



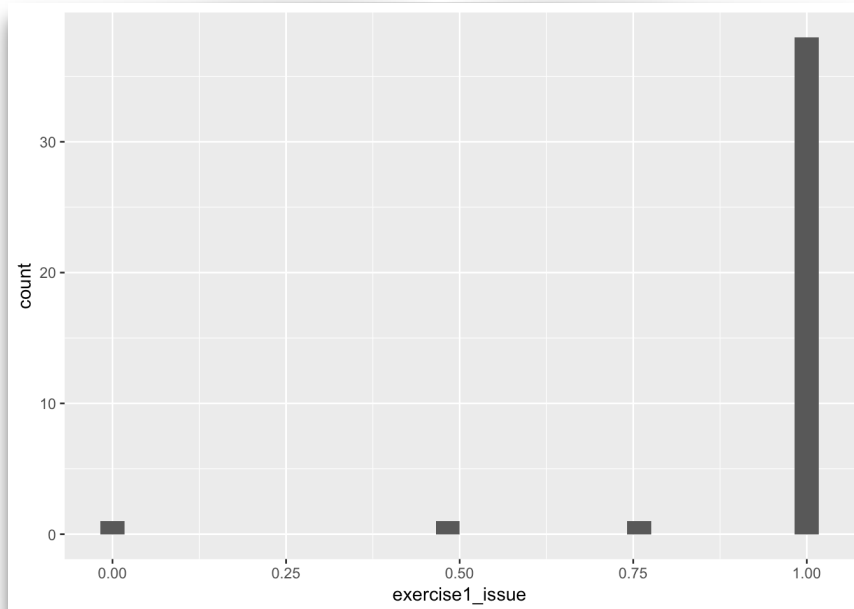
Week 3 notes:

- Exercises
- limma paper
- Journal Club
- .. brainstorm

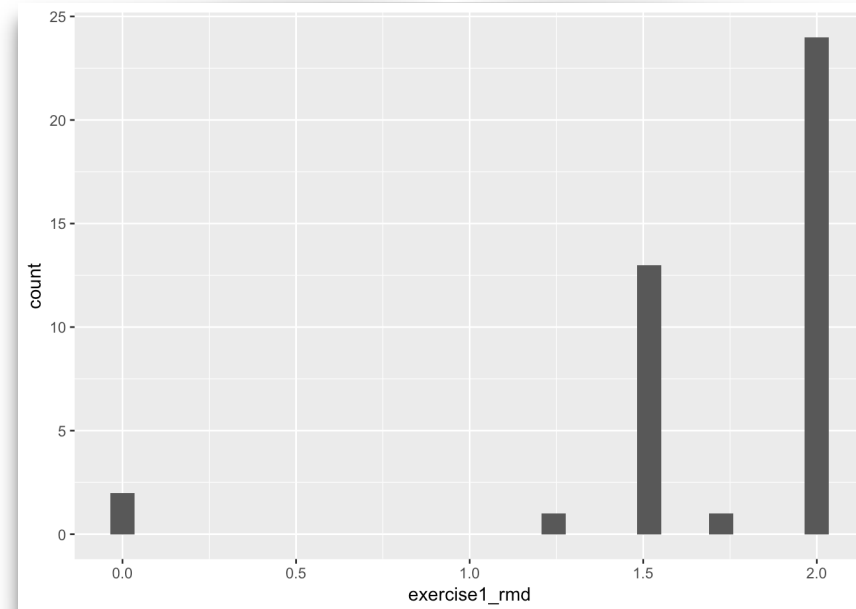
Exercises

- The GitHub system is working: 41 students registered; 40 repos created
- Emails: received? changes? —> send to me on Slack

Part 1



Part 2





Exercises

- Week 2 exercise due later today
- No R markdown for this week's exercise
- Instead, we will:
 1. do a group assignment via GitHub
 2. practice a pull request (<https://help.github.com/articles/creating-a-pull-request/>)
- Post describing how I manage marking/feedback (e.g., Googlesheets):
<http://rpubs.com/markrobinsonca/423663>

