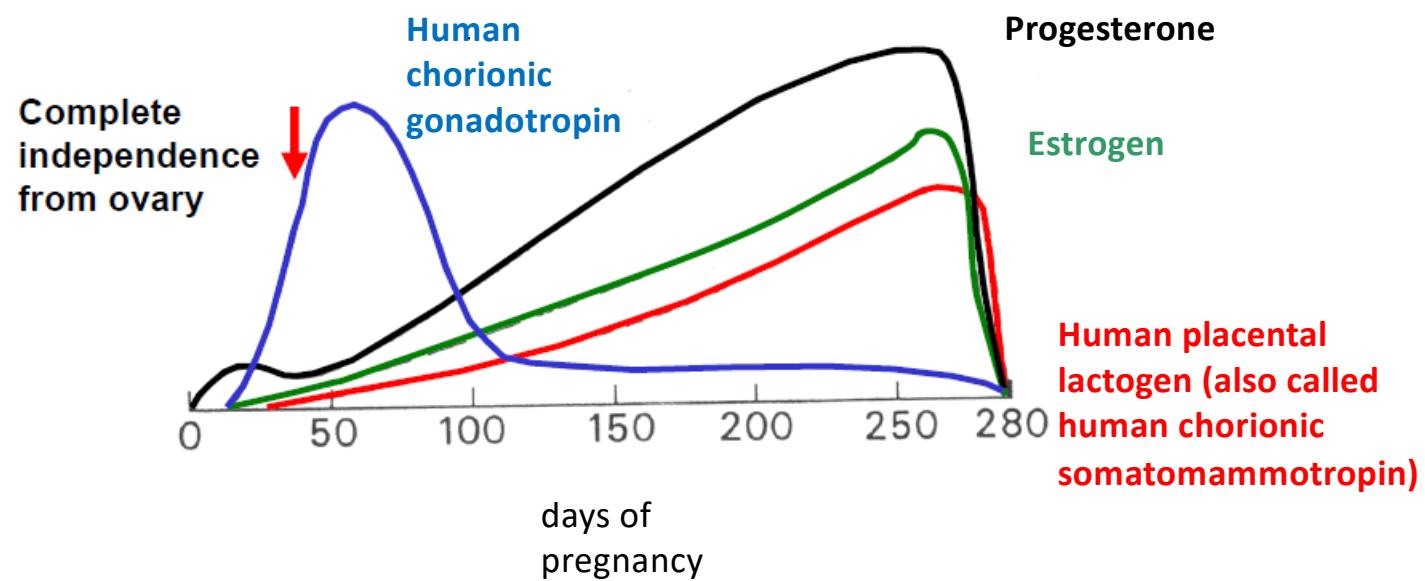


12 weeks



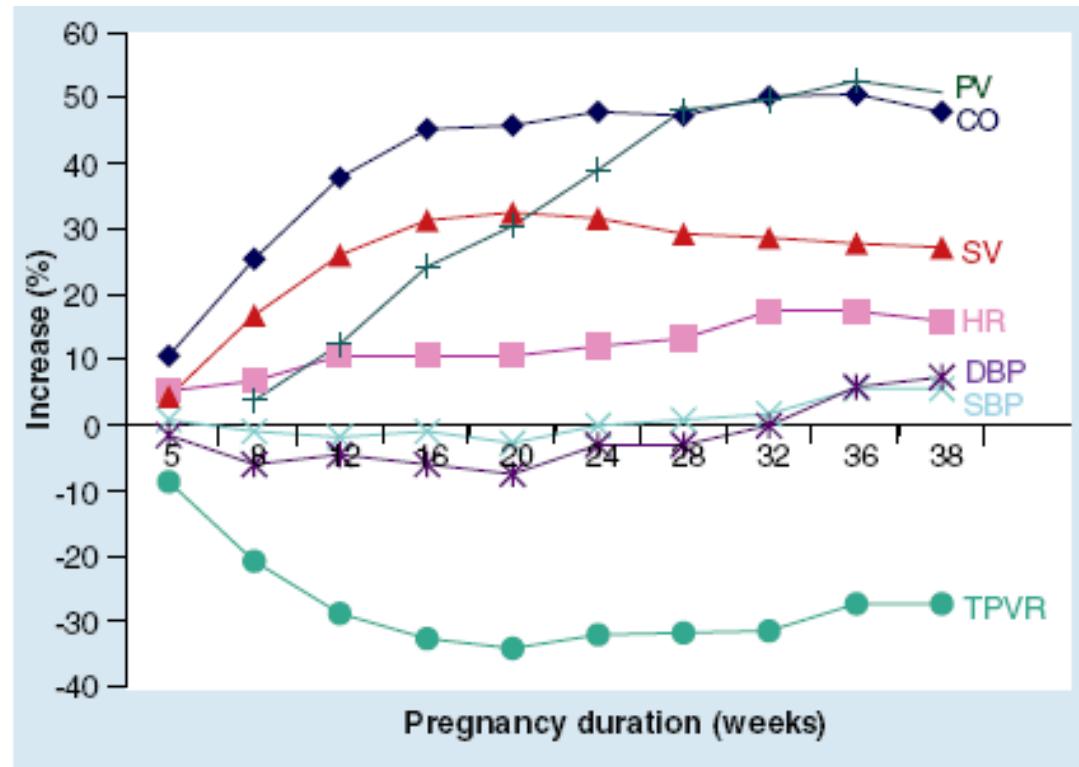
Term (40 weeks)

