

## ddRAD Data Delivery Specifications

### Output data info

- FASTQ.GZ files are provided as de-multiplexed and uniformed length after removal of inline barcodes (de-multiplexed reads are 15bp shorter than nominal run cycles). Adapter contaminants are masked by N characters in the sequence.

### ddRAD standard bioinformatics analysis

- A consensus sequence of all the detected loci in FASTA format (de novo approach).
- A list of all sites included in the catalog of loci in VCF format.
- Population-wise SNP and haplotype calls in VCF format.
- Polymorphic sites in Structure, GenePop and PLINK formats.
- Population genetic statistics calculated for each variant site and locus in tabular format.
- Sequencing and coverage metrics in tabular format.
- A REPORT file describing the library preparation and analysis flow.