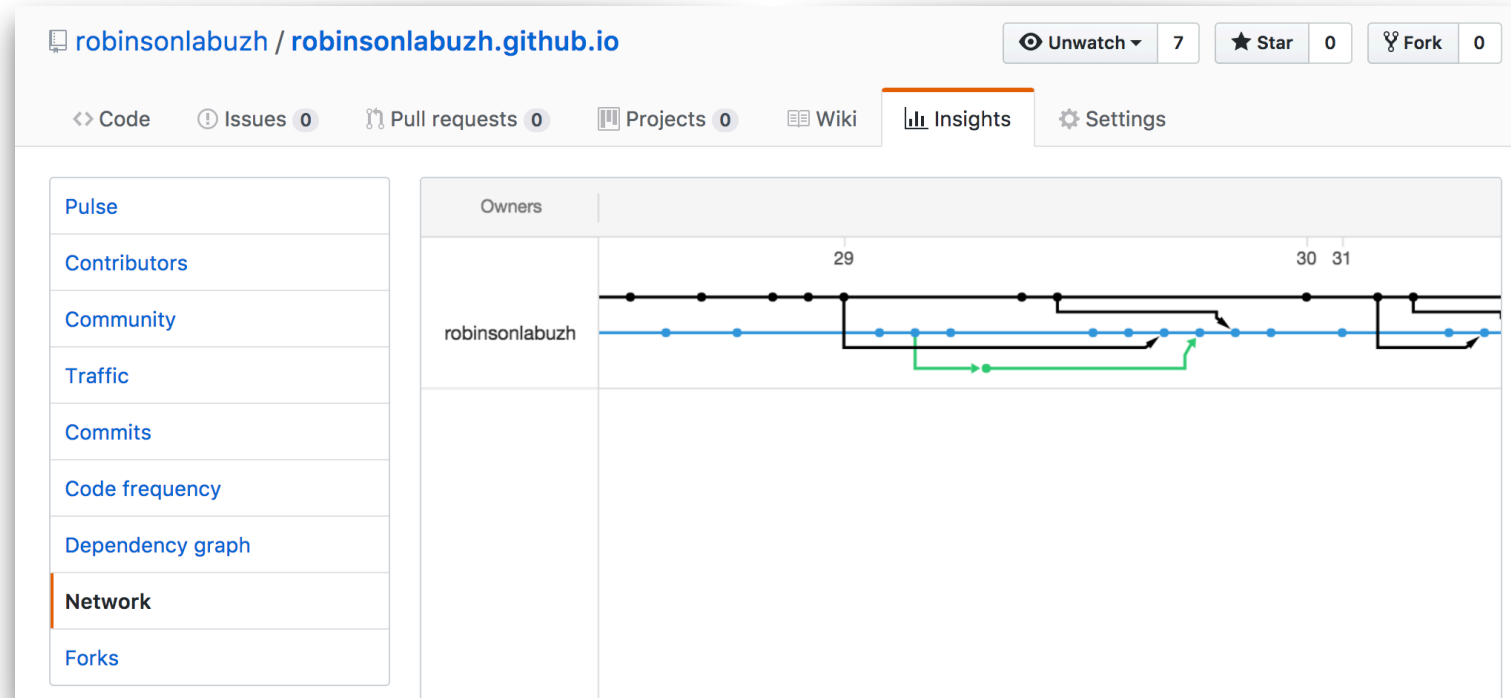


Branches, forking and pull requests

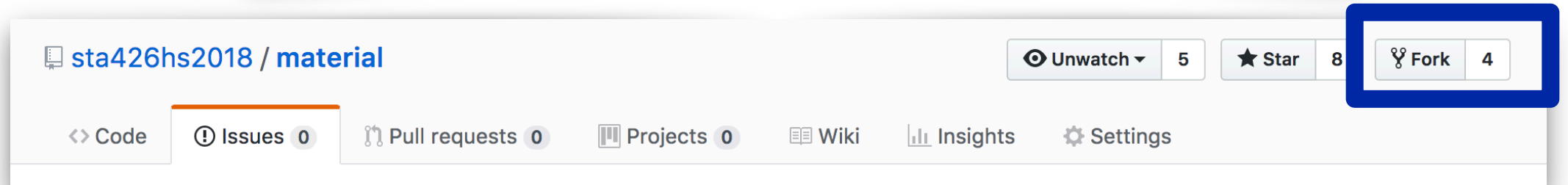
General rule: if you are coding together with others, don't push direct to master branch.

Instead, create a branch, make changes there, submit a pull request .. merge into master after review.

If no ownership to repository, make changes in a fork, then submit PR.



←
black = master





limma (= linear models for microarray data) paper

Linear models and empirical bayes methods for assessing differential expression in microarray experiments

[GK Smyth](#) - Statistical applications in genetics and molecular ..., 2004 - [degruyter.com](#)

... This section describes how **gene-wise linear models** arise from experimental designs and states the distributional assumptions about the ... There may be more or fewer contrasts than coefficients for the **linear model**, although if more than the contrasts will be **linearly** dependent ...

