

Top Computational biology Resources

Ming Tommy Tang

4/10/2023

12 websites to learn computation and many others!

1. coursera <https://www.coursera.org/> The first website I used. I took a data science Specialization <https://www.coursera.org/specializations/jhu-data-science> and <https://www.coursera.org/learn/bioinformatics>
2. edx <https://www.edx.org/> I took this one first <https://www.edx.org/professional-certificate/harvardx-data-analysis-for-life-sciences> and <https://www.edx.org/course/statistical-learning> and recently some management classes
3. udacity <https://www.udacity.com/> I took R courses, ggplot2, github and intro to ML
4. udemy <https://www.udemy.com/> I took several python courses there.
5. MIT opencourseware <https://ocw.mit.edu/> 1806 linear algebra and many others!
6. harvard CS50 <https://pll.harvard.edu/course/cs50-introduction-computer-science?delta=0>
7. youtube channel 3blue1brown https://www.youtube.com/channel/UCYO_jab_esuFRV4b17AJtAw blow you away with Crystal clear explanations. I watched the linear algebra series
8. Khan academy <https://www.khanacademy.org/>
9. free code camp <https://www.freecodecamp.org/>
10. Josh Starmer's Statquest for machine learning. The best ever <https://www.youtube.com/channel/UCtYLUtTgS3k1Fg4y5tAhLbw>
11. EBML-EBI training <https://www.ebi.ac.uk/training/online/> bioinformatics courses
12. skillup <https://www.simplilearn.com/skillup-free-online-courses> this is new to me as well.

10 courses to get you started with bioinformatics

1. <http://rafalab.dfci.harvard.edu/pages/harvardx.html> by Rafa Irizarry at Dana-Farber.
2. <https://github.com/quinlan-lab/applied-computational-genomics#course-lecture-slides>

by Aaron Quinlan, the creator of bedtools and many other cool tools.

3. <https://www.bioinformaticsalgorithms.org/> You can find the video classes on Coursera
4. <http://www.personal.psu.edu/iua1/courses/2014-BMMB-852.html> by Istvan Albert, the creator of biostars.

5. Introduction to Bioinformatics and Computational Biology <https://liulab-dfci.github.io/bioinfo-combio/> by @XShirleyLiu

glad to contribute a little myself.

6. data carpentry workshops <https://datacarpentry.org/lessons/#genomics-workshop> I am honored to serve as the curriculum committee chair
7. Computational Genomics: Applied Comparative Genomics <https://github.com/schatzlab/appliedgenomics2018>
8. Introduction to Computational Biology <https://biodatascience.github.io/compbio/> by Mike Love
9. MIT Computational Biology: Genomes, Networks, Evolution, Health - Fall 2018 - 6.047/6.878/HST.507 by Manolis Kellis
10. An introduction to Applied Bioinformatics <https://github.com/applied-bioinformatics/iab2> Very nice book with python code.

16 resources for re-analyzing public expression data.

1. <https://rnama.com/docs/search-evaluation> RNA meta Analysis has ~26,700 studies (5,717 RNA-Seq and 20,955 Microarray)
2. refine.bio will have harmonized over 60,000 gene expression experiments
3. BioJupies <https://maayanlab.cloud/biojupies/>
4. [Recount2-FANTOM](https://recount2.fantom.org/) Recounting the FANTOM Cage Associated Transcriptome. Long non-coding RNAs.
5. Recount3 <https://rna.recount.bio/>
6. [dee2](https://dee2.org/) Digital Expression Explorer 2. Digital Expression Explorer 2 (DEE2) is a repository of uniformly processed RNA-seq data mined from public data obtained from NCBI Short Read Archive. By Ziemann Mark et.al! Version 2 of dee.
7. Extracting allelic read counts from 250,000 human sequencing runs in Sequence Read Archive <https://www.biorxiv.org/content/10.1101/386441v1?rss=1>
8. [MetaSRA: normalized sample-specific metadata for the Sequence Read Archive](https://meta.sra.hawaii.edu/)
9. [ARCHS4: Massive Mining of Publicly Available RNA-seq Data from Human and Mouse](https://archs4.org/) ARCHS4 provides access to gene counts from HiSeq 2000, HiSeq 2500 and NextSeq 500 platforms for human and mouse experiments from GEO and SRA.
10. [DEP-reads: Uniformly processed public RNA-Seq data](https://dep.reads.org/) Read counts data for 5,470 human and mouse datasets from ARCHS4 v6 and 12,670 datasets from DEE2 for 9 model organisms by Steven Ge.
11. [SRA-explorer](https://sra-explorer.org/) This tool aims to make datasets within the Sequence Read Archive more accessible.
12. [intropolis](https://intropolis.org/) is a list of exon-exon junctions found across **21,504** human RNA-seq samples on the Sequence Read Archive (SRA) from spliced read alignment to hg19 with Rail-RNA.
13. [batch recompute](https://batch-recompute.org/) ~20,000 RNA-seq samples from target sequencing project such as TCGA, TARGET and GETEX. Used hg38 and gencode v21 as annotation.
14. [A cloud-based workflow to quantify transcript-expression levels in public cancer compendia](https://cloud-bioinformatics.github.io/) used kallisto for TCGA/CCLE datasets and gencode v24 as annotation.

