**RADseq Cheatsheet**

*Restriction Site:* 4-8 nucleotide sequences in genome; recognized by restriction enzymes *Restriction Enzyme:* Enzyme to cut at a restriction site *Read:* A set of bp obtained via RADseq. Of a target size *FastQ:* A filetype storing both quality score information and the raw data. *Barcode:* Added sequence of nucleotides so samples can be identified *Demultiplex*: Sort fastq files by barcode to get sets of reads tagged to individual samples.

Two Major pipelines for today: STACKS and pyRAD. Stacks is often controlled as individual pieces (i.e.; you will have multiple scripts to run this pipeline), pyRAD has one control file, though you may choose to run steps separately to do error-checking. The following table compares some of the major parameters (that I've used) between these two software packages:

|  |  |  |
| --- | --- | --- |
| **Stacks Parameter** | **Meaning** | **pyRAD equivalent** |
| Process\_rad: Barcode Distance | How many mismatches are tolerated | option 19: MaxM |
|  | between barcode in read and provided barcode |  |
| ustacks: m | Minimum depth of coverage required to create a stack | Option 8: MinDepth |
| ustacks: M | Maximum mismatches allowed between stacks | Option 10: wclust (clustering percentage) |
| populations: r | minimum percentage of individuals in a population | Option 12: MinCoV |
|  | required to process a locus for that population. |  |
| populations: p | minimum number of populations a locus must | Option 12: MinCov |
|  | be present in to process a locus. |  |