

1. Select read type: single-end or paired-end (cannot have both types)
2. Click '+' button on the bottom left to add sample

Fig 1.1: Add sample

3. Sample name
 - a. Each sample name entered will be used as the folder name containing 1 folder (for single-end reads) or 2 folders (for paired-end reads), so follow folder naming conventions.
 - b. Ensure that each sample name is **unique**
4. Uploading fastq zip file
 - a. Zip file format: When extracted, **folder name** must be the **same as zip file name**

	dm_female_a.zip		Today at 1:12 AM	1.83 GB
	dm_female_a		when extracted	Today at 12:29 AM
	SRR999252.fastq.gz		Today at 12:31 AM	665.2 MB
	SRR999251.fastq.gz		Today at 12:29 AM	628.4 MB
	SRR999250.fastq.gz		Today at 12:11 AM	534.8 MB

Fig 1.2: Zip file format

- b. Upload for single-end reads: 1 zip file per sample

Fig 1.3: Single-end upload

c. Upload for paired-end reads:

- 1 zip file containing only R1 reads, 1 zip file containing only R2 reads per sample.
- Ensure that each pair of files (belonging to the same fastq run) have the same order in both zip files.

The form is titled 'Choose Read Type' and has two radio buttons: 'single' and 'paired', with 'paired' selected. Below this, there is a 'Sample Name:' text input field. Under 'Select Read 1 FASTQ File', there is a 'Browse...' button and a 'No file selected' status. Similarly, under 'Select Read 2 FASTQ File', there is a 'Browse...' button and a 'No file selected' status. At the bottom, there is a 'FASTQ Suffix:' text input field and a blue 'Unzip FASTQ' button. A '+' button is located at the very bottom left of the form.

Fig 1.4: Paired-end upload

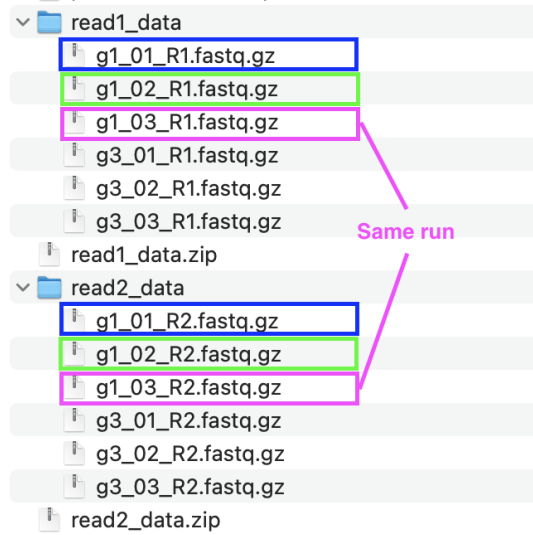


Fig 1.5: File order (Paired-end upload)

- d. Individual raw fastq files can easily exceed 3GB and running multiple concurrently may crash the server. Hence with large fastq files, it's recommended to upload each fastq file as a separate zip file and under a different sample name.

5. Fastq suffix

- Suffix refers to the text to be excluded from the end of the fastq file name, with the remaining text used to identify the fastq file within the sample.
(note: These identifiers will serve as column names for the fastq files in the raw count .csv files.)

b. Single-end reads:

Filename (SRRXXX.fastq.gz) = Identifier (SRRXXX) + Fastq suffix (.fastq.gz)



Fig 1.6: Fastq suffix (single-end)

c. Paired-end reads:

- i. Filename of **R1** (SRRXXX_R1.fastq.gz)
= Identifier (SRRXXX) + Fastq suffix (_R1.fastq.gz)
- ii. For paired-end reads, suffixes will **only be taken from R1** and both R1 and R2 fastq files (belonging to the same fastq run) will use the remaining text from R1 file name as an identifier.

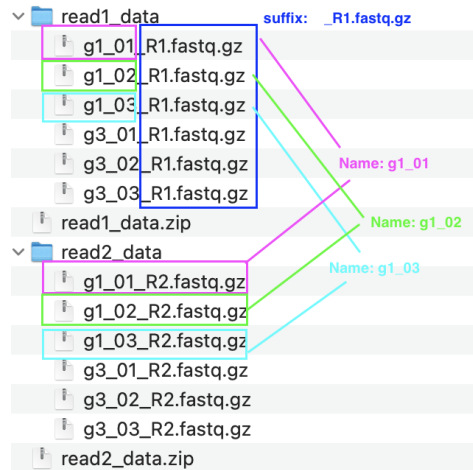


Fig 1.7: Fastq suffix (paired-end)

6. Ensure that each zip file has finished uploading completely before uploading another zip file or proceeding (else shiny application might get stuck especially with large files).

Fig 1.8: Upload incomplete

Fig 1.9: Upload completed, file unzipped

7. Click the unzip button and wait for the icon to change to 'Unzipped!'