15. [MiPanda](#) is an online resource for the interrogation and visualization of gene expression data from the myriad of publicly available cancer and normal next generation sequencing datasets.
16. [Curation of over 10,000 transcriptomic studies to enable data reuse](#)

5 websites to analyze GEO RNAseq data without a single line of code

1. GEOdiver <https://www.geodiver.co.uk/>
2. shinyGEO <https://gdancik.shinyapps.io/shinyGEO/>
3. scanGEO <http://scangeo.dartmouth.edu/ScanGEO/>
4. GREIN : GEO RNA-seq experiments interactive navigator for re-analyzing GEO RNA-seq data <https://hub.docker.com/r/ucbd2k/grein/>
5. ImaGEO: Integrative Meta-Analysis of GEO Data <https://imageo.genyo.es/>
6. Gemma <https://gemma.msl.ubc.ca/home.html>

12 web tools to explore genomics data

1. cbiportal <https://cbiportal.org> explore genomic datasets at the tips of your fingers
2. xena <https://xena.ucsc.edu>, a UCSC effort. Everyone needs to learn how to use UCSC genome browser <https://genome.ucsc.edu>
3. depmap portal <https://depmap.org/portal/interactive/> Cancer Cell Line Encyclopedia
4. TCGA RNA fusion portal <https://tumorfusions.org>
5. <http://gepia2.cancer-pku.cn/#index> Gene expression profile interactive analysis
6. Tumor Immune Syngeneic MOuse (TISMO) database <http://tismo.cistrome.org>
7. PDX models <https://lnkd.in/ezy9kns>
8. <http://tide.dfci.harvard.edu/login/> Tumor Immune Dysfunction and Exclusion
9. <http://timer.cistrome.org> TIMER is a comprehensive resource for systematical analysis of immune infiltrates across diverse cancer types
10. genePattern <https://genepattern.org>
11. <https://epiviz.github.io/index.html> Epiviz is an interactive visualization tool for functional genomics data.
12. draw mutation for a protein <https://pecan.stjude.cloud/proteinpaint/TP53/>

6 links on workflow to make your life easier

Bioinformatics analysis involves a lot of steps, 6 links on workflow to make your life easier:

1. over hundreds of workflow tools and engines <https://github.com/pditommaso/awesome-pipeline>
2. see also from the CWL wiki <https://github.com/common-workflow-language/common-workflow-language/wiki/Existing-Workflow-systems>

3. A review of bioinformatic pipeline frameworks <https://academic.oup.com/bib/article/18/3/530/2562749>
4. discussion on biostars <https://www.biostars.org/p/115745/>
5. two papers by Titus Brown [Ten simple rules and a template for creating workflows-as-applications](#)
6. Streamlining Data-Intensive Biology With Workflow Systems <https://dib-lab.github.io/2020-workflows-paper/>

8 links to bookmark for better data visualization

Data visualization is a critical step in data analysis, 8 links to bookmark for better data visualization :

1. Nature Methods point of view data visualization <http://blogs.nature.com/methagora/2013/07/data-visualization-points-of-view.html> the columns on color mapping and heatmap are very nice.
2. Ten simple rules to colorize biological data visualization <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008259>
3. data visualization resources <https://sabahzero.github.io/dataviz/resources>
4. Fundamentals of Data Visualization <https://clauswilke.com/dataviz/>
5. Data Visualization <https://socviz.co/> by Kieran Healy. I've read book and 4 and 5.
6. [R Graphics Cookbook](#) by Winston Chang.
7. [ggplot2: Elegant Graphics for Data Analysis](#) by Hadely Wickham.

