

Users guide to DNA sample submission

There are a number of things to take into consideration when preparing amplicon pools from environmental samples for large-scale sequencing. Source, quality and amount of DNA are important for amplicon preparation. This is a guide of issues to take into consideration when preparing amplicon pools.

1. If you plan to send environmental samples for DNA extractions, please contact us to discuss methodology. Only samples in aliquots ready for extraction is accepted.
2. **DNA quality.** The DNA quality is affected by sample type and extraction method. If possible, design your experiment in such a way that a certain loss of samples is acceptable. This is because a number of samples will not amplify during PCR due to low DNA quality. Since PCR primers for large-scale sequencing preparation are longer than regular primers, the PCR tends to work less efficiently. When receiving samples, we measure quantity and quality of the DNA using Nanodrop. If the samples pass this test, we do a PCR to test if the DNA can be amplified. After this, we will contact you with information about the likelihood of success for your samples.
3. **Amount of DNA.** The DNA amount needed for PCR varies with sample type and target gene, but minimum requirements are concentration, >5 ng/ μ l and volume >20 μ l. However, more DNA is desirable. A general rule is that the more DNA, the better. If the DNA concentration is low it is important that there is no EDTA or other substances which can disturb the PCR, e.g. humic substances, in it.

The final amount of DNA needed for large-scale amplicon sequencing varies with sequencing method and where sequencing is done, but here are some guidelines for preparing amplicons to send to SciLifeLab.

PacBio: a pool should have at least 250-500 ng amplicon DNA at a concentration of at least 1 ng/ μ l. To reach this amount we recommend running the PCR in at least duplicates if there are less than 60 samples in a pool. With more samples in a pool, PCR

replicates are usually not needed. This the standard method for ITS sequencing.

Illumina (MiSeq paired-end reads): The amount and concentration of DNA are dependent of the sequencing protocol; contact the sequencing facility to get latest information. This is the standard method for 16S and 18S rDNA gene sequencing (if not too long fragments).

- 4. Primers.** The primers used to prepare amplicons for large-scale sequencing consist of several parts. The actual PCR primer with appropriate specificity, a tag for sample identification and sometimes an adaptor. The primer is often very long, which reduces PCR efficiency. The primer should be kept as short as possible. The gene-specific primer should be kept to 18-20 bases, if possible. A tag with 8 bases is usually enough for several hundred of tags. The last base of the tag (5' of the specific primer) should not bind to any of the potential targets. If necessary, insert an extra non-binding base between the tag and the specific primer. The adaptor does not have to be part of the primer, it can be ligated on afterwards. The 5' end of the primer should not vary, one or two identical bases should be included 5' of the tag.

- 5. Samples and sequencing depth.** Normally, each sample is tagged individually to be able to pool a number of samples. The number of samples in each pool determines the amount of sequences that can be obtained per sample. This needs to be decided before submitting your samples for sequencing. If you want to run sample replicates you need to submit the replicates as individual samples so they get individual tags.