

[illegible]

1. Mio reads output

Standard outputs are 130 mio (Mid output) or 400 mio reads (High output).

Users can ask that we perform a **spiking** of a library into another run, when it is a non-standard library type and a significant chance exists that sequencing could fail. When requesting a spike:

- Avoid absolutely mistakes of indexes (see comments in index section).
- Let us know if you plan to run it later in a full run, to make sure we will still have enough of it.
- If you have several samples to spike, let us know if they can be sequenced in distinct runs.
- Spikes cannot be allocated more than 2% of a run in total. If you need significantly more than that, please rather consider a shared run.

When users can't fill up a full run, they can request **sharing** a run with another lab. It may take several weeks to fulfill such demand. When requesting a shared run:

- Ask beforehand if a run is already available to share, in order to use compatible indexes.
- Avoid absolutely mistakes of indexes during prep (see comments in index section).
- Shared runs allow to minimize costs for the users but come with more uncertainties regarding read numbers, so extra safety margin must be included.

2. Reads configuration

For paired-end reads, the reads length you want should be consistent with the mean size of your libraries to avoid too many overlapping reads. For that, remember that your mean library size comprises ca 140bp of adapters. For PE75 (2x75nt), overlapping reads should not be an issue unless your mean size is very short (<320bp). For PE150 (2x150nt) run, though, the mean size should be of around 500bp ideally.

The quality of the ends of 150nt reads decreases significantly after 130nt, so depending on your aims you may want to use only the first 130 nt of the reads in downstream analyses.

Whenever possible (depending on different parameters such library method, indexes...), we generate reads that are a little longer than specified in the kit, for instance 80-85nt instead of the 75nt mentioned in our list price. If you plan to compare fastq files over a large series of sequencing runs, you may want us to stick to the nominal read length for consistency (for instance 75nt). Please indicate us to do so in the comments section, if needed.

When requesting a spike, indicate minimal reads configuration requirements.

3. Library method

Indicate the library preparation method/kit. When running custom/non-commercial libraries:

- If using a new technique for library prep, tell us and we'll first do a spike. Without that, we cannot give guarantees in case of run failure, as the profile and quantity may look fine but the library structure may be faulty.
- Custom libraries (ATAC-seq, HiC...) should be quantified by qPCR to get consistent results. In many cases results are good-enough without qPCR quantification, but we cannot give guarantees regarding outcome without it. In some extreme cases, the qPCR does not perfectly reflect the

behavior in the sequencer, in these cases we cannot be held responsible for issues of reads abundance (provided we did not make an error).

- CAUTION: the NextSeq will only work with modern P5 and P7 adapters also known as paired-end adapters, but not with the original obsolete version (such as used in STRT-seq or in the initial 4C-seq protocol for instance).
- When doing a custom library type, no amount of reads can be guaranteed. We do our best to give you satisfactory results though.

4. Indexes

Indicate unambiguous index names (original names when commercial). We'll ask for sequences if we don't know these indexes.

If less than 4 samples are sequenced, base diversity of index reads may be too low, please contact us to discuss which indexes should thus be picked up. In particular, when using less than 4 indexes, avoid using nextera indexes N705 and N718, truseq LT index 11, NEBNext index 11, custom index Ad2.5, and sureselect XT indexes B05 and F05.

When mixing libraries indexed with different sets of indexes, you can use our homemade website tool to check compatibility, Tindex: <https://gecftools.epfl.ch/tindex> . Ask us in case of doubt.

IMPORTANT: it is crucial that users requesting spikes are 100% sure of the indexes they used for generating their libraries. In case of error, the main run may be destroyed if the main user also used that index. In such case, the user having requested the spike will be charged for the whole run. In the same manner, in case the main user announces a wrong index, its library may be affected by a spike. In that case, the main user is responsible and no discount will be granted. The same logic applies with shared runs.

5. Qubit values

To get accurate and consistent qubit values, we recommend: using the "1x dsDNA HS" kit; measuring all samples in parallel; running regularly new standard curves; and measuring regularly a known positive control library.

In case you use other methods (dsDNA BR, picogreen, etc.), please specify it in the comments.

Provided enough volume is available, we'll check qubit values, and make a mean of both measurements.

We cannot be held responsible if a run fails due to inaccurate quantification values provided by a user.

6. Profiles analysis

Submit the pdf with mean smear size in a region of 150-1'000nt (we can help).

Profiles labels in the pdf must match the library names or tubes labels in excel file.

We have no way to double-check raw profiles when they were not performed at GECF, so we'll have to blindly trust the values you give us in that case.

In case only an agarose gel image is available, indicate the average size of the library and submit a properly labeled image (including ladder labels).

Profiles should meet following criteria:

- no adapter dimers (120-140nt peak). Small amounts and in no more than 20% of the samples can be acceptable, but will still lead to loss of informative reads. An additional bead selection will help.
- main part of the smear should be below 700bp (if not, discussion is needed, as qPCR quantification and spiking may be necessary). Fragments bigger than that will be sequenced only inefficiently.
- smear should be bell-shaped unless specific case
- reasonable amounts of contaminating primers is not an issue (<60nt).

7. Mio reads requested

When libraries were prepared by GECF, we guarantee a read output per sample of at least 80% of the anticipated number of reads.

When mixing different libraries types, it is difficult to predict reads output accurately due to their different sequencing efficiency. Therefore, in such case, allow for extra safety margin when designing your run.

Custom libraries often give 5-10% more percentage of undetermined indexes (with homemade Tn5 for instance) → take that into account when computing required reads/sample.

Spikes cannot be allocated more than 2% of a run in total.

8. Run # (when distributing libraries over several runs)

Think about batch effects when designing the sequencing runs. Typically, different biological groups should not be processed isolated in different runs, but should be mixed within these runs. We can also run a single pool of libraries over several runs, and then pool data.

Versions log

- v1.01: initial version.

- v1.02: Modified the minimum quantity to submit to “7ul at at least 1ng/ul”.