5 tools to visualize genomic datasets

1. Karyoploter https://bernatgel.github.io/karyoploter_tutorial/Tutorial/PlotCoverage/PlotCoverage.html I used that to plot single-cell ATACseq tracks <https://github.com/crazyhottommy/scATACutils/#plot-atacseq-tracks-for-each-cluster-of-cells>, more examples https://rpubs.com/crazyhottommy/scATAC_tracks
2. plotgardener is a genomic data visualization package for R. Using `grid` graphics, `plotgardener` empowers users to programmatically and flexibly generate multi-panel figures

<https://github.com/PhanstielLab/plotgardener>

3. The goal of `g(r)osling` <https://github.com/gosling-lang/gosling> is to help you build interactive genomics visualizations with `Gosling`. This package uses `reticulate` to provide an interface to the `Gos` Python package. <https://github.com/gosling-lang/gosling>

4. Intervene: a tool for intersection and visualization of multiple gene or genomic region sets

<https://bitbucket.org/CBGR/intervene/src/master/>

5. <https://42basepairs.com/> saw it yesterday by @RobAboukhalil

8 links to BETTER understand principal component analysis (PCA)

1. <https://divingintogeneticsandgenomics.rbind.io/post/pca-in-action/> PCA in action, my blog post to calculate SVD and PCA with #rstats
2. <https://www.youtube.com/watch?v=rYz83XPxiZo> MIT 1806 linear algebra on SVD
3. <https://peterbloem.nl/blog/pca-4> THE SINGULAR VALUE DECOMPOSITION (SVD)
4. <http://rafalab.github.io/pages/harvardx.html> High Dimension data analysis, week 2.
5. <https://towardsdatascience.com/why-pca-looks-triangular-a642daac721a> why PCA looks triangular.
6. <https://www.nxn.se/valent/2017/6/12/how-to-read-pca-plots> How to read PCA plots for single-cell data.
7. <https://twitter.com/AedinCulhane/status/1007110262187544577> PCA horseshoe artifact
8. https://www.youtube.com/watch?v=_UVHneBUBW0 by Josh Starmer

9 tools for interactive exploring single-cell RNAseq data

1. cellxgene <https://github.com/chanzuckerberg/cellxgene>
2. cellar <https://github.com/euxhenh/cellar>
3. scSVA: an interactive tool for big data visualization and exploration in single-cell omics <https://www.biorxiv.org/content/10.1101/512582v1>
4. ASAP: a web-based platform for the analysis and interactive visualization of single-cell RNA-seq data <https://academic.oup.com/bioinformatics/article/33/19/3123/3852081?login=false>
5. iSEE Provides functions for creating an interactive Shiny-based graphical user interface for exploring data stored in SummarizedExperiment objects, including row- and column-level metadata
6. VISION A high-throughput and unbiased module for interpreting scRNA-seq data.
7. DISCO: Deep Integration of Single-Cell Omics. Want to visual millions of cell online and annotate cell type automatically? Try it!!! Make single cell easier and make life easier!
8. TISCH Tumor Immune Single-cell Hub (TISCH) is a scRNA-seq database focusing on tumor microenvironment (TME).
9. CancerSCEM To date, CancerSCE version 1.0 consists of 208 cancer samples across 28 studies and 20 human cancer types

12 resources to bookmark for reproducible computational research

1. a reproducible workflow. <https://www.youtube.com/watch?v=s3JldKoA0zw> This two minute video will change your mind on reproducible research
2. Parallel sequencing lives, or what makes large sequencing projects successful <https://academic.oup.com/gigascience/article/6/11/gix100/4557140?login=false>
3. Common-sense approaches to sharing tabular data alongside publication <https://www.sciencedirect.com/science/article/pii/S2666389921002300>
4. A Reproducible Data Analysis Workflow with R Markdown, Git, Make, and Docker <https://psyarxiv.com/8xzqy/>

