

# Bioinformatics Training Resources

<http://stephenturner.us/p/edu>

Stephen Turner, Ph.D.

August 2013

## Contents

<b>In-person Workshops</b>	<b>1</b>
<b>Online Material</b>	<b>3</b>
<b>Other Resources</b>	<b>5</b>

## In-person Workshops

**Cold Spring Harbor Courses:**[meetings.cshl.edu/courses.html](http://meetings.cshl.edu/courses.html)

Cold Spring Harbor has been offering advanced workshops and short courses in the life sciences for years. Relevant workshops include [Advanced Sequencing Technologies & Applications](#), [Computational & Comparative Genomics](#), [Programming for Biology](#), [Statistical Methods for Functional Genomics](#), the [Genome Access Course](#), and others. Unlike most of the others below, you won't find material from past years' CSHL courses available online.

**Canadian Bioinformatics Workshops:**[bioinformatics.ca/workshops](http://bioinformatics.ca/workshops)

Bioinformatics.ca through its Canadian Bioinformatics Workshops (CBW) series began offering one and two week short courses in bioinformatics, genomics and proteomics in 1999. The more recent workshops focus on training researchers using advanced high-throughput technologies on the latest approaches being used in computational biology to deal with the new data. Course material from [past workshops](#) is freely available online, including both audio/video lectures and slideshows. Topics include [microarray analysis](#), [RNA-seq analysis](#), genome rearrangements, copy number alteration, [network/pathway analysis](#), genome visualization, gene function prediction, functional annotation, data analysis using R, [statistics for metabolomics](#), and [much more](#).

**UC Davis Bioinformatics Training Program:**[training.bioinformatics.ucdavis.edu](http://training.bioinformatics.ucdavis.edu)

The UC Davis Bioinformatics Training program offers several intensive [short bootcamp workshops](#) on RNA-seq, data analysis and visualization, and cloud computing with a focus on Amazon's computing resources. They also offer a week-long [Bioinformatics Short Course](#), covering in-depth the practical theory and application of cutting-edge next-generation sequencing techniques. Every course's documentation is [freely available online](#), even if you didn't take the course.

**MSU NGS Summer Course:**[bioinformatics.msu.edu/ngs-summer-course-2013](http://bioinformatics.msu.edu/ngs-summer-course-2013)

This intensive two week summer course will introduce attendees with a strong biology background to the practice of analyzing short-read sequencing data from Illumina and other next-gen platforms. The first week will introduce students to computational thinking and large-scale data analysis on UNIX platforms. The second week will focus on mapping, assembly, and analysis of short-read data for resequencing, ChIP-seq, and RNAseq. [Materials from previous courses are freely available online](#) under a CC-by-SA license.

**Genetic Analysis of Complex Human Diseases:**[hihg.med.miami.edu/edu...](http://hihg.med.miami.edu/edu...)

The Genetic Analysis of Complex Human Diseases is a comprehensive four-day course directed toward physician-scientists and other medical researchers. The course will introduce state-of-the-art approaches for the mapping and characterization of human inherited disorders with an emphasis on the mapping of genes involved in common and genetically complex disease phenotypes. The primary goal of this course is to provide participants with an overview of approaches to identifying genes involved in complex human diseases. At the end of the course, participants should be able to identify the key components of a study team, and communicate effectively with specialists in various areas to design and execute a study. The course is in Miami Beach, FL. (Full Disclosure: I teach a section in this course.) Most of the course material from previous years is not available online, but [my RNA-seq & methylation lectures are on Figshare](#).

**UAB Short Course on Statistical Genetics and Genomics:**[soph.uab.edu/ssg/...](http://soph.uab.edu/ssg/...)

Focusing on the state-of-art methodology to analyze complex traits, this five-day course will offer an interactive program to enhance researchers' ability to understand & use statistical genetic methods, as well as implement & interpret sophisticated genetic analyses. Topics include GWAS Design/Analysis/Imputation/Interpretation; Non-Mendelian Disorders Analysis; Pharmacogenetics/Pharmacogenomics; ELSI; Rare Variants & Exome Sequencing; Whole Genome Prediction; Analysis of DNA Methylation Microarray Data; Variant Calling from NGS Data; RNAseq: Experimental Design and Data Analysis; Analysis of ChIP-seq Data; Statistical Methods for NGS Data; Discovering new drugs & diagnostics from 300 billion points of data. Video recording from the 2012 course are [available online](#).

