Week 3 notes:

- handing in exercises
- Journal club
- limma paper
- .. brainstorm

Handing in exercises

- Will always be done via GitHub, but we will use (mostly) private repositories
- Link for Exercise 2: https://classroom.github.com/a/qucI12PB
- Further links will be given in class notes / Slack
- (may say something like pairing to matrikulation numbers required, but I think you can skip that)
- In class today, we will:
  1. do a group assignment via GitHub
  2. practice a pull request (https://help.github.com/articles/creating-a-pull-request/)

Journal club

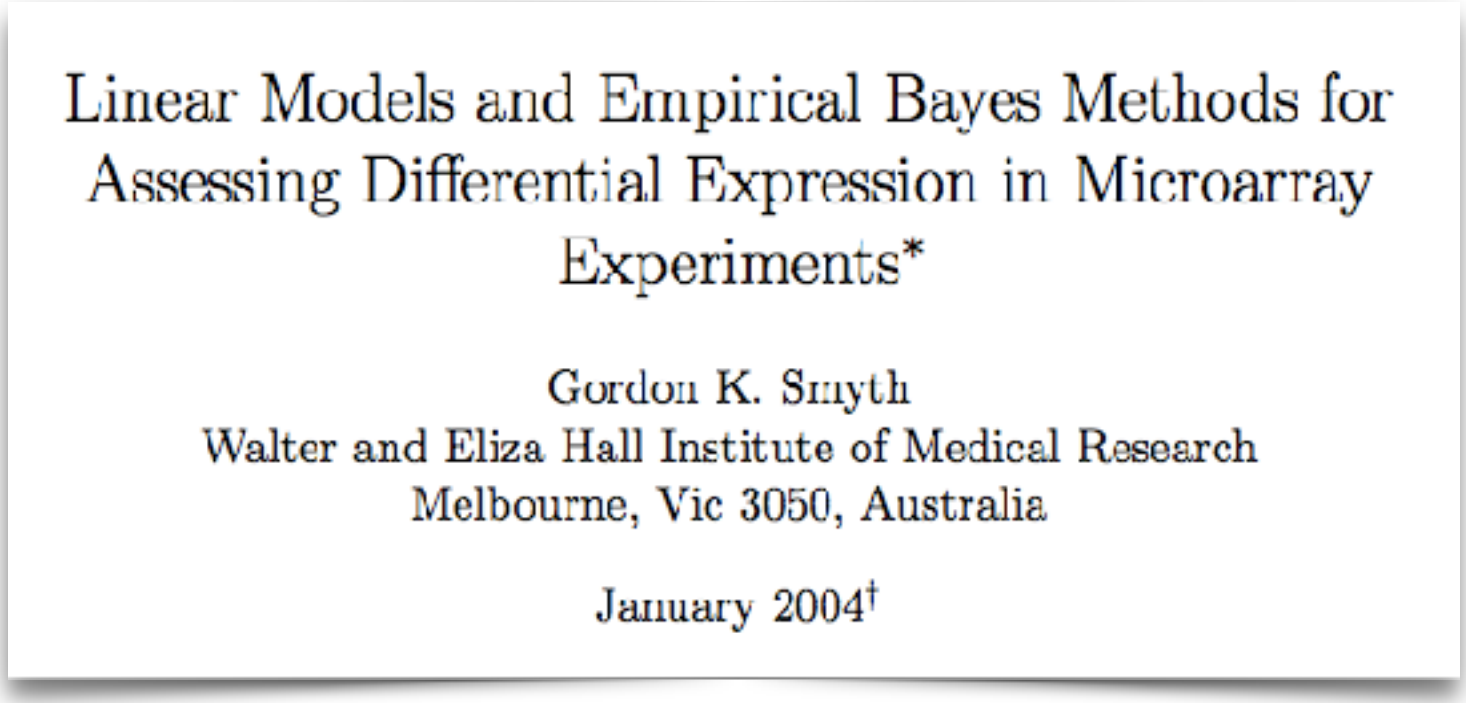
Papers to be picked by 17.00 on 16th October; suggest a date (not 13 Nov or 11 Dec)

Journal Club schedule to be finalised by 23rd October

Given the number of students, it is allowed to do Journal Club either alone or in groups of 2.