5. Practical Computational Reproducibility in the Life Sciences [https://www.cell.com/cell-systems/fulltext/S2405-4712\(18\)30140-6](https://www.cell.com/cell-systems/fulltext/S2405-4712(18)30140-6)
6. A video by Dr.Keith A. Baggerly from MD Anderson [The Importance of Reproducible Research in High-Throughput Biology](#) highly recommended.
7. Ten Simple Rules for Reproducible Computational Research <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003285>)
8. Good Enough Practices in Scientific Computing <http://arxiv.org/abs/1609.00037>
9. Best Practices for Scientific Computing <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1001745>
10. A Quick Guide to Organizing Computational Biology Projects <http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.100042> A must read for computational biologists!
11. Reproducibility of computational workflows is automated using continuous analysis <https://www.nature.com/articles/nbt.3780>
12. Five selfish reasons to work reproducibly <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-015-0850-7>

7 FREE Books to learn data science

1. Data science: A first introduction <https://datasciencebook.ca/>
2. Introduction to Data Science <http://rafalab.dfci.harvard.edu/dsbook/>
3. Agile Data Science with R <https://edwinth.github.io/ADSwR/index.html>
4. Tidy Modeling with R <https://www.tmrw.org/>
5. Feature Engineering and Selection: A Practical Approach for Predictive Models <https://bookdown.org/max/FES/>
6. Another Book on Data Science <https://www.anotherbookondatascience.com/> compare R and python side by side
7. Research Software Engineering with Python <https://merely-useful.tech/py-rse/>

7 Books for you to learn bioinformatics

1. Data Analysis for the Life Sciences <https://leanpub.com/dataanalysisforthelifesciences> You can get it for free!
2. practical computing for biologist <https://practicalcomputing.org/> My first ever book to start learning computational biology.
3. A Primer for Computational Biology <https://open.oregonstate.education/computationalbiology/>
4. Computational Genomics with R <http://compgenomr.github.io/book/>
5. The Biologist's Guide to Computing <https://book.biologistsguide2computing.com/en/stable/>
6. Bioinformatics Data Skills <https://www.oreilly.com/library/view/bioinformatics-data-skills/9781449367480/> A must read to upgrade your bioinformatics skills once you know the basics.
7. Bioinformatics Workbook: A tutorial to help scientists design their projects and analyze their data. <https://bioinformaticsworkbook.org/#gsc.tab=0>

8 Resources to study Transcription factor binding, enhancers and histone modification distribution

1. ENCODE <https://www.encodeproject.org/>
2. The International Human Epigenome Consortium (IHEC) epigenome data portal <http://epigenomesportal.ca/ihec/index.html?as=1>
3. Blueprint epigenome <http://dcc.blueprint-epigenome.eu/#/home>
4. EpiFactors <http://epifactors.autosome.ru/> is a database for epigenetic factors, corresponding genes and products.
5. CistromeDB <http://cistrome.org/db/#/> by Shirley Liu group
6. Remap <https://remap2022.univ-amu.fr/> is a large scale integrative analysis of DNA-binding experiments for Homo sapiens, Mus musculus, Drosophila melanogaster and Arabidopsis thaliana transcriptional regulators.
7. ChIP-Atlas <http://chip-atlas.org/> An integrative, comprehensive database to explore public Epigenetic dataset, including ChIP-Seq, DNase-Seq, ATAC-Seq, and Bisulfite-Seq data: ChIP-Atlas covers almost all public data archived in Sequence Read Archive of NCBI, EBI, and DDBJ with over 224,000 experiments.
8. Fantom5 <https://fantom.gsc.riken.jp/5/>

8 R/command line tools to deal with excel, tsv and csv files

R packages:

- [readxl](#)
- [tidyxl](#)
- [janitor](#)

command line tools:

- [VisiData](#) is an interactive multitool for tabular data. It combines the clarity of a spreadsheet, the efficiency of the terminal, and the power of Python, into a lightweight utility which can handle millions of rows with ease.
- [csvkit](#)
- [csvtk](#) a cross-platform, efficient and practical CSV/TSV toolkit.
- [Miller](#) is a command-line tool for querying, shaping, and reformatting data files in various formats including CSV, TSV, JSON, and JSON Lines.
- [eBay's TSV Utilities](#)

