ChIPed DNA submission guidelines for ChIP-seq

How to submit your ChIPed DNA samples?

- Fill in the **ChIPed DNA submission excel file** found on our website and email it to us. All relevant information should be included in this form, nothing should be communicated only by email or orally.

- The minimal QC to be provided is DNA quantification, ideally by qubit. Nanodrop can also be suitable for DNAs that are concentrated-enough to allow for a reliable measurement (above ca 30ng/ul).

- A profile analysis is recommended before starting the prep. We can do it, or we can train you on our Fragment Analyzer for you to do it (recommended as it streamlines the workflow).

- Bring DNAs on ice, in DNase-free PCR-grade 0.5ml-2ml microtubes, but not in PCR tubes/strips/PCR plates. Leave some space on the side of tubes for GECF codes.

- Indicated tube labels must match exactly the ones on your physical tubes. If available, labels of profiles must also match perfectly either tube labels or sample names.

- Avoid using "S1", "S2"… in your samples name. These codes are already appended by the sequencer to fastq files, which causes confusion.

- If you need to **modify** or comment your submission in any way, do it by email, even when you have already told us orally.

ChIPed DNA preparation

- Measuring **DNA concentration** by nanodrop may be very inaccurate, in particular at low concentrations (<20ng/ul). Therefore, always report 260/280 and 260/230 ratios when nanodrop is used. Alternatively, a more accurate and sensitive method is Qubit DNA HS (available at GECF).

- Depending on your project/aims, the lower limit to the **DNA quantity** required for library prep will differ. Thus please contact us in advance to discuss it.

- **DNA elution** should ideally be done in PCR-grade H2O, Tris or "low-EDTA TE" (10mM Tris, 0.1mM EDTA).

- **DNA storage** should be done at -20°.

Recommendations for ChIP-seq experimental design

These recommendations are just a help to avoid most significant errors during Chip-seq, but for more in depth ones have a look at https://www.encodeproject.org/data-standards/chip-seq/

- When **designing your ChIP-seq experiment**, consider these points:

- Include a total input for each chromatin sample. We will perform the total input library prep starting from the amount of DNA corresponding to the chipped samples with the highest concentration. If this is between 0.5 and 10ng, we'll do it with 10ng to insure good input sample

data. If this is below 0.5ng, we'll do it with 5 ng to reduce the delta of concentration. Let us know if you would like a different strategy.

- validate antibody specificity in standard protein IP at least once at beginning of experiment.
- If available, validate antibody/method efficiency by qPCR quantification on a known enriched target.
- At least duplicate biological replicates are needed, but many/most bioinformaticians ask for **triplicates**, check that in advance.
- Facultative: have a cell line not expressing your IP target (not transduced, KO…)
- Facultative: for ChIP of a modified histone, have a control ChIP on the non-modified histone, to check difference is not simply due to difference in quantity of that histone but really due to difference of modification level.

Determining library prep method and sequencing parameters

- The choice of a library prep method depends directly on results from DNA quantification, and will be a shared decision between the users and the GECF.

- Single-end reads (SR) or paired-end (PE) reads? consensus is clearly PE for chip-seq. This allows to identify clearly the bound region, which is not possible with just SR, as the ChIPed fragment could be +/- 200nt in a typical library.

- Read length? If the targets are typically not thought to bind specifically to repetitive regions and if the user wants to minimize costs, PE37 is enough. If cost is not a critical issue, and/or if targets are repetitive, PE75 is preferred as mapping will be better (fewer multiple mappings).

- Mio reads/sample? This strongly depends on the target ChIPed protein, and on the size of the genome. For mammalian TFs, a good starting point is 15-20 mio reads/sample.

Bioinformatics

- The GECF does not offer in depth bioinformatics analysis. If needed contact for instance the EPFL bioinformatics facility.

Disclaimers

- The GECF guarantees successful library prep when using a commercial method starting with ChIPed DNAs of concentrations and quality matching that method's specifications. In case a library prep fails or does not lead to satisfactory results due to technical problems, the GECF will prepare it again without additional charges. Yet, when it turns out that the issue does not come from the GECF (e.g. problem is reproducible, and/or a control library is successfully prepared in parallel), the libraries prep will be billed to user. Typical, but very rare, examples of GECF-independent issues are:

- Library prep is inhibited by a contaminant present in the gDNA samples. These can be difficult to detect and can potentially harm library preparation.

- A wrong DNA concentration was indicated by user. In case of suspicion of a quantification issue and when sample quantities allow, we independently check user-provided concentrations. Yet sometimes sample quantities do not allow for this double-checking.

- When the GECF and users agree to perform an exploratory method, or to process samples even though they do not comply with recommended specs for the chosen method, the GECF cannot be held responsible for failures or unsatisfactory results (outside obvious errors from the GECF side).

- The GECF is an open access facility, therefore we cannot give guarantee regarding the confidentiality of data. If this is an issue, the easiest solution is to code your sample names.

Versions log

- v1.01: initial release.

- v1.02: minor edits. Clarification regarding the input control starting amounts that we'll use for library prep. Clarified that the submission form itself must contain all information, and that DNA quantification is mandatory. Suggested that users perform profile analysis to streamline the workflow.