**MBL Molecular Evolution Workshop:**[hermes.mbl.edu/education/...](http://hermes.mbl.edu/education/...)

One of the longest-running courses listed here (est. 1988), the Workshop on Molecular Evolution at Woods Hole presents a series of lectures, discussions, and bioinformatic exercises that span contemporary topics in molecular evolution. The course addresses phylogenetic analysis, population genetics, database and sequence matching, molecular evolution and development, and comparative genomics, [using software packages](#) including AWTY, BEAST, BEST, Clustal W/X, FASTA, FigTree, GARLI, MIGRATE, LAMARC, MAFFT, MP-EST,

MrBayes, PAML, PAUP\*, PHYLIP, STEM, STEM-hy, and SeaView. Some of the course materials can be found by digging around the [course wiki](#).

**UW Summer Institute in Statistical Genetics:**[sisg.biostat.washington.edu](http://sisg.biostat.washington.edu)

consists of a series of two-and-a-half day modules over a 3-week period designed to introduce geneticists to modern methods of statistical analysis and to introduce statisticians to the statistical challenges posed by modern genetic data. Prerequisites are minimal, and the modular nature of the Institute enables participants to design a program best suited to their backgrounds and interests. Most participants take two or three modules. See the [schedule](#) and the [list of SISG modules](#) for details.

## Online Material

**Canadian Bioinformatics Workshops:**[bioinformatics.ca/workshops](http://bioinformatics.ca/workshops)

(*In person workshop described above*). Course material from [past workshops](#) is freely available online, including both audio/video lectures and slideshows. Topics include [microarray analysis](#), [RNA-seq analysis](#), genome rearrangements, copy number alteration, [network/pathway analysis](#), genome visualization, gene function prediction, functional annotation, data analysis using R, [statistics for metabolomics](#), and [much more](#).

**UC Davis Bioinformatics Training Program:**[training.bioinformatics.ucdavis.edu](http://training.bioinformatics.ucdavis.edu)

(*In person workshop described above*). Every course's documentation is [freely available online](#), even if you didn't take the course. [Past topics](#) include Galaxy, Bioinformatics for NGS, cloud computing, and RNA-seq.

**MSU NGS Summer Course:**[bioinformatics.msu.edu/ngs-summer-course-2013](http://bioinformatics.msu.edu/ngs-summer-course-2013)

(*In person workshop described above*). [Materials from previous courses](#) are freely available [online](#) under a CC-by-SA license, which cover mapping, assembly, and analysis of short-read data for resequencing, ChIP-seq, and RNAseq.

**EMBL-EBI Train Online:**[www.ebi.ac.uk/training/online](http://www.ebi.ac.uk/training/online)

Train online provides free courses on Europe's most widely used data resources, created by experts at EMBL-EBI and collaborating institutes. Topics include [Genes and Genomes](#), [Gene Expression](#), [Interactions](#), [Pathways](#), and [Networks](#), and others. Of particular interest may be the [Practical Course on Analysis of High-Throughput Sequencing Data](#), which covers Bioconductor packages for short read analysis, ChIP-Seq, RNA-seq, and allele-specific expression & eQTLs.

**UC Riverside Bioinformatics Manuals:**[manuals.bioinformatics.ucr.edu](http://manuals.bioinformatics.ucr.edu)

This is an excellent collection of manuals and code snippets. Topics include [Programming in R](#), [R+Bioconductor](#), [Sequence Analysis with R and Bioconductor](#), [NGS analysis with Galaxy and IGV](#), basic [Linux skills](#), and others.

**Software Carpentry:**[software-carpentry.org](http://software-carpentry.org)

Software Carpentry helps researchers be more productive by teaching them basic computing skills. We recently ran a [2-day Software Carpentry Bootcamp here at UVA](#). Check out the [online lectures](#) for some introductory material on Unix, Python, Version Control, Databases, Automation, and many other topics.

**Coursera:**[coursera.org/courses](http://coursera.org/courses)

Coursera partners with top universities to offer courses online for anyone to take, for free. Courses are usually 4-6 weeks, and consist of video lectures, quizzes, assignments, and exams. Joining a course gives you access to the course's forum where you can interact with the instructor and other participants. Relevant courses include [Data Analysis](#), [Computing for Data Analysis using R](#), and [Bioinformatics Algorithms](#), among others. You can also [view all of Jeff Leek's Data Analysis lectures on Youtube](#).