6 training resources for data management

- Best Practices for Biomedical Research Data Management <https://learn.canvas.net/courses/1854>
- Research Data Management Librarian Academy (<https://rdmla.github.io/>)
- DataONE Data Management Skillbuilding Hub (<https://dataoneorg.github.io/Education>)
- Data Management Training Clearinghouse (<https://dmtclearinghouse.esipfed.org/>)
- Research data management open training materials Zenodo Community (<https://zenodo.org/communities/dcc-rdm-training-materials>)
- Consortium of European Social Science Data Archives (CESSDA) Training Resources (<https://www.cessda.eu/Training-Resources>)

Bonus:

Learn from TCGA # Collaborative Genomics Projects: A Comprehensive Guide <https://www.sciencedirect.com/book/9780128021439/collaborative-genomics-projects-a-comprehensive-guide>

7 links to deeply understand heatmap

Making a heatmap is an essential skill for a bioinformatician. Just check how many figures are heatmap or heatmap variants in the genomics or single cell paper.

But you probably do not understand heatmap. 7 reading resources to understand heatmap!

1. Mapping quantitative data to color <https://www.nature.com/articles/nmeth.2134>
2. Heat map from Nature Method column <https://www.nature.com/articles/nmeth.1902>
3. A tale of two heatmap functions <https://rpubs.com/crazyhottommy/a-tale-of-two-heatmap-functions>
An old post by me.
4. Heatmap demystified https://rpubs.com/crazyhottommy/heatmap_demystified yet another post by me
5. understand color mapping is key <https://jokergoo.github.io/ComplexHeatmap-reference/book/a-single-heatmap.html#colors>
6. understand rastering <https://jokergoo.github.io/2020/06/30/rasterization-in-complexheatmap/>
7. what happens when you have a huge matrix 20,000 rows/genes x 50 columns to plot? <https://gdevailly.netlify.app/post/plotting-big-matrices-in-r/>

I learned so much from Zuguang Gu, thanks for his awesome Complexheatmap package <https://jokergoo.github.io/ComplexHeatmap-reference/book/index.html> . it is my go-to tool for making heatmaps.

23 tools to work with (single-cell) TCR/BCR-seq immune repertoire data

1. [immunarch](#)
2. [scRepertoire](#)
3. [dandelion](#) python package for analyzing single cell BCR/TCR data from 10x Genomics 5' solution!

4. [TRUST4](#) developed in Shirley Liu's group. Use it to extract TCR/BCR information from bulk RNAseq or 5' scRNAseq data.
5. a dramatic speedup for one of the core computations for adaptive immune receptor repertoire (AIRR) analysis - the discovery and counting of receptors that overlap between repertoires! Check out [CompAIRR](#). With 10^4 repertoires of 10^5 sequences each, CompAIRR ran in 17 minutes while the fastest existing tool took 10 days, amounting to a $\sim 1000\times$ speedup
6. [ClusTCR](#): a Python interface for rapid clustering of large sets of CDR3 sequences with unknown antigen specificity;
7. [GLIPH2](#)
8. [GIANA allows computationally-efficient TCR clustering and multi-disease repertoire classification by isometric transformation](#) from Bo Li.
9. [tcrdist3](#) is a python API-enabled toolkit for analyzing T-cell receptor repertoires
10. [TCRex](#): a web tool for the prediction of TCR-epitope recognition
11. [ImRex](#) TCR-epitope recognition prediction using combined sequence input representation for convolutional neural networks.
12. [NetTCR - 2.0](#) Sequence-based prediction of peptide-TCR binding
13. [CellaRepertorium](#)
14. [enclone](#) from 10x. we should give this a try if we want to cluster TCR and BCR clonotypes.
15. [migec](#): A RepSeq processing swiss-knife.
16. [MiXCR](#) is a universal software for fast and accurate analysis of T- and B- cell receptor repertoire sequencing data.
17. [ImReP](#) is a computational method for rapid and accurate profiling of the adaptive immune repertoire from regular RNA-Seq data.
18. [TcellMatch](#): Predicting T-cell to epitope specificity. cellMatch is a collection of models to predict antigen specificity of **single T cells** based on CDR3 sequences and other single cell modalities, such as RNA counts and surface protein counts
19. [scirpy](#): A scanpy extension for single-cell TCR analysis.
20. [Tessa](#) is a Bayesian model to integrate T cell receptor (TCR) sequence profiling with transcriptomes of T cells. Enabled by the recently developed single cell sequencing techniques, which provide both TCR sequences and RNA sequences of each T cell concurrently, Tessa maps the functional landscape of the TCR repertoire, and generates insights into understanding human immune response to diseases.
21. [DeepTCR](#) Deep Learning Methods for Parsing T-Cell Receptor Sequencing (TCRSeq) Data

