COMPUTATIONAL BIOLOGY SOFTWARE AND RESOURCES

PROGRAMMING

Unix Resources

Run UNIX commands online: http://www.compileonline.com/execute_ksh_online.php

R Resources

CRAN R Tutorial: http://cran.r-project.org/doc/manuals/R-intro.html#The-R-environment

GUI software: http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/

RStudio: https://www.rstudio.com

Bioconductor: http://bioconductor.org

Python Resources

Python3 (download latest Python release): https://www.python.org/downloads/

Python book: http://openbookproject.net/thinkcs/python/english2e/

Python tutorial: http://www.learnpython.org

BioPython tutorial: http://biopython.org/DIST/docs/tutorial/Tutorial.html#sec3

Coursera tutorial: https://www.coursera.org/learn/python

Anaconda: https://www.continuum.io/downloads

Perl Resources

Write and run Perl programs online: http://www.compileonline.com/execute_perl_online.php

Perl tutorial: http://perlmaven.com/perl-tutorial

Another Perl tutorial: http://www.tutorialspoint.com/perl/index.htm

BioPerl: http://www.bioperl.org/wiki/Main_Page

SOFTWARE DOWNLOADS AND INSTRUCTIONS

General Software for PC Users

Install Perl: http://www.perl.org/get.html

Install Cygwin for UNIX capabilities: https://cygwin.com/install.html

After installing Cygwin, activate OpenSSH in Cygwin: http://www.howtogeek.com/howto/41560/how-to-get-ssh-command-line-access-to-windows-7-using-cygwin/

Install Notepad++: http://notepad-plus-plus.org/download/v6.6.8.html

Install Putty (optional): http://www.chiark.greenend.org.uk/~sgtatham/putty/download.html

General Software for Mac Users

Install TextWrangler: http://www.barebones.com/products/textwrangler/

Install Xcode: https://developer.apple.com/xcode/downloads/

Linux Users

Presumably, if you use Linux, you have everything you need for basic computational biology.

Sanger Sequence Trace Viewers

4peaks (Mac): http://nucleobytes.com/index.php/4peaks

FinchTV (PC): http://www.geospiza.com/Products/finchtv.shtml

ApE Plasmid Editor (PC or Mac. More functionality than 4peaks or FinchTV but less intuitive): http://biologylabs.utah.edu/jorgensen/wayned/ape/

SeqTrace (PC, Mac. Requires Python): https://code.google.com/p/seqtrace/

FTP and SFTP Browsers

Cyberduck: https://cyberduck.io/?l=en

FileZilla: https://filezilla-project.org

Next-Gen Sequencing and Data Visualization

General Next-Gen sequencing data processing and quality control

FastQC (for FASTQ file quality control): http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

FASTX Toolkit (for data processing): http://hannonlab.cshl.edu/fastx_toolkit/

Trimmomatic (for FASTQ read trimming): http://www.usadellab.org/cms/?page=trimmomatic

Bowtie2 (for sequence alignment): http://bowtie-bio.sourceforge.net/bowtie2/index.shtml

BWA (for sequence alignment): http://bio-bwa.sourceforge.net

SAM Tools (for searching and manipulating sequence alignments): http://samtools.sourceforge.net

Picard Tools (for manipulation of sequence and alignment files): http://broadinstitute.github.io/picard/

bedtools (for manipulation of sequence and alignment files): http://bedtools.readthedocs.org/en/latest/

RNA-seq

TopHat (for mapping reads across slice junctions): http://ccb.jhu.edu/software/tophat/index.shtml

Cufflinks and cuffdiff (for RNA-seq data analysis): http://cufflinks.cbcb.umd.edu

CummeRbund (for plotting Cufflinks results; an R package): http://compbio.mit.edu/cummeRbund/manual_2_0.html

DESeq (alternative to cuffdiff for differential gene expression analysis; an R package): http://bioconductor.org/packages/release/bioc/html/DESeq.html

Trinity (for de-novo transcript assembly): https://github.com/trinityrnaseq/trinityrnaseq/wiki

Corset (for differential gene expression without a reference genome): https://github.com/Oshlack/Corset/wiki

- Corset-based example pipeline for differential gene expression without a reference genome: https://github.com/Oshlack/Corset/wiki/Example
- Homebrew (to facilitate software installation on Macs - may need for Corset installation on a Mac): http://brew.sh

edgeR (for differential gene expression analysis; requires R and bioconductor, see link above): https://bioconductor.org/packages/release/bioc/html/edgeR.html

Genome assembly (examples of short-read assemblers)

ALLPATHS: http://www.broadinstitute.org/software/allpaths-lg/blog/

MaSuRCA: http://www.genome.umd.edu/masurca.html

SOAPdenovo: http://soap.genomics.org.cn/soapdenovo.html

SPAdes: http://bioinf.spbau.ru/spades

Velvet: https://www.ebi.ac.uk/~zerbino/velvet/

Microbiome Diversiry

mothur: http://www.mothur.org/

QIIME: http://qiime.org

Data visualization and functional analysis

IGV (for visualizing sequencing data and mapped reads): https://www.broadinstitute.org/igv/

Tablet (for visualizing mapped reads and genome assemblies): https://ics.hutton.ac.uk/tablet/

Circos (for making genomic figures data): http://circos.ca

ggplot2 (for general graphics and plotting data; an R package): http://ggplot2.org

DAVID (for assigning biological function to gene lists; GO terms and KEGG pathways): https://david.ncifcrf.gov

KAAS Webserver (for mapping genes to KEGG pathways): http://www.genome.jp/tools/kaas/

Prodigal Webserver (for gene prediction in bacterial genomes): http://prodigal.ornl.gov/server.html

Local NCBI BLAST (for performing sequence similarity searches): ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/

Variant Detection

GATK (for SNP and indel calling): https://www.broadinstitute.org/gatk/

Stacks (for analysis of RAD-seq data): http://catchenlab.life.illinois.edu/stacks/