

## Sample requirements for DNA or 1st round PCR products submitted for generation of Illumina amplicons

The CGR cannot guarantee good results from samples that do not meet the requirements set out in this document. The submitted gDNA samples will not undergo QC, unless otherwise requested, but will be entered into the 1st round PCR at the submitted concentration. Optimal extraction methods vary depending on the input material and you may have to consult manufacturers to decide which is most suitable for your samples

### Experimental design

We recommend that no more than 94 samples are included in any submitted 96-well plate. This allows us to include a positive control for some projects as well as a negative control with each project/plate. These PCR control samples will be processed by the CGR at no additional project cost.

We recommend inclusion of control samples the following:

- **Extraction kit negative control:** A mock nucleic acid extraction, using the reagents provided in the extraction kit and following the protocol, but without addition of any actual sample material. If applicable, lot-to-lot variation between extraction kits should be avoided. This may be easier to achieve if kits are purchased at the same time. However, if multiple lots of kits are used, this should be incorporated into the statistical analyses.
- **Extraction kit positive control:** Extraction from a mock community of known composition is recommended to determine the success of the extraction method.
- **PCR negative control:** A no-template control (NTC) omits any DNA or RNA template from the PCR reaction and serves as a general control for extraneous nucleic acid contamination.

In addition, we strongly recommend that you include the extraction kit controls as part of your own samples, which will be included on the same sequencing run.

Some suggested publications addressing considerations in experimental design are listed on page 3.

### Amplicon design strategy

We strongly recommend an overlap of >100 bases for the forward and reverse reads of the intended run configuration. The quality of reads will always tail away towards the end of the sequencing. This is particularly prominent for longer reads (2x250 or 2x300 cycle sequencing) and any experimental design should take this into account.

For amplicons targeting the 16S rRNA gene V4 region (16S v4), we have successfully used the following primers, which (in red) include annealing sites for the 2<sup>nd</sup> round PCR as well as the Illumina sequencing primer:

- Forward primer: 5'**ACACTCTTCCCTACACGACGCTCTTCCGATCT**NNNNNGTGCCAGCMGCCGCGGTAA3'
- Reverse primer: 5'**GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GGACTACHVGGGTWTCTAAT3'

### Pre-submission amplification test

Samples are likely to contain a varying proportion of gDNA from non-target organisms and, thus, it is difficult to normalise the actual input into the PCRs although an increasing number of PCR cycles will introduce more biases during the amplicon generation.

Our standard 16S v4 protocol includes a total of 25 PCR cycles (10 first round PCR cycles to amplify the target region and 15 second round PCR cycles to incorporate indexed Illumina adaptors). These parameters work well for most samples, but we occasionally encounter samples which require a greater number of PCR cycles.

We request a test PCR before shipping the samples. You will perform a PCR experiment with 25, 30 and 35 cycles for a subset (or all) of your samples and check that these produce visible bands of the expected size on an agarose gel/automated electrophoresis instrument such as Bioanalyzer, TapeStation, Fragment Analyzer or equivalent. We suggest that your test include samples that span the entire range of sample sites, conditions and concentrations obtained for your samples as well as a positive and a negative control. Please upload the test outcome document and indicate the total number of PCR cycles required to obtain visible correctly-sized products. If you fail to provide

the information, we are not able to guarantee any library generation and/or sequencing performance and you may incur in additional cost.

### **First-round amplicon submission**

Please ensure that the introduced overhangs are compatible with the target regions for the primers used in the 2nd round PCR and for sequencing at the CGR. More details can be found in the document [Dual index nested PCR primer design](#).

If the Pre-submission amplification tests (see above) indicated that a total of 25, 30 or 35 PCR cycles were required to obtain visible correctly-sized products, please carry out 10, 15 or 20 PCR cycles, respectively, for the first round PCR products submitted to the CGR. This allows us to complete the amplicons by use of 15 PCR cycles to add the indexed adaptors without the risk of over-amplification. If more are required, it is important to make us aware of this in advance.

Unfortunately, if the amplification generation fail due to the nature of the samples, we must charge for the work performed up until that point.

### **Sample submission requirements**

We ask for the following:

- Amplicons are free from contaminants, as enzymatic steps in the library preparations could be inhibited. This, however, does not apply if your quote is for submission of un-purified first round PCR products
- Only when DNA is provided, quantify the samples by use of a dye-based method such as Qubit (ThermoFisher Scientific) rather than a spectrophotometric method such as NanoDrop.
- use the KAPA HiFi Hotstart ReadyMix and, if possible, we recommend that this is also used in the pre-submission amplification test (see above), as alternative polymerases may be less sensitive to potential inhibitors in your samples.
- no more than 94 samples are included in any submitted 96-well plate. This allows us to include a positive control for some projects as well as a negative control with each project/plate. These PCR control samples will be processed by the CGR at no additional project cost.
- we strongly recommend that you include the extraction kit controls (see experimental design paragraph) as part of your own samples, which will be included on the same sequencing run
- samples be sent to the CGR in 96-well plates rather than in tubes. Please refer to sample positions within the plate on the online order form when the samples are submitted. We further ask that the samples are ordered by column rather than row to aid automation of the workflows.
- If submit samples in tube please label them in numerical order for ease of sample identification. Please underline any numbers that could be misread upside-down (e.g. 6/9, 16/91).

For projects that involve data analyses at the CGR, we ask that information relating to sample metadata is uploaded to the online order form at the time of sample submission to increase the speed at which the analysis component of a project is completed. This information can be uploaded in the form of an Excel sheet/tab-delimited table in which the first column contains sample identifiers and subsequent columns contain different metadata for each sample.

If you are unable to meet the stated requirements, please contact us at [CGR.Enquiries@liverpool.ac.uk](mailto:CGR.Enquiries@liverpool.ac.uk) and we will be happy to offer further advice.

For more recommendations on the design of metagenetics studies, we suggest consulting the following publications:

Salter *et al.*, BMC Biology 2014 <http://www.biomedcentral.com/1741-7007/12/87>

Tighe *et al.*, Journal of Biomolecular Techniques 2017 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5345951/>

Kim *et al.*, Microbiome 2017 <https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0267-5>

Debelius *et al.*, Genome Biology 2016 <https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1086-x>

D'Amore *et al.*, Genome Biology 2016 <http://bmcbgenomics.biomedcentral.com/articles/10.1186/s12864-015-2194-9>

Pollock *et al.*, AEM 2018 <http://aem.asm.org/content/84/7/e02627-17.full.pdf+html>

Frau *et al.*, Scientific Reports 2019 <https://www.nature.com/articles/s41598-019-44974-x>