Centre for Genomic Research

UNIVERSITY OF LIVERPOOL

Submission of samples for PacBio amplicon libraries

As PacBio platforms sequence single molecules, any defects (strand breaks, abasic sites, DNA adducts, cross-linking etc.) can interfere with the library preparation and sequencing processes. The quality and quantity of the submitted DNA will determine the quality of the library, the number of SMRT cells the library can be sequenced over and the quality of the resulting sequence data. High quality DNA is crucial for optimal performance and the CGR cannot guarantee good results from samples that do not meet the requirements set out in this document.

Maximising sample quality

To maximise quality, it is essential that your amplicons:

- are double-stranded. Single-stranded DNA is not compatible with PacBio library preps.
- have not been exposed to high temperatures or extremes of pH.
- have a 260:280 ratio of 1.8-2.0 and a 260:230 ratio of 2.0-2.2.
- do not contain insoluble material.
- are free from RNA contamination.
- have been eluted and stored in a neutral, buffered solution, preferably QIAGEN EB Buffer with no EDTA. Avoid storing samples in unbuffered solutions, RNase-free water or AE Buffer.
- have not been vortexed or shaken, as this can cause shearing of the DNA.
- have not been exposed to intercalating fluorescent dyes or ultraviolet radiation. SYBR dyes do not damage DNA, but we would strongly advise against using ethidium bromide.
- do not contain denaturants (such as guanidinium salts or phenol), divalent metal cations (such as Mg^{2+}) or detergents (such as SDS or Triton-X100).
- do not contain contamination from the original organism/tissue (haeme, humic acid, polyphenols, etc.).

Clean, target-specific PCR products are extremely important for obtaining high-quality sequence data. Non-specific products can represent a substantial percentage of the sequencing reads if they are not removed. To minimize the presence of non-specific products, consider the following recommendations for generating high-quality amplicons suitable for SMRTbell library preparation and sequencing.

1. Begin with high-quality nucleic acids and work in a clean environment.
   a. If extracted nucleic acids must be stored, freeze at high concentrations in appropriately buffered solutions.
   b. To minimize possible contamination and degradation caused by multiple freeze/thaw cycles, aliquot DNA into smaller volumes for storage.
   c. Set up PCR reactions in an environment free from sources of non-specific primer and template contaminants; ideally a laminar flow hood, using dedicated pre-PCR pipettor, tips and reagents.

2. Use PCR reagents and conditions for generating target-specific, full-length amplicons.
   a. Use the highest-fidelity polymerase compatible with your PCR amplification system.
   b. Use desalted or HPLC-purified oligos; damaged bases at the ends of the amplicons cannot be repaired by DNA damage repair enzymes.
   c. Optimize PCR conditions to minimize total time spent at high (>65°C) temperatures, particularly during denaturation.
   d. PCR extension time should be long enough to ensure complete extension, taking into consideration the polymerase used and target amplicon size. For mixed samples with similar targets, it is important to complete extension at every step to avoid generating chimeric products in subsequent steps. As a general guideline, use extension times of one minute per 1000 base pairs (e.g., 3 minutes for a 3 kb product).

3. Use the lowest number of cycles required for obtaining adequate yields (ng) of PCR products to proceed with SMRTbell library construction. Avoid over-amplification.

4. If non-specific products are present, optimize PCR conditions or perform AMPure PB Bead-based size selection to enrich for PCR amplicons with the desired target size (see recommendations below).

It is highly recommended to purify amplicon samples before library preparation to remove PCR reagents, buffers, primer dimers and short non-specific PCR products. Depending on the size of the target amplicon, the concentration of AMPure PB beads required for purification varies. Refer to the table below for the appropriate concentration of AMPure PB beads to use for purification.

Amplicon size (bp)	AMPure PB beads:sample ratio
250 - 500	1.8:1
500 - 1000	1:1
1000 - 3000	0.6:1
3000 - 10000	0.45:1

For optimal sequencing data output, only amplicons of similar size (within ± 15% of the mean size) should be pooled prior to library preparation.

In-line barcodes

Please be aware that we do not recommend the use of in-line barcodes for multiplexed PacBio amplicon libraries, as they can be difficult to demultiplex and exonuclease activity during the end-repair step of the library preparation process can result in partial or total loss of the in-line barcodes. Instead, we would strongly recommend that each sample is given a unique PacBio-validated barcode, which can be included in the library preparation service we offer at the CGR.

If you do decide to use in-line barcodes, we would recommend including at least 100 nucleotides of additional sequence outside of these barcodes to prevent against exonuclease activity. Please note that the CGR does not offer demultiplexing of PacBio libraries at the in-line barcode level and we cannot guarantee the performance of such libraries.

Assessing the quality and quantity of samples prior to submission

As part of the sample submission process, we will ask you to provide quantification data for your samples. It is important that the DNA is quantified accurately – we would recommend a dye-based, dsDNA-specific method, such as Qubit. NanoDrop readings alone are not sufficient for accurate quantification but can help with assessing the quality of the sample.

Concentration measurements by Qubit and NanoDrop should not differ significantly. A significant difference in those values may indicate that the sample contains single-stranded gDNA, RNA and/or other contaminating compounds (which may not be reflected in reduced NanoDrop 260:280 and 260:230 ratios). Unfortunately, we do not have an exact acceptable difference between the readings.

Please provide a gel image or trace of all submitted samples to confirm amplification of the intended target, including the type of ladder and/or indication of fragment size(s). If there is more than one band or a smear, the sample may contain non-specific products or degraded DNA, be contaminated with RNA, or contain a contaminant that could affect the library preparation.

Sample submission requirements

Refer to the table below for DNA input requirements for library construction.

Amplicon size (bp)	Minimum input per sample
250 - 1000	20 ng per amplicon for ≥25-plex
1001 - 3000	100 ng per amplicon for ≥10-plex
3001 - 10000	200 ng per amplicon for ≥10-plex

We request that samples are submitted in volumes of 10-20 µl. We also request that samples are clearly labelled in numerical order for ease of sample identification. Please underline any numbers that could be misread upside-down (e.g. 6/9, 16/91).

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