

# Recommended links and literature

Alexander Jueterbock, Martin Jakt\*

## PhD course: High throughput sequencing of non-model organisms

The following sections list relevant literature, links to analysis pipelines, as well as to tutorials and learning environments that get you closer to become bioinformatics-experts.

### 1 Recommended Reading for RAD sequencing

- Baird, Nathan A., et al. "Rapid SNP discovery and genetic mapping using sequenced RAD markers." *PloS one* 3.10 (2008): e3376.
- Baker, Monya. "De novo genome assembly: what every biologist should know." *Nature methods* 9.4 (2012): 333-337
- Catchen, Julian, et al. "Stacks: an analysis tool set for population genomics." *Molecular ecology* 22.11 (2013): 3124-3140.
- DaCosta, Jeffrey M., and Michael D. Sorenson. "Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol." *PloS one* 9.9 (2014): e106713.
- Davey, John W., and Mark L. Blaxter. "RADSeq: next-generation population genetics." *Briefings in Functional Genomics* 9.5-6 (2010): 416-423.
- Hohenlohe, Paul A., et al. "Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags." *PLoS genetics* 6.2 (2010): e1000862.
- Hohenlohe, Paul A., et al. "Nextgeneration RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and westslope cutthroat trout." *Molecular ecology resources* 11.s1 (2011): 117-122.
- Peterson, Brant K., et al. "Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species." *PloS one* 7.5 (2012): e37135.
- Rasic, Gordana, et al. "Genome-wide SNPs lead to strong signals of geographic structure and relatedness patterns in the major arbovirus vector, *Aedes aegypti*." *BMC genomics* 15.1 (2014): 275.

---

\*University of Nordland, Norway

- Schweyen, Hannah, Andrey Rozenberg, and Florian Leese. "Detection and Removal of PCR Duplicates in Population Genomic ddRAD Studies by Addition of a Degenerate Base Region (DBR) in Sequencing Adapters." *The Biological Bulletin* 227.2 (2014): 146-160.
- Puritz, Jonathan B., Christopher M. Hollenbeck, and John R. Gold. "dDocent: a RADseq, variant-calling pipeline designed for population genomics of non-model organisms." *PeerJ* 2 (2014): e431.
- Puritz, Jonathan B., et al. "Demystifying the RAD fad." *Molecular ecology* 23.24 (2014): 5937-5942.
- [Blog on RAD seq](#)

## 2 Guidelines for pooled sequencing data

- Schlötterer, Christian, et al. "Sequencing pools of individuals - mining genome-wide polymorphism data without big funding." *Nature Reviews Genetics* (2014).

## 3 Useful programs and analysis pipelines

- *Molecular Ecology Resources* Jan 2017; Special Issue: Population Genomics with R
- [Popoolation](#): Population genomic analysis of pooled samples, see also [this presentation](#)
- [Popoolation 2](#): allows comparison of allele frequencies between two or more populations
- [The Simple Fool's Guide to Population Genomics via RNA-Seq](#)
- [Bioconductor](#): R packages for genomic data analysis
- [Rosalind](#): Learning python
- [Biopython](#)
- [BioPerl](#)
- [STACKS](#): building loci from short sequences and analyzing RADseq data
- [Ddocent](#): ddRAD analysis pipeline

## 4 Recommended books

- [Unix and Perl to the Rescue](#)

- Computational Biology
- Primer to Analysis of Genomic Data Using R (Use R!)
- Bioinformatics Data Skills: Reproducible and Robust Research with Open Source Tools
- Practical Computing for Biologists

## 5 Upcoming Coursera courses

- R programming
- Algorithms, Biology, and Programming for Beginners
- Epigenetic Control of Gene Expression
- Bioconductor for Genomic Data Science
- Statistics for Genomic Data Science
- Comparing Genes, Proteins, and Genomes (Bioinformatics III)
- Python for Genomic Data Science
- Command Line Tools for Genomic Data Science

Emacs 24.5.1 (Org mode 8.3beta)