Endocrine organ



The placenta regulates cardiovascular function

- Increased demand of placenta
- Large plasma volume
- Increased heart rate
- Decreased blood pressure
- Reduce peripheral vascular resistance
- Increased capillarization
- Increased lung alveoli

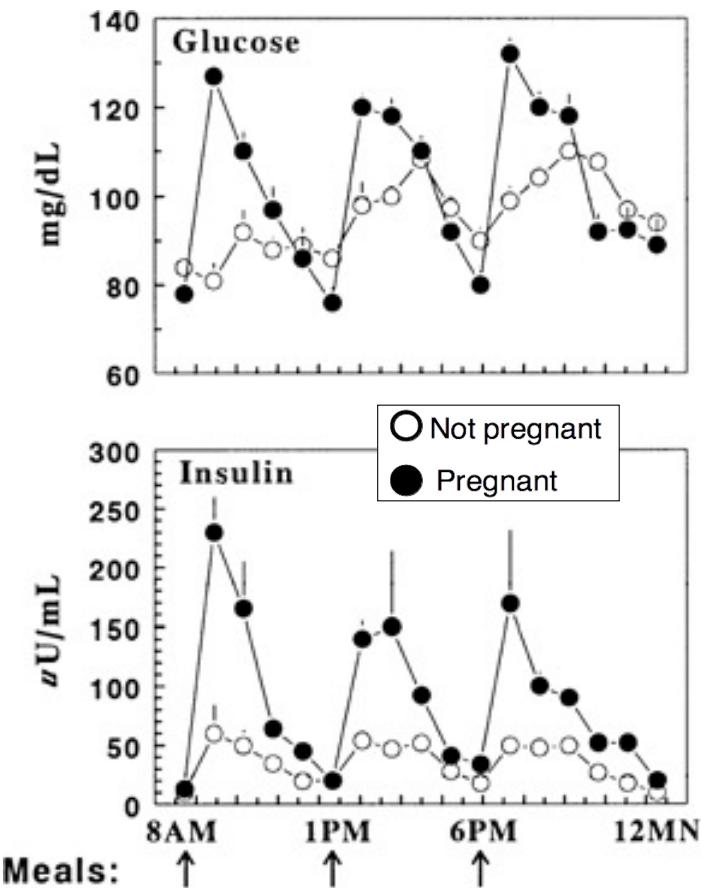


CO: Cardiac output; DBP: Diastolic blood pressure; HR: Heart rate; PV: Plasma volume; SBP: Systolic blood pressure; SV: Stroke volume; TPVR: Total peripheral vascular resistance.

Taken from [3,6].

The placenta regulates metabolism

- Increased placental/fetal demand
- Maintain higher energy load in the blood
- Insulin resistance



Am J Obstet Gynecol, 1981; 140:730-736



What can go
wrong?