☆ Cited by 10031 Related articles All 39 versions Import into BibTeX

Limma: linear models for microarray data

[GK Smyth](#) - ... and computational biology solutions using R and ..., 2005 - [Springer](#)

... **limma** uses **linear models** to analyze designed microarray experiments (Yang and Speed, 2003; Smyth, 2004) ... reference microarray data, **linear modeling** is much the same as ordinary ANOVA or multiple **regression** except that a **model** is fitted for ... **Linear Models** Case Study 401 ...

☆ Cited by 4799 Related articles All 18 versions Import into BibTeX

limma powers differential expression analyses for RNA-sequencing and microarray studies

[ME Ritchie](#), [B Phipson](#), [D Wu](#), [Y Hu](#)... - Nucleic acids ..., 2015 - [academic.oup.com](#)

... **limma** powers **differential expression** analyses for RNA-sequencing and microarray studies. Matthew E. Ritchie. 1. Molecular Medicine Division, The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, Victoria 3052, Australia ...

☆ Cited by 3437 Related articles All 14 versions Import into BibTeX

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 digest this into a single lecture (15 Oct)

How to read a scientific paper



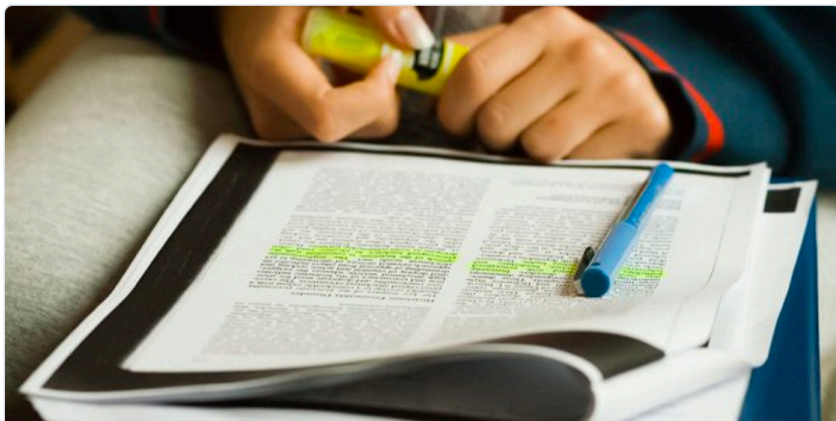
Lipi Thukral

@Sci_Lipi

Follow

Title->Abstract->Conclusion ->Hop to figures
-> intro. -> results

How to (seriously) read a scientific paper



How to (seriously) read a scientific paper

Reading becomes easier with experience, but it is up to each scientist to identify the techniques that work best for them.

sciencemag.org

11:00 PM - 19 Mar 2017

<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.



Journal club

Papers to be picked by 18.00
on 15th October; discuss it
with Hubert and I.

Journal Club schedule to be
finalized by 22nd October

Given the number of students,
groups of 2-3.

Use the #papers channel. I
have put several suggestions
there.

Signup by pull request.

Schedule

Date	Lecturer	Topic	Exercise	JC1	JC2
17.09.2018	Mark + Hubert	admin; mol. bio. basics	R markdown; git(hub)		
24.09.2018	Hubert	NGS intro; exploratory data analysis	EDA in R		
01.10.2018	Mark + Hubert	interactive technology/statistics session	group exercise: technology pull request		
08.10.2018	Hubert	mapping	Rsubread		
15.10.2018	Mark	limma + friends	linear model simulation + design matrices		
22.10.2018	Hubert	RNA-seq quantification	RSEM		
29.10.2018	Charlotte	hands-on session #1: RNA-seq	FASTQC/Salmon/etc.		X
05.11.2018	Mark	edgeR+friends 1	basic edgeR/voom		
12.11.2018	Mark	edgeR+friends 2	GLM/DEXSeq		
19.11.2018	Mark	single-cell dim. reduction + clustering; FDR	conquer		
26.11.2018	Lukas	hands-on session #2: cytometry	cytof null comparison		X
03.12.2018	Hubert	classification	MLInterfaces		
10.12.2018	Mark	loose ends: HMM, EM, robustness	segmentation, peak finding		
17.12.2018	Mark	hands-on session #3: single-cell RNA-seq	full scRNA-seq pipeline		X