SERVICE GUIDE

AGRF-UQ PacBio Service

PacBio-UQ@agrf.org.au
1300 247 301
www.agrf.org.au

SVGPCB2206

Service Guide:
AGRF-UQ PacBio Service

Service Guide:
AGRF-UQ PacBio Service

1.0 About the Service

Third generation sequencing technologies build upon the advances made in short-read, second generation sequencing and introduce long-read sequencing as the new frontier in molecular biology. One such technology is the Pacific Biosciences (PacBio) Sequel II long-read sequencing platform which leverages fluorescently labelled nucleotides (Sanger, sequencing-by-synthesis) within nanoscopic wells called Zero-Mode Waveguides (ZMW), and a DNA Polymerase to produce high-quality long-read sequencing data (Figure 1). This technology is called Single-Molecule, Real-Time (SMRT) sequencing and produces High-Fidelity (HiFi), long-read data with simultaneous epigenetics (in specific applications).

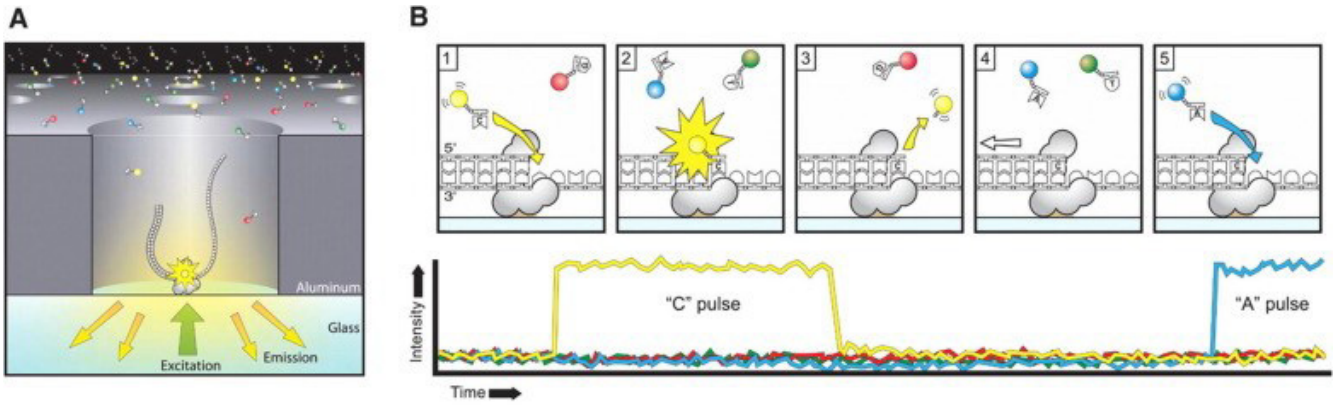


Figure 1: Visual representation of SMRT sequencing in action (Rhoads, 2015). A) Template material is bound within the ZMW with a DNA polymerase incorporating a fluorescently labelled nucleotide complimentary to the template material. B) As the polymerase moves along the template material, the fluorophore fluoresces as it's incorporated, and the emission is detected using a nano-sized camera with each colour representing a different nucleotide.

However, prior to sequencing, the target material must undergo library preparation which is the process of preparing your samples for sequencing. This process typically involves fragmentation and attaching of oligo-nucleotide adapters to facilitate sequencing.

PacBio libraries require circularisation which is the process of attaching SMRTbells to the template, allowing the polymerase to repeatedly sequence the same template for improved accuracy (Figure 2). This allows the generation of circular consensus (CCS) reads which are up to Q50 in quality.