Preterm birth

- Labor and iatrogenic (caused by the cure)
- Prevalent, 8% in Canada over half are unknown cause (idiopathic)
- No real treatments, anecdotal
- Premature births cost about \$25 billion in the USA

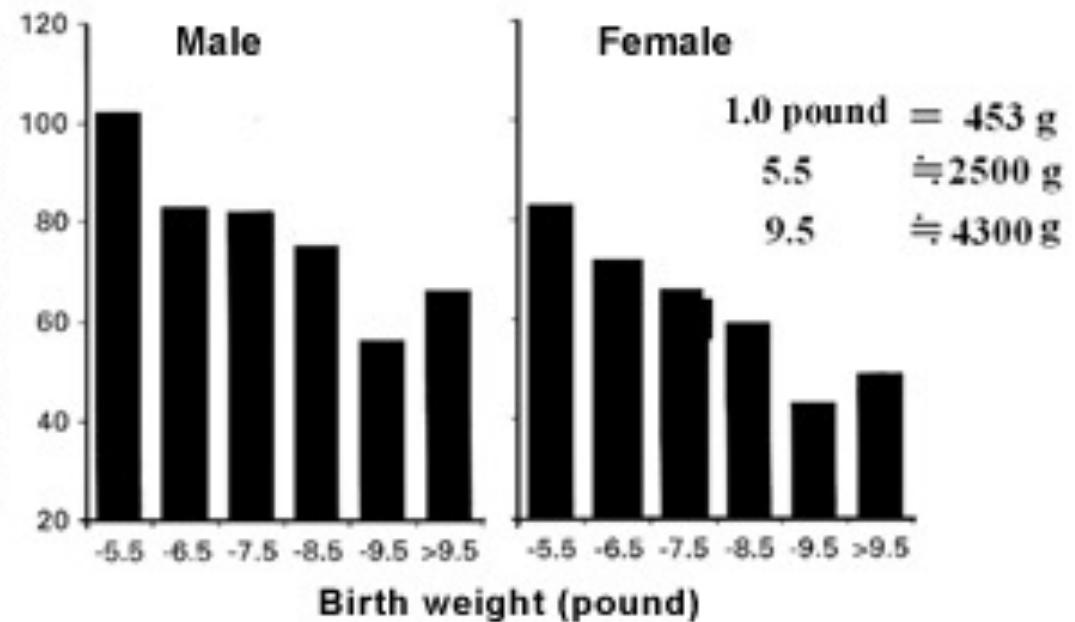
Preeclampsia

- Systemic disease, hypertension, lethal
- Begins around the 25th week of pregnancy
- Affects 5-8% of all pregnancies globally
- Leading cause of fetal and maternal death in the developing world, >70,000 maternal and 500,000 neonatal deaths per year
- One third of all hospital admissions during pregnancy

Long term effects on the child

Low birth weight correlates to multiple chronic health problems

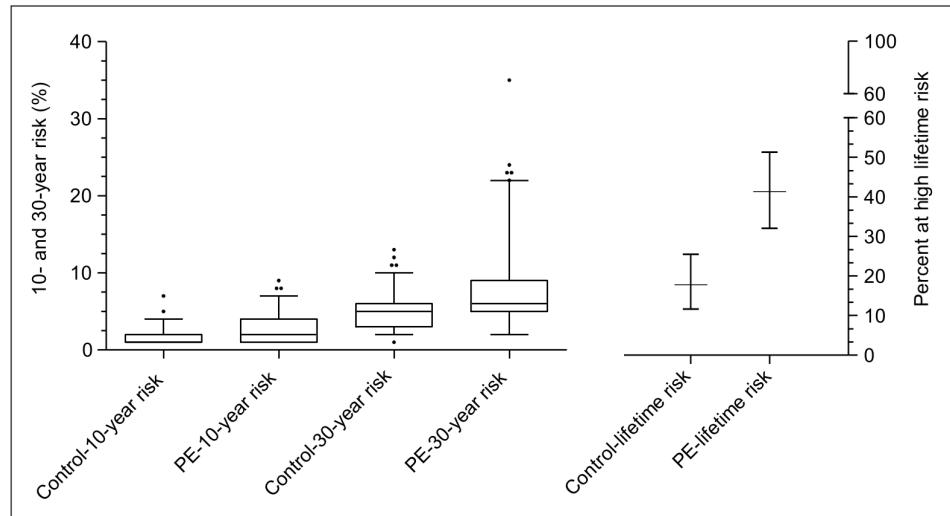
Correlation between birth weight and mortality from ischemic heart disease



Modified from Osmond C. D. Barker, *BMJ* 307: 1519, 1993

Long term maternal health suffers

Different cardiovascular risk scores for control and preeclamptic women at one year postpartum. Ten- and 30-year risks of developing CVD (left-side y axis) are presented as box and whisker plots ($\pm 95\% \text{ CI}$) and proportion at high lifetime risk of CVD (right-side axis) is presented as a line with whiskers ($\pm 95\% \text{ CI}$) for control subjects ($n = 118$) and women with PE ($n = 99$) one year postpartum.



