



Libraries submission guidelines for NGS

Submission process

In short: Fill in the excel submission form found on GECF website, and submit it by email to gecf@epfl.ch together with the profiles. Submit your tubes in a clear bag, labeled with your name, lab PI, short description and date. Place the bag in submission drawer of our -20°C. Once we have received both the submission form and the samples, the project enters the queue.

Important new information: custom oligos used to prepare libraries for Aviti sequencing must be HPLC-purified. If you're new to Aviti, look for the few other Aviti-specific recommendations along the file.

Notes about the form:

- Fields highlighted in red in the form are mandatory. In case of significant issues in the form, profiles or tubes labels, the submission may go down in the queue until the issues are fixed.
- If you need to modify or comment your submission, do it by email even when you have already told us in person.

Notes about the tubes:

- Libraries must have been cleaned up before submission with beads or columns. For elution, H₂O, Tris or TE are acceptable.
- Preferred tubes are 0.5ml microtubes or larger. Strips are accepted for large number of samples.
- Never take back a tube from that drawer, or swap a tube, without notifying us in advance.

The GECF is an open-access facility, therefore we cannot formally guarantee the confidentiality of data. If this is an issue, the easiest solution is to code your samples names. Upon request we can create access-controlled folders for data delivery.

Libraries amount to submit

The minimum amount for each library (of standard size) is >7ul at >1ng/ul, or >4nM by qPCR, but larger volumes and higher concentrations are preferred, and sometimes required (see below). If you cannot reach these values, we will apply when possible a special protocol. Yet, this may reflect an underlying prep issue, often leading to low yield/high duplication rates and thus without guarantee of reads output.

In addition, for the NovaSeq there are constraints on the minimal total volume to submit for the full set of libraries (again more volume is better):

NovaSeq 800mio	55ul
NovaSeq 10'000mio	200ul

If these values cannot be reached, we can discuss an alternative workflow for an additional fee.

Practical examples, considering libraries at >1ng/ul:

Aviti with 1 lib: >7ul of this library
 “ “ 20 lib: >7ul of each library

NovaSeq 10'000mio with 1 lib: >200ul of this library
 “ “ “ 5 lib: >40ul of each library
 “ “ “ 20 lib: >10ul of each library

These examples assume equal distribution of reads across libraries. The volumes will need to be adjusted in case of very uneven distribution.

Additional details regarding fields in the form:

Submission date:		DD.MM.YYYY										For GECF, do not use.
User(s):		First name, last name										
Lab PI:		Last name (not Unit)										
Very brief exp description:		E.g. RNA-seq on mouse T cells, HiC of human breast tumours, diverse chip-seq...										
Billing to (PI):		Also indicate billing address if not at EPFL										
1 Sequencer and total mio reads requested:		Scroll for all aviti, novaseq and miseq options.										
2 Reads configuration		First select sequencer. If spikes, indicate in comments the minimal configuration required.										
Custom sequencing primer needed?		If yes provide it, or check with us that we have it in sufficient amount.										
Commercial or custom indexes adapters used?		If custom, indicate if HPLC-purified, as this is important for aviti sequencing.										
Low complexity libraries?		E.g. amplicons, libraries containing invariable regions (4C primers, "samples barcodes")...										
Trimming needed?		If you are unsure whether trimming is needed, inquire with us or with your bioinformatician.										
Similar to a tricky previous run?		Indicate if libraries quality/settings are peculiar and similar to a previous tricky run of yours.										
If >30 samples, still check all qubits or only pool?		We charge a modest fee for checking quantification above a certain number of samples.										
Are these aliquots that can be discarded?		If yes we will discard them without prior notice after a few months.										
Is turnaround of >2 weeks OK?		If you or the bioinformatician										
3 Comments? 4 Special requests? 5		E.g.: CellRanger needed?; If										
GECF ULX code. Do not use.	6 Tube label please max 10 characters	7 Name for fastq files • only _ and - allowed • Avoid "S1", "S2" ... • Avoid upper case "R1" & "R2"	8 Library method	9 User name	10 Index set	Index1 (i7)	Index2 (i5)	Qubit (ng/ul)	Smear mean size (region 150-1000nt)	qPCR-based nM	Mio reads /sample	

1. Sequencer and total mio reads requested

Read configurations options change depending on the sequencer and mio reads selected. Sometimes the choice between Aviti and NovaSeq will require discussions, in such case choose "TBD".

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 fails. When requesting a spike:

- Avoid absolutely errors 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.
- Spikes cannot be allocated more than 10-20 mio reads in total. If you need significantly more, rather consider an Aviti 100mio, a MiSeq run, or a shared run if the project is not time-sensitive.

2. Reads configuration

For paired-end reads, the selected reads length should be consistent with the mean size of your libraries to avoid overlapping reads. Remember that your mean library size comprises 140bp of adapters.

On Illumina instruments, 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 130nt of the reads in downstream analyses.

