

Draft notes

matthieu bruneaux

March 3, 2015

Outline

Course about RAD-tag sequencing - Roscoff 2015

Problematic

- ▶ High coverage
- ▶ Genome reduction
- ▶ Non-model species

Presentation of the RAD-tag approach (history)

- ▶ Protocol
- ▶ Single or double digest
- ▶ Single end or paired ends
- ▶ Advantage over SNP-chip: no prior knowledge of the genome is needed, less costly to put in place for non-model species
- ▶ As explained in Peterson 2012, SNP-chip or similar, once designed, can only test for the variability which was known at design time and cannot detect rare alleles or population-specific variants, even though those can be of high interest in population genetics studies

Practicals

Pipelines include STACKS and dDocent <https://ddocent.wordpress.com/ddocent-pipeline-user-guide/>:

[//ddocent.wordpress.com/ddocent-pipeline-user-guide/](https://ddocent.wordpress.com/ddocent-pipeline-user-guide/).

Teaching objectives

- ▶ Sequence quality check: trimming, quality control
- ▶ Stack building (STACKS or other tools, with or without paired-ends information)
- ▶ SNP calling and genotyping (VCFtools)
- ▶ Allele frequency estimates (individuals or pools?)
- ▶ Population genetics: comparison of microsats and SNPs trees (cf. Anti)
- ▶ Population genomics: localisation, genome profiling and Gst calculation, kernel smoothing
- ▶ Functional analysis (outside the scope of RAD-tags *sensus stricto*)
- ▶ Predicting the number of fragments based on published genome sequence and restriction sites with python scripts

Important people and groups

Three central groups

To the best of my knowledge, there are three groups which are involved in the origin of the RAD tags method and which have produced a lot of the initial papers. They are also producing a lot of the software that can be used for RAD data analysis. The groups are:

- ▶ Eric A. Johnson's group (Institute of Molecular Biology, U. Oregon): [link](#)
- ▶ William A. Cresko's group (Institute of Ecology and Evolution, U. Oregon): [link](#). The Cresko lab is responsible for the Stacks pipeline that can be used to analyse RAD data.
- ▶ Paul A. Hohenlohe (Depts of Biological Science and Statistics, U. Idaho): [link](#). They also have some software and their news page has some interesting information about what's going on with RAD.

People

Other interesting papers

- ▶ Hohenlohe 2010, "Using population genomics to detect selection in natural populations: key concepts and methodological considerations" ([link](#)).
- ▶ Ellengren 2014, "Genome sequencing and population genomics in non-model organisms" ([link](#))