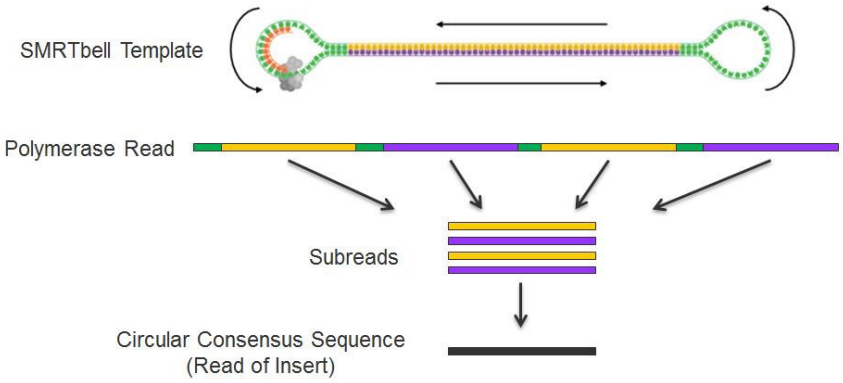


Figure 2: Graphical representation of the SMRTBell adapters as part of the template. This allows the polymerase to perform multiple passes of the same template molecule to generate high quality long read data. The multiple passes are polymerase reads and then the consensus of the polymerase reads are the CCS reads.

Service Guide:
AGRF-UQ PacBio Service

The AGRF-UQ PacBio Service offers the following applications:

1. Iso-Seq (RNA-seq): Full-length transcript sequencing of transcripts up to 10Kbp in length.
    a. Discovery of new genes
    b. Improved gene annotations
    c. Improved accuracy to detect isoforms
    d. dentify genetic structures, regulatory elements and coding regions
2. HiFi (Whole-genome Sequencing): Whole-genome sequencing (WGS) generating long-reads >10Kbp in length
    a. High-quality reference genome assembly
    c. Sequence and characterise entire genes within a single read
    d. Improved structural variant detection

2.0 Iso-Seq

Iso-Seq libraries are generated from Total RNA and follow typical RNA-Seq steps except with the addition of SMRTbell library construction and size selection is not required (Figure 3). Iso-Seq libraries are initially prepared using NEBNext Low Input cDNA Synthesis & Amplification module (Figure 4) followed by SMRTBell Express 2.0 library construction (adapter ligation of SMRTbell). RNA submission requirements are listed in Table 1.

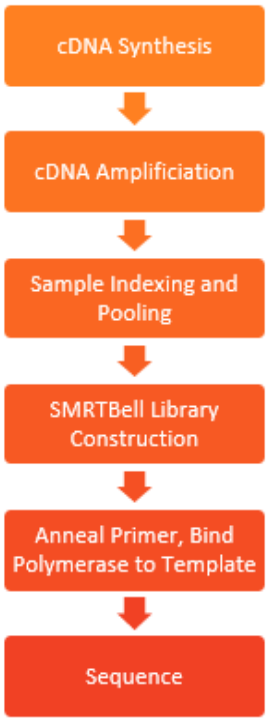
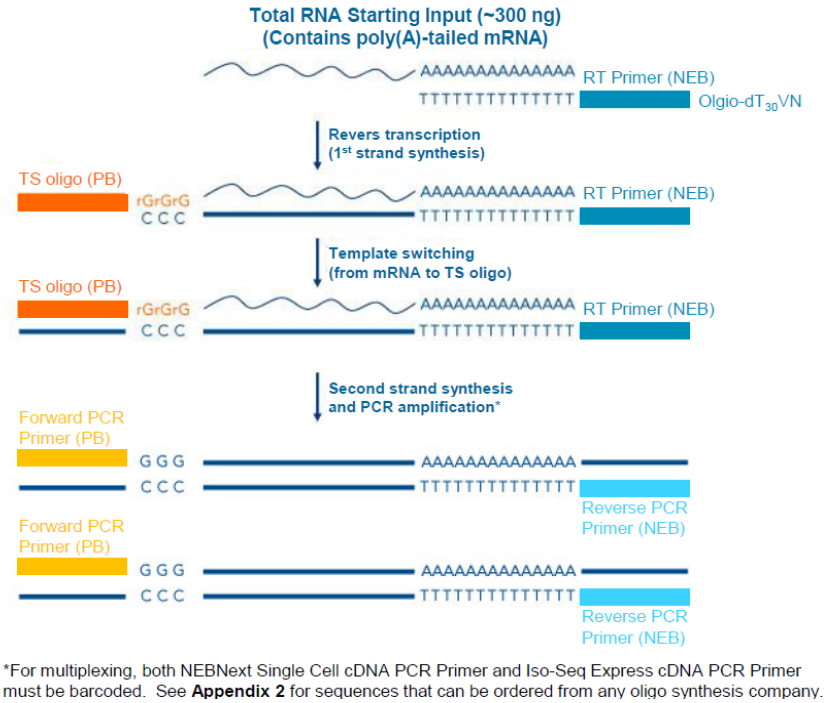


Figure 3: Iso-Seq library preparation workflow summary.



