Read data

Read data can be provided in one the following formats:

BAM

BAM file names need to end up with the .bam suffix and prefix shall be sample_id (e.g. ' sam1.bam'). It also needs to be readable by standard bioinformatics softwares. It must be readable by SAMtools.

CRAM

CRAM file names need to end up with the .cram suffix and prefix shall be sample_id (e.g. ' sam1.cram'). It also needs to be readable by standard bioinformatics softwares. It must be readable by SAMtools.

FASTQ

Fastq file names need to use sample_id as its prefix. The suffix shall be '.fq', '.fq1', or '.fq2'. '.fq' shall be used for single end read, and '.fq1', or '.fq2' shall be used for paired end reads. Example of accepted name Single and paired reads are accepted a fastq files that meet the following the requirements:

Quality scores must be in Phred scale and offset of either 33 or 64 are acceptable.

Pleas don't provide the technical reads (adapters, linkers, barcodes).

The first line for each read shall start with '@'.

The base calls and quality scores must be separated by a line starting with '+'.

Please always compress the fastq files. The Fastq files must be compressed using gzip or bzip2.