UNIVERSITY OF LIVERPOOL

Centre for Genomic Research

Submission of DNA samples for preparation of PacBio 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, high molecular weight 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 DNA samples:

- are double-stranded. Single-stranded DNA is not compatible with PacBio library preps.
- have been stored at 4°C (short term) or -20/-80°C (long term) and have not undergone multiple freeze-thaw cycles, which can affect DNA quality.
- 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 Mg2+) or detergents (such as SDS or Triton-X100).
- do not contain contamination from the original organism/tissue (haeme, humic acid, polyphenols, etc.).

PacBio have validated the extraction methods and kits listed in the table below:

Kit	Method	Sample amount	Typical yields
QIAGEN MagAttract HMW DNA Kit	Magnetic bead	• Blood: 200 µL • Bacterial cells: 2 x 10^9 • Tissue: up to 25 mg	• Blood: 4-8 µg • Gram(-) bacteria: up to 14 µg • Gram(+) bacteria: up to 3.5 µg • Tissue: 0.5-2.8 µg per mg tissue
QIAGEN PaxGene Blood DNA Kit	Precipitation	• Blood: 8.5 mL	• 150-500 µg depending on number of nucleated cells
QIAGEN Gentra PureGene Kit	Precipitation	• Cells: up to 6.7 x 10^9 • Tissue: up to 100 mg	• 7 µg per million cells • Tissue: 5-100 µg
QIAGEN Genomic Tip 20/G Kit	Anion-exchange column	• Blood: 1 mL • Cultured cells: 5 x 10^6 • Yeast cells: 1.5 x 10^9 • Bacterial cells: 4.5 x 10^9 • Tissue: up to 20 mg	• 1-20 µg
Circulomics Nanobind CBB Kit	Nanobind disc	• Blood: 200 µL • Cells: 1 x 10^6 • Bacterial cells: 5 x 10^9	• 5-34 µg depending on sample type and input amount
Circulomics Nanobind Tissue Big DNA Kit	Nanobind disc	• Tissue: ~25 mg	• 5-100 µg
Circulomics Nanobind Plant Nuclei Big DNA Kit	Nanobind disc	• Plant nuclei: up to 10 g	• 5-20 µg
Lucigen MasterPure Kit	Precipitation	• Cells: 1 x 10^6 • Blood: 200 mL • Bacterial cells: 3.5 x 10^6	• Cells: 3-12 µg • Blood: 3-9 µg • Bacteria: 1.3-1.6 µg
NEB Monarch Genomic DNA Purification Kit	Anion-exchange column	• Blood: 100 µL • Bacterial cells: 2 x 10^9 • Cells: 5 x 10^6 • Tissue: 10 mg	• Blood: 2.5-4 µg • Gram(-) bacteria: 6-10 µg • Gram(+) bacteria: 6-9 µg • Mammalian cells: 7-9 µg • Tissue: 5-30 µg

Kit	Method	Sample amount	Typical yields
Macherey-Nagel NucleoBond HMW DNA Kit	Anion-exchange column	• Blood: 2 mL • Plant leaves: 1.5 g • Bacteria: up to 100 mg • Cells: $1x10^7$ • Animal tissue: up to 300 mg	Sample dependent
QIAGEN DNeasy PowerMax Soil Kit	Anion-exchange column	• Soil: up to 10 g	Sample dependent

If using any other kit, we advise you to consult the manufacturer's technical support team to find out if they have recommendations for the organism you are working with, specifically with regard to long-read sequencing.
Please see www.extractdnaforpacbio.com for "homebrew" protocols for DNA extraction from a range of different organisms. During the extraction process, we recommend the inactivation of nucleases and other DNA binding proteins with proteinase K, as well as removal of RNA with RNase A.

General recommendations to follow before DNA extraction to help maintain high molecular weight and clean DNA:

- avoid incubation in complex or rich media.
- harvest from several cultures rather than a single, high-density culture.
- harvest during early- to mid-logarithmic growth phase.
- extract small volumes rather than large volumes to avoid accumulation of inhibitors.

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.

If you are trying to quantify HMW DNA that is viscous, typically >100 kb in size on average, then we recommend diluting a small aliquot and pipette up and down 10-20 times to shear and better homogenize the DNA in solution. For even more accurate quantification, consider making 3 aliquots and take the mean of all three readings. The CV between the three readings should be less than 30%. If not, try additional mixing of the samples to homogenize the DNA, and repeat quantification.

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.

In order to obtain the true size of the gDNA, samples should be assessed on one of the following instruments:

- CHEF Mapper XA Pulsed Field Electrophoresis System (Bio-Rad)
- Femto Pulse (Agilent)
- Pippin Pulse (Sage Science)

If this is not possible, please run the gDNA on a 0.5% agarose gel overnight at 30 - 35 V for 17 - 18 hours. The ladder on the gel should have a marker of >40 Kb (we recommend the GeneRuler High Range DNA Ladder from Thermo or the 1 Kb DNA Extension Ladder from Life Technologies). Please provide a gel image or trace of all submitted samples to confirm sample integrity, 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 degraded gDNA, be contaminated with RNA, or contain a contaminant that could affect the library preparation.

Sample submission requirements

The concentrations required will depend on the type of libraries being generated. The table below outlines the minimum input for each type of library, but higher amounts of DNA will generally lead to higher quality libraries and will enable us to sequence over more SMRT cells.

Library type	Minimum DNA per sample	Additional information
Ultra-low input	20 - 50 ng	Not recommended for genomes >500 Mb.
Low input	1.5 µg	Not recommended for genomes >1 Gb.
Multiplexed microbial	2 µg	Maximum 96 samples per SMRT cell. See Best Practices guide.
Structural variant detection	10 µg	N/A
HiFi sequencing	5 µg	5 µg is the minimum input requirement for this library prep, but we require 5 µg DNA per 1 Gb genome for larger genomes.

We request that samples are submitted in volumes of 100-200 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).

When shipping, we recommend DNA and SMRTbell libraries be shipped frozen on dry ice. Keeping the DNA, or library, frozen helps prevent shearing which can occur from the jostling experienced during shipping.

If you are unable to meet the stated requirements for your library type, please contact us at CGR_Lab@liverpool.ac.uk and we will be happy to offer further advice.