*For multiplexing, both NEBNext Single Cell cDNA PCR Primer and Iso-Seq Express cDNA PCR Primer must be barcoded. See **Appendix 2** for sequences that can be ordered from any oligo synthesis company.

Figure 4: Visualisation of the cDNA synthesis process using Template Switching (TS) and Reverse Transcription (RT) oligos (PacBio, 2020)

3.0 HiFi (Whole Genome Sequencing)

Samples submitted for HiFi library construction are prepared from gDNA and undergo mechanical shearing followed by SMRTbell Express 2.0 library construction, BluePippin size selection which selects for fragments between 10-20Kbp, and then sequencing (Figure 5). DNA sample requirements for both workflows are listed in Table 1.

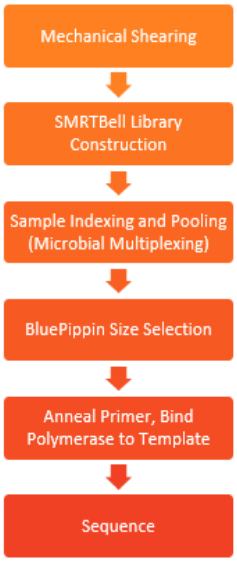


Figure 5: HiFi and Microbial Multiplexing workflow summary

4.0 Sequencing Recommendations and Sample Requirements

Sequencing and input requirements for all PacBio services offered can vary considerably depending on the type of sample and the intended analysis type.

	Iso-Seq	HiFi
Input	>300ng	> 20µg
Multiplexing per SMRT Cell	12 (Human)	2 (Human) or up to 48 (Microbial)
Target Insert Size	≤2Kbp or ≥3Kbp	10-15Kbp
Fragmentation	N/A	Mechanical
Sample Type	Total RNA	Complex Genomes (gDNA)

Table 1: Summary of input requirements for AGRF offered PacBio library preparation kits.

Sample Types	Application/Analysis	Coverage
Small Genomes <200Mb (32-48 genomes per SMRT Cell)	resequencing - SNV, MLSTs de novo genome assembly† m6A and m4C detection	20-100x
Complex Genomes >200Mb (Genomes up to 2Gb per SMRT cell)	resequencing - SNVs resequencing - Indels resequencing - CNVs hypo and hyper-methylated CpG de novo genome assembly†	>15x
Iso-Seq (12 transcriptomes for gene annotations or 1 human transcriptome for deep profiling per SMRT Cell)	Discovery of new genes Improve gene annotations	200K-300K reads
	Improved accuracy to detect isoforms Identify genetic structures, regulatory elements and coding regions	Up to 4M reads

Table 2: General recommendations for sequence depth per SMRT Cell

The standard input for any genome sequencing project is high quality genomic DNA (DNA Integrity Number - DIN ≥ 8) or for any transcriptome sequencing project is high quality RNA (RNA Integrity Number - RIN ≥ 7). Recommendations for isolation of DNA and RNA for PacBio projects are listed below:

DNA (HiFi)

Column-based extraction protocols or equivalent, specific for high-molecular weight recovery (example kits below):

- Qiagen Genomic-Tip 20/100/500/G kit
- Circulomics Nanobind CBB/Plant kit
- GeneJET Plant Genomic DNA Purification kit
- Lucigen Masterpure Kit

- Please ensure the extraction protocol includes an RNase treatment to ensure removal of RNA.
- Purified DNA should be eluted/resuspended in 10 mM Tris-Cl (pH 8.5 – also known as buffer EB), or nuclease-free water. Do not use EDTA-containing buffers as this can inhibit PCR reactions.

For best results we recommend following the PacBio sample and extraction guides below:
1. Technical Note - Preparing DNA for PacBio HiFi Sequencing - Extraction and Quality Control
2. Technical Note - Preparing Samples for PacBio Whole Genome Sequencing for de novo Assembly - Collection and Storage

Organism specific protocols recommended by PacBio can also be found here.