Date	Lecturer	Topic	JC
18.09.2017	Mark	admin, mol. biology basics, R markdown	
25.09.2017	Hubert	NGS intro; exploratory data analysis	
02.10.2017	Mark + Hubert	interactive technology session	
09.10.2017	Hubert	mapping	
16.10.2017	Mark	limma 1	
23.10.2017	Mark	limma 2	
30.10.2017	Hubert	RNA-seq quantification	
06.11.2017	Mark	edgeR+friends 1	
13.11.2017	Charlotte	hands-on session #1: RNA-seq	X
20.11.2017	Mark	edgeR+friends 2	
27.11.2017	Hubert	classification	
04.12.2017	Mark	single-cell	
11.12.2017	Gosia	hands-on session #2: mass cytometry	X
18.12.2017	Mark	epigenomics, DNA methylation, ChIP data, gene set analysis	

limma (= linear models for microarray data) paper



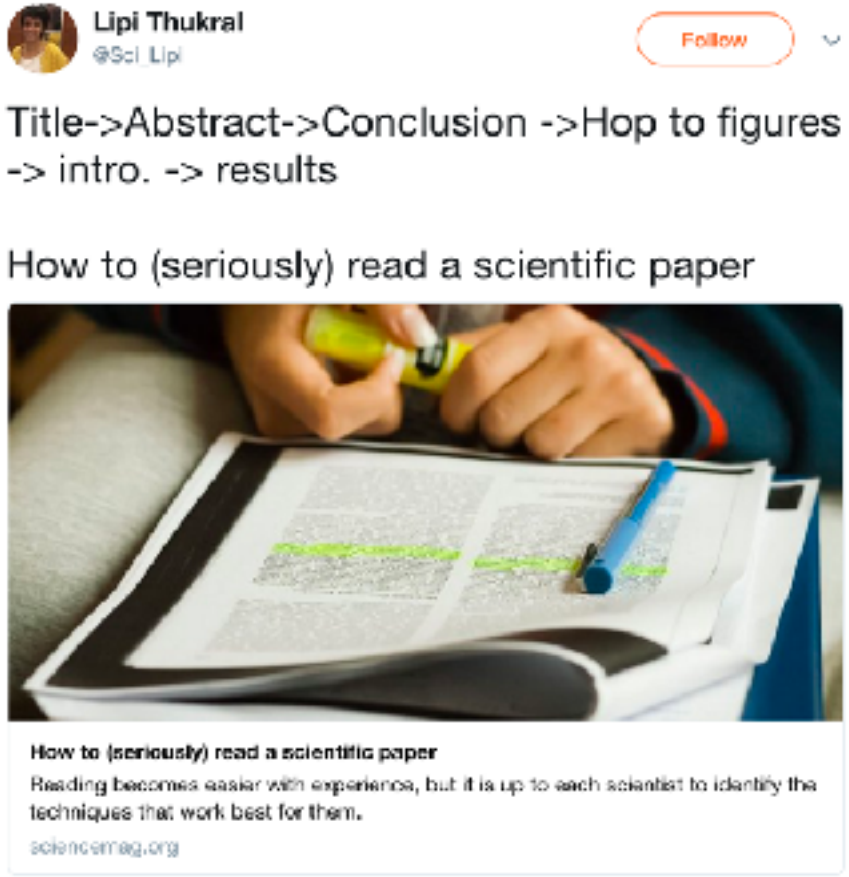
Linear Models and Empirical Bayes Methods for Assessing Differential Expression in Microarray Experiments*

Gordon K. Smyth
Walter and Eliza Hall Institute of Medical Research
Melbourne, Vic 3050, Australia

January 2004†

- seminal paper (cited ~10,000 times)
- provides the foundation for a lot of (statistical) research in genomics
- we will take two lectures to dissect the details

How to read a scientific paper



https://github.com/jtleek/readingpapers

How much should you read?

Academic papers come out all the time. Thousands are published every year, including hundreds in any given specific area. Unless you devote yourself full time to reading academic papers you won't be able to keep up with them all. I believe in the idea that you should read papers that you find interesting. Science is awesome and you shouldn't waste your time on the boring parts if you can avoid it.

In general there are two main ways to find papers that I like. The way I used to do it was set up an aggregator with the RSS feeds from journals that I like, then I use the following (approximate) rates of reading parts of papers.

- 100% - read the title
- 20-50% - read the abstract
- 5-10% - look at the figures/captions
- 1-3% - read the whole paper

The new way that I do it is follow bioRxiv and a bunch of other people who have similar interests on Twitter. I use the above percentages for papers tweeted from aggregators and if I see a paper tweeted by 2-3 people I trust I usually end up reading that paper.