Whenever possible (depending on parameters such library method, indexes...), we generate reads that are a little longer than specified, for instance 80nt instead of the 75nt mentioned in our list price. If you plan to compare fastq files over different projects, ask us to stick to the nominal read length for consistency.

When requesting a spike, indicate minimal reads configuration requirements.

Aviti-specific note: Aviti chemistry is based on rolling circle amplification, therefore DNA clusters on the flow cell consist of multiple copies of the libraries forming concatemers. If the extreme case of an insert (=library size minus 140nt) >100nt shorter than the read1 length, the sequencing will extend deep into the next copy in the concatemer, impacting negatively on the quality of the run. If the pool is highly enriched in such fragments, this can lead to very poor read2 quality and lower yield.

3. Tube label

Tube labels must be unambiguous, *ideally* max 10 characters, and perfectly match the ones indicated in the excel sheet. Leave some space on the side of tubes for our internal codes.

4. Sample name

Do not use “S1”, “S2”..., nor “R1” or “R2”... in your samples name, as these codes are already appended to fastq files by the sequencers.

Only special symbols tolerated by the sequencer are – and _ .

5. Library method

Indicate the library preparation method/kit. If you use a new custom technique for library prep, tell us and we may 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. A qPCR is also useful (see below).

6. Adaptors and Indexes

Indicate unambiguous index names (original names when commercial).

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

NovaSeq is subject to index hopping issues (indexes “jumping” from a fragment to another in the libraries pool). To alleviate this issue, which can impact a few percent of reads, use only “UDI” indexes. In addition, remove free adaptors (sharp peaks below 100nt). Aviti is less sensitive to index hopping, but the issue is still present to a lower extent, therefore the usage of UDI indexes is universally recommended.

IMPORTANT: it is crucial that users requesting spikes are 100% sure of the indexes they used. In case of error and thus indexes overlap, the main run may be destroyed and the spiking user may 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.

7. 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 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 (for free up to 30ish samples), 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.

8. 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 given to us in that case.

If only an agarose gel is available, indicate the average size of the library and submit a properly labeled image (including ladder sizes). Gels are not recommended as they often fail to detect bands below 200nt.

Profiles should meet following criteria:

- no adapter dimers (120-150nt 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. *Aviti is more sensitive to presence of adapter dimers, so they should be avoided as much as possible, this is even more important with reads longer than 75 nt. If a cleanup is not possible due to low concentration, you can expect that yield may be more impacted than for Illumina runs, but this in general only reflects the more stringent quality filtering operated by the Aviti software.*
- main part of the smear should ideally be below 800bp (if not, discussion is needed, as qPCR quantification and spiking may be necessary). Fragments bigger than that will be sequenced only inefficiently and may lead to overall decrease in run quality.
- Note: it is difficult to balance the amounts of reads when profiles have very different shapes.
- *Aviti-specific note:* Sizes of each library type should be roughly similar (e.g. all <300nt, or all 300-700nt, or all >700nt).

9. qPCR-based nM

Certain complex custom libraries (HiC, 4C-UMI...) should be quantified by qPCR to get consistent results. Results are often good-enough without qPCR quantification, but we cannot give guarantees. 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 output.

Ask us for a protocol if needed. We recommend the JetSeq Library Quantification Lo-ROX Kit.

10. 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 for Illumina sequencers, and 75% for the Aviti.

Aviti-specific points:

- For commercial libraries methods, the yield of Aviti runs corresponds to the expected nominal yield of the flow cell. For custom methods, it is expected to obtain a little less, especially for first runs (optimization needed).
- For custom preps, if oligos are not HPLC-purified, it is expected that yield may be lower overall, and very significantly lower for certain libraries. HPLC oligos are thus very strongly recommended for Aviti runs. If not possible, let us know and we may run a spike of the pool to check for an issue.
- When libraries have issues, Aviti yield may be more strongly impacted than for Illumina runs, but this in general only reflects the more stringent quality filtering operated by the Aviti software.

When running a custom library type, no amount of reads can be guaranteed. We do our best to give you satisfactory results though.

When mixing different libraries types:

- Plan a significant safety margin in reads distribution (at least 25%), as the behavior of each library type is difficult to predict and cannot be guaranteed, for the first run at least.
- In particular for the Aviti, sizes of each library type should be roughly similar (e.g. all <300nt, or all 300-700nt, or all >700nt).
- It is difficult to pool different libraries containing low complexity regions at different places.

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 10-20 mio reads in total.

Versions log

- v1.04-v1.07: Major edits. Clarified to use strips only for large numbers of libraries. Added recommendations for qPCR. When quantified by qPCR, libraries should be >4nM. Updated “Sequencer and total mio reads requested” fields guidelines.
- v1.08: aviti-specific changes
- v1.09 (18.09.2025): More Aviti-specific points. Pooling different library types disclaimers. Removed the mention that free adapters are not an issue. HPLC requirement for custom oligos for Aviti. Tindex link fixed.