https://twitter.com/John_Will_I_Am/status/1570837756787691527

<https://www.science.org/doi/10.1126/sciadv.abq5089>

22. [Integrating T cell receptor sequences and transcriptional profiles by clonotype neighbor graph analysis \(CoNGA\)](#)
23. [Echidna: Integrated simulations of single-cell immune receptor repertoires and transcriptomes](#)

15 tools/papers for multi-sample multi-group single-cell RNAseq differential expression analysis

1. [An Empirical Bayes Method for Differential Expression Analysis of Single Cells with Deep Generative Models](#) [scVI-DE](#)
2. [muscat](#)
3. [Confronting false discoveries in single-cell differential expression](#) “These observations suggest that, in practice, pseudobulk approaches provide an excellent trade-off between speed and accuracy for single-cell DE analysis.” One needs to consider biological replicates, pseudobulk works well.
4. [Modelling group heteroscedasticity in single-cell RNA-seq pseudo-bulk data](#)
5. [BSDE: barycenter single-cell differential expression for case-control studies](#)
6. [distinct](#) Both are from Mark Robinson group.
7. [nebula](#) <https://www.biorxiv.org/content/biorxiv/early/2020/09/25/2020.09.24.311662.full.pdf>
8. [Fast identification of differential distributions in single-cell RNA-sequencing data with waddR](#) <https://github.com/goncalves-lab/waddR>
9. [CoCoA-diff: counterfactual inference for single-cell gene expression analysis](#)
10. [Bias, robustness and scalability in single-cell differential expression analysis](#) From Mark Robinson group.
11. [Comparative analysis of differential gene expression analysis tools for single-cell RNA sequencing data](#) “We observed that current methods designed for scRNAseq data do not tend to show better performance compared to methods designed for bulk RNAseq data.”
12. [Tree-based Correlation Screen and Visualization for Exploring Phenotype-Cell Type Association in Multiple Sample Single-Cell RNA-Sequencing Experiments](#) TreeCorTreat is an open source R package that tackles this problem by using a tree-based correlation screen to analyze and visualize the association between phenotype and transcriptomic features and cell types at multiple cell type resolution levels.
13. [Quantifying the effect of experimental perturbations in single-cell RNA-sequencing data using graph signal processing](#) read this thread <https://twitter.com/krishnaswamylab/status/1328876444810960896?s=27>
14. [Causal identification of single-cell experimental perturbation effects with CINEMA-OT](#)

github <https://github.com/vandijklab/CINEMA-OT>

15. [IDEAS: individual level differential expression analysis for single-cell RNA-seq data](#)

10 tips for learning git

1. Several basic commands will serve you a long way:

```
git clone
git add
git commit -m
git push
```

Those are enough to get you started. To be honest, those are still the most frequent commands I use.

2. understand git and github. You use git to track files locally, and github can host your repos. You can start with the github skill page <https://buff.ly/3tO2iaf> gitlab <https://buff.ly/3JlGA69> is an alternative to github
3. software carpentry git workshop is a nice resource to learn git <https://buff.ly/3kUhqB7>
4. An open source game about learning Git! <https://buff.ly/2ZPXUrX>
5. Learn it for free on Udemy <https://buff.ly/3RvTCA9>
6. The best interactive tutorial for learning git branching <https://buff.ly/2tQTJN4> I had a lot of fun playing it.
7. <https://buff.ly/2w5p9zi> Oh Shit, Git!?! You know, sometimes it messed up so much locally I just delete my local copy and do a fresh git clone :)
8. <https://buff.ly/2U9C8hC> How to use git with R.
9. git cheatsheet <https://buff.ly/3H2PrWa>
10. if you collaborate with others, you need to understand the github flow <https://buff.ly/3CcvTio>