RNA (Iso-Seq)
Column-based extraction protocols, specific for high-molecular weight recovery (example kits below):
- Qiagen RNeasy Plus kit
- Ambion Poly(A) PuristTM MAG kit
- Sigma Aldrich Spectrum Plant Total RNA kit
- iNtRON Easy Spin Total RNA
- RNALater

- Please ensure the extraction protocol includes a DNase treatment to ensure removal of DNA.
- Purified RNA should be eluted/resuspended in nuclease-free water or RNALater and shipped using dry ice.
- Purity of nucleic acid samples can be assessed by measuring the absorbance spectra on a spectrophotometer (e.g. Nanodrop). The ratio of absorbance values 260nm and 280nm or 230nm provides an estimate of sample purity or the presence of common contaminants (Table 3).

Ratio	Target (RNA)	Target (DNA)	Low Ratio (<1.6) indications
A260/280	2.0	1.8-2.0	Residual phenol from extraction, proteins or very low conc. of nucleic acids (<1ng/µl)
A260/A230	≥2.0		Residual guanidine from the extraction protocol Carryover of carbohydrates (e.g. plant polysaccharides)

Table 3: Recommendation for sample purity as assessed by absorbance spectra.

Sample quality is a key factor for successful PacBio experiments. AGRF will perform quality control (QC) prior to commencing a project, however we recommend you check your material before submission.

DNA Integrity can be assessed by agarose gel electrophoresis or microfluidic assay such as an Agilent gDNA TapeStation assay, Agilent FemtoPulse, Sage Science PippinPulse or PerkinElmer GX systems (Figure 6). DNA should appear as a clear, high molecular weight band. There should be no indication of RNA contamination, as evident by faint bands or smears toward the bottom of the gel. Quantification of gDNA by dsDNA assay using fluorescence such as either Quantifluor or PicoGreen or Qubit is highly recommended (Table 4).

RNA Integrity can be assessed by electrophoresis using systems such as the Agilent Bioanalyzer or PerkinElmer LabChip GX. These systems provide a measure of the “intactness” of RNA, based on the profile of the sample electropherogram, reporting a RIN or RNA Quality Score (RQS). Samples with a RIN (or RQS) ≥7 are considered high quality (Figure 7). Quantification of RNA using an RNA fluorescence assay such as either Quantifluor or RiboGreen or Qubit is highly recommended (Table 4).

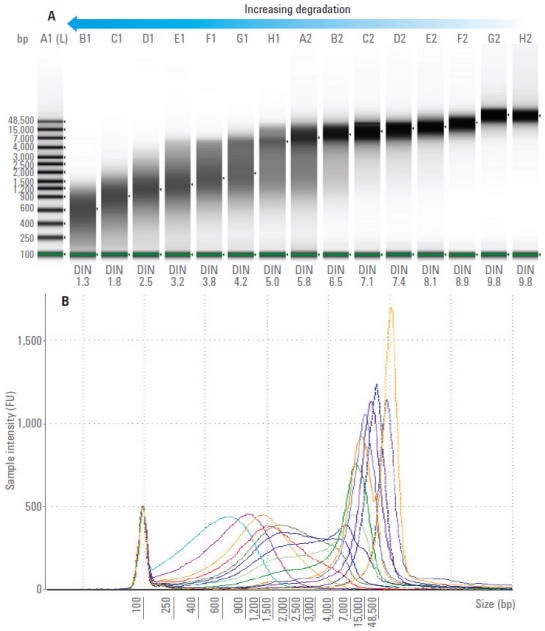


Figure 6: Example of DIN assessment using an Agilent 2200 gDNA TapeStation assay (Agilent, 2015).