**OpenHelix:**[openhelix.com](http://openhelix.com)

OpenHelix provides tutorials and training materials (video lectures, screencasts, slides, exercises) on over 100 bioinformatics tools, databases, and web servers. [Most of the tutorials](#) require a subscription, but there are a handful of [free tutorials](#) on topics such as [OMIM](#), UCSC genome browser ([beginner](#), [advanced](#), and [more](#)), and a few others.

**Rosalind:**<http://rosalind.info>

Quite different from the others listed here, Rosalind is a platform for learning bioinformatics through gaming-like problem solving. Visit the [Python Village](#) to learn the basics of Python. Arm yourself at the [Bioinformatics Armory](#), equipping yourself with existing ready-to-use bioinformatics software tools. Or storm the [Bioinformatics Stronghold](#), implementing your own algorithms for computational mass spectrometry, alignment, dynamic programming, genome assembly, genome rearrangements, phylogeny, probability, string algorithms and others.

**Bioconductor Course Materials:**[bioconductor.org/help/course-materials](http://bioconductor.org/help/course-materials)

Bioconductor provides training in computational and statistical methods for the analysis of genomic data. In particular, the course materials from the [useR! 2013 R/Bioconductor for Analyzing High-Throughput Genomic Data](#) are especially useful, containing R code and a PDF tutorial on sequence alignment, RNA-seq, ChIP-seq, and sequence annotation with various Bioconductor packages.

**Google Developers R Programming Video Lectures:**<http://www.youtube.com/...>

Google Developers recognized that most developers learn R in bits and pieces, which can leave significant knowledge gaps. To help fill these gaps, they created a series of [introductory R programming videos](#). These videos provide a solid foundation for programming tools, data manipulation, and functions in the R language and software. The series of short videos is organized into four subsections: intro to R, loading data and more data formats, data processing and writing functions. Start watching the [YouTube playlist here](#).

## Other Resources

- [Titus Brown's list bioinformatics courses](#): Includes a few others not listed here (also see the comments).
- [GMOD Training and Outreach](#): GMOD is the Generic Model Organism Database project, a collection of open source software tools for creating and managing genome-scale biological databases. This page links out to tutorials on [GMOD Components](#) such as Apollo, BioMart, Galaxy, GBrowse, MAKER, and others.
- [Seqanswers.com](#): A discussion forum for anything related to Bioinformatics, including Q&A, paper discussions, new software announcements, protocols, and more.
- [Biostars.org](#): Similar to SEQanswers, but more strictly a Q&A site.
- [BioConductor Mailing list](#): A very active mailing list for getting help with Bioconductor packages. Make sure you do some Google searching yourself first before posting to this list.
- [Bioconductor Events](#): List of upcoming and prior Bioconductor training and events worldwide.
- [Learn Galaxy](#): Screencasts and tutorials for learning to use Galaxy.
- [Galaxy Event Horizon](#): Worldwide Galaxy-related events (workshops, training, user meetings) are listed here.
- [Galaxy RNA-Seq Exercise](#): Run through a small RNA-seq study from start to finish using Galaxy.
- [Rafael Irizarry's Youtube Channel](#): Several statistics and bioinformatics video lectures.
- [PLOS Comp Bio Online Bioinformatics Curriculum](#): A perspective paper by David B Searls outlining a series of free online learning initiatives for beginning to advanced training in biology, biochemistry, genetics, computational biology, genomics, math, statistics, computer science, programming, web development, databases, parallel computing, image processing, AI, NLP, and more.
- [Getting Genetics Done](#): Shameless plug – I write a blog highlighting literature of interest, new tools, and occasionally tutorials in genetics, statistics, and bioinformatics. I recently wrote [this post about how to stay current in bioinformatics & genomics](#).
- [NGS Conferences](#): A comprehensive list of next-generation sequencing conferences.
- [I want to learn bioinformatics!](#): A guide by Nick Loman for helping a non-computer-literate person to learn bioinformatics
- [Bioinformatics not something you are taught; it's a way of life](#): Blog post by Mick Watson about how one doesn't simply learn bioinformatics in a week-long workshop. It takes years of practice and hands-on doing to learn how to solve problems in bioinformatics, and this post gives some great suggestions about how to get started.