- 1. Smith GN, Jessica P, Walker M, Wen S, Emerging PN. Cardiovascular Disease Risk Estimates Preeclampsia. 2012

Reproductive organs are generally poorly researched

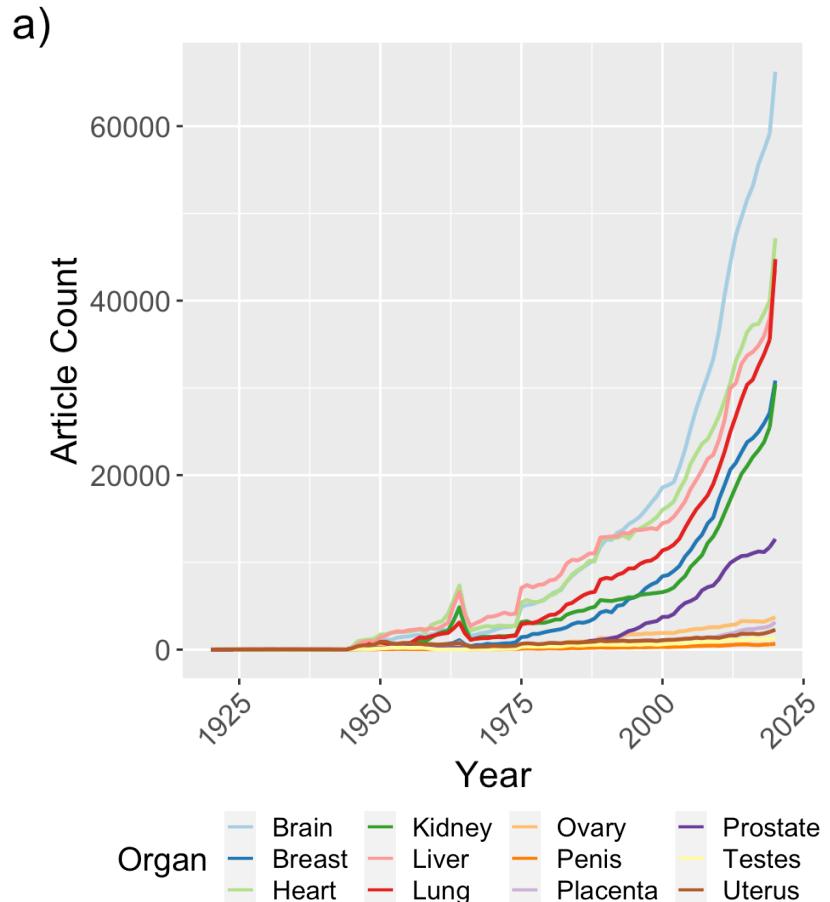
- Many reproductive organs are poorly researched
- Some are exceptions, Breast and Prostate
- Why?

| Keyword | Total Matching Articles |
|----------|-------------------------|
| Brain | 1 024 560 |
| Heart | 857 851 |
| Liver | 837 791 |
| Lung | 628 410 |
| Kidney | 444 876 |
| Breast | 443 590 |
| Prostate | 190 916 |
| Ovary | 86 000 |
| Uterus | 68 024 |
| Placenta | 58 776 |
| Testes | 33 069 |
| Penis | 15 738 |

Mercuri and Cox, in Revision

Worse over time

- Reproductive organs other than breast and prostate become flat
- Up to 80% of breast and prostate research involves cancer
- Lack of basic biology and reproductive disease focused research



Summary

- Pregnancy invokes many adaptive changes
 - Metabolic
 - Cardiovascular
 - Immunological
- Largely driven in part of hormonal cues from the placenta, ovary
- Unfortunately, least understood area of physiology

Our experiment

- Mice aged 8-12 weeks
- Two groups
 - Not pregnant
 - Pregnant E14.5 (14 day post fertilization)
- Collected replicates of
 - Heart left ventricle
 - Liver
 - Lung
- What are the adaptive changes in these organs?
 - Differential expression
 - Ontology enrichments