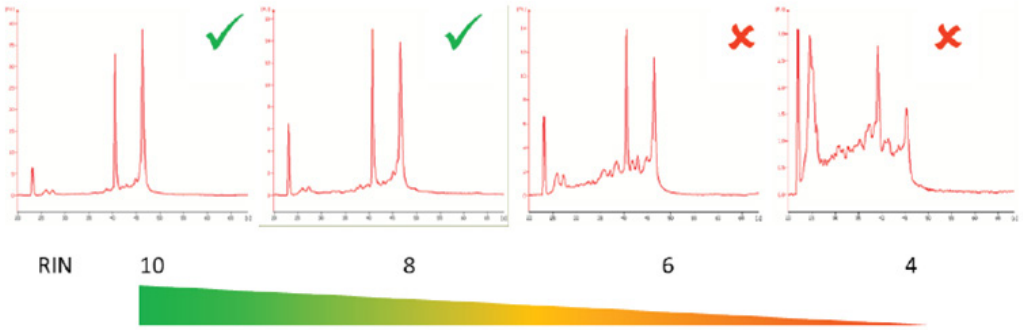


Figure 7: Example of RIN assessment using an Agilent Bioanalyzer 2100.

Library Type	Sample Type	Quantity	DNA Concentration
HiFi	gDNA (≥8 DIN) >20-40Kbp peak size	≥20µg	≥150ng/µl
Multiplexed Microbial	gDNA (≥8 DIN) >20-40Kbp peak size	≥2µg	≥150ng/µl
Iso-Seq	Total RNA (≥7 RIN)	≥300ng	≥20ng/µl

Table 4: Sample Requirements of DNA library preparation services.

5.0 Data Output

All PacBio sequencing projects will undergo quality control (QC) analysis to assess the quality of both the sequencing and library preparation. AGRF will provide CCS reads in FASTQ format.

AGRF also provides options for sequencing analysis including genome alignment, transcriptome assembly and mapping and de novo assembly. If you have specific analysis requirements, please discuss these during the quotation process.

6.0 Sample Submission

6.1 Online Submission

- In the client portal, select 'Next Generation Sequencing' from the service dropdown menu.
- Enter your species and submission format (tube or plate).
- Complete and upload the template file.
- Submit the form and print the submission receipt to be included with your sample package.

6.2 Packaging of Samples

- DNA samples can be sent at ambient temperature or on ice blocks.
- If you are sending your samples in plates, please use strip caps to seal the plates.

Sample tubes or plates should be in a zip-lock bag or box to avoid direct contact with dry ice. Post/send/deliver samples to the addresses below:

Postal Address (mail)
AGRF Ltd
Gehrmann Laboratories
Research Rd
University of Queensland
Brisbane QLD 4072

Physical address (courier)
AGRF Ltd
Level 5, Gehrmann Laboratories
Research Rd
University of Queensland
Brisbane QLD 4072

6.3 Return of Samples

If you require your samples to be returned to you post-processing, please let your Account Manager know at the time of quoting. Please note that a fee will be charged for return of samples.

7.0 Quality Statement

All works carried out by AGRF are performed following the strict requirements of ISO17025: 2005. AGRF Ltd is accredited in the field of Biological Testing (Scope: DNA Analysis) according to the ISO17025: 2005 standard by the National Association of Testing Authorities (NATA). Staff follow Standard Operating Procedures, which define their responsibilities and provide guidance on achieving standards; compliance is monitored at regular reviews and internal audits. All work is supervised by a person with relevant qualifications and is checked while in progress and upon completion to ensure that it met the necessary ISO17025: 2005 standards.
The AGRF-UQ PacBio Service is a Certified Service Provider (CSP) for PacBio sequencing. CSP programs are collaborative service partnerships between facilities and vendors to ensure best practice and the highest data quality.

8.0 References

1. Agilent. (2015). DNA Integrity Number (DIN) with the Agilent 2200 TapeStation System and the Agilent Genomic DNA ScreenTape. DNA Integrity Number (DIN) with the Agilent 2200 TapeStation System and the Agilent Genomic DNA ScreenTape. Agilent.
2. PacBio. (2020). Iso-Seq Express Library Preparation Using SMRTbell Express Template Prep Kit 2.0. Iso-Seq Express Library Preparation Using SMRTbell Express Template Prep Kit 2.0. Pacific Biosciences.
3. Rhoads, A. a. (2015). PacBio sequencing and its applications. Genomics, proteomics & bioinformatics 13.5 (2015): 278-289.