GBS data analysis

Step 1: Demultiplex

Here is the link to the blog post for demultiplexing:

<http://www.zoology.ubc.ca/~rieseberg/RiesebergResources/gbs-two-enzyme-demultiplexing/>

Here is the script:

<https://github.com/owensgl/reformat/blob/master/GBS_fastq_Demultiplexer_v9_2Enzyme2barcode_withmismatch.pl>

Step 2: Quality check

Blast using Ugene <http://ugene.net/>

FastQC report: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>

Step 3: Align using PyRAD or Ddocent

<http://ddocent.com/>

<http://ddocent.com/assembly/>

Step 4: Filter further

This data was further filtered in *vcfR* using read depths (DP) and mapping qualities (MQ). Data was filtered as follows:

* A minimum DP of 5x.
* Variants in the top 5% of the DP distribution were removed.
* Only variants with a MQ greater than 40 were retained.
* Variants with more than 60% missing data were removed.

Step 4: Population genetic analysis on SNP’s:

<https://grunwaldlab.github.io/Population_Genetics_in_R/gbs_analysis.html>