# Illumina Sample Submission Guidelines

## Introduction

The BRF defines a "sample" as nucleic acids (or material containing nucleid acids) that have not been prepared with Illumina index adapters. Once the sample has been prepared with Illumina index adapters, it is referred to as a "library". Two or more uniquely indexed libraries combined into a single tube is referred to as a "library pool".

Next-generation sequencing (NGS) is a massively parallel sequencing technology that offers ultra-high throughput, scalability, and speed. The technology is used to determine the order of nucleotides in entire genomes or targeted regions of DNA or RNA. It leverages sequencing by synthesis (SBS) technology – tracking the addition of labelled nucleotides as the DNA chain is copied – in a massively parallel fashion.

Some examples of NGS applications include:

- whole genome sequencing (WGS)
- exome sequencing (deeply sequence targeted regions)
- RNA sequencing (RNA-seq) to discover novel RNA variants and splice sites, or quantify mRNAs for gene expression analysis
- epigenetics such as genome-wide DNA methylation and DNA-protein interactions
- rare somatic variants, tumour subclones etc. in cancer samples
- metagenomics
- identifying novel pathogens
- spatial genome organisation and genome contact maps to aid genome assembly (HiC sequencing)

## Illumina Sequencing Instruments at the BRF

The BRF has a range of Illumina Sequencers to accommodate a wide range of sequencing data output from 0.3 Gb to 8,000 Gb, depending on the project requirements. The BRF Illumina sequencer fleet includes a NovaSeq X Plus, a NovaSeq 6000, a NextSeq 2000 and a MiSeq. The NovaSeq 6000 and NovaSeq X Plus feature dual flow cell configuration, meaning two flow cells can run simultaneously.

Sample Submission Procedure for Illumina Sequencing at the BRF

The BRF Illumina sequencing services offers:

- library preparation services followed by sequencing
- sequencing of libraries that have been prepared by the customer

The following is a step-by-step guide on how to submit samples and libraries for sequencing.

## Sample Submission for BRF Library Preparation and Sequencing

1. Complete and sign an order form, which can be found on the BRF Illumina Sequencing website. There is a form for submitting samples for library preparation and sequencing, and a form for library submission and sequencing. Please use the "BRF Library Preparation & Sequencing" order form, and email a completed copy to brf@anu.edu.au (**do not print the form**).
2. Prepare QC documents, which includes fragment analysis reports and a spreadsheet listing sample numbers, sample IDs, concentrations, volumes and quality metrics. These reports and spreadsheets must be emailed to brf@anu.edu.au with the order form.
3. Please use a simple numbering convention to label samples (e.g. 1, 2, 3 etc.). Samples can have unique names as long as they are listed in the QC spreadsheet with their corresponding sample number. If submitting ≤ 24 samples, please use 1.5 mL tubes with the sample number written clearly and neatly on the lid. If submitting > 24 samples, please use 96-well plates with the samples plated column-wise. The QC spreadsheet should include sample well position in the plate.
4. After the BRF has assessed the documentation and accepted the sample submission, please bring the samples either in person or use a courier to deliver them to the BRF. It is recommended to send them on Monday to Wednesday to ensure they are received before a weekend.

## Library Submission for Sequencing

1. Complete and sign an order form, which can be found on the BRF Illumina Sequencing website. There is a form for submitting samples for library preparation and sequencing, and a form for library submission and sequencing. Please use the "Sequencing-Ready Libraries" order form, and email a completed copy to brf@anu.edu.au (**do not print the form**).
2. Prepare QC documents, which includes fragment analysis reports and a spreadsheet listing library numbers, library IDs, concentrations, volumes and quality metrics. Please list the Index 1 and Index 2 sequences for each library in this spreadsheet. These reports and spreadsheets must be emailed to brf@anu.edu.au with the order form.
3. After the BRF has assessed the documentation and accepted the library submission, please bring the libraries either in person or use a courier to deliver them to the BRF on ice or cold packs. It is recommended to send them on Monday to Wednesday to ensure they are received before a weekend. Each Illumina sequencer requires different amounts of DNA for sequencing, and these are detailed in the "Sequencing-Ready Libraries" order form.

## Data Management & Delivery Options

Illumina sequencers generate demulitplexed sequencing data in .fastq.gz format. The size of these files can range from a few hundred megabytes to a few hundred gigabytes or even terabytes. Large amounts of data can be challenging to share and store, so the BRF highly advises that the customer organises a long-term storage solution prior to receiving the sequencing data.

There are multiple options for data sharing:

1. **BaseSpace Sequence Hub:** is the preferred data sharing option for large sequencing runs. Users are able to create a free Illumina BaseSpace account that can store up to 1 TB of data. The BRF can upload the data to BaseSpace and share a link to the user to access it.
2. **NCI**: the BRF uses NCI to collect NGS data from the sequencers. For ANU users that are a member of an NCI project, temporary access to the user's project can be granted to the BRF to transfer the sequencing data from the BRF's project to the user's project. This is recommended for very large sequencing runs that will be analysed using NCI compute power.
3. **AARNet FileSender**: is a free online file transfer service. After the data is uploaded, the customer will receive a link to access and download the data to a location of their choosing.
4. **Globus**: is an online data sharing portal. A Globus account needs to be created on a University level (not on an individual user level), and the BRF can share a secure link to the sequencing data.
5. **Hard drive**: a hard drive can be supplied or purchased from the BRF. This method of data delivery should only serve as a last resort, and keeping data on a hard drive as a long-term solution is not advised.

# Sample Requirements

## DNA Library Preparation

For **each sample**, the requirements are:

| | |
|---|---|
| **gDNA Prep** | for large genomes, 100-500 ng gDNA. For small genomes, at least 1-10 ng gDNA |
| **Plasmid Prep** | 100-500 ng |
| **DNA Amplicon Prep** | 20 ng DNA amplicons (or fragmented DNA) |
| **Blood Prep** | 12 µL of fresh, nucleated whole blood in EDTA/heparin tubes (anticoagulated) |
| **Methyl-seq Prep** | 10-200 ng gDNA |

| | |
|---|---|
| **DNA NanoDrop Absorbance Ratio** | 260/280 = 1.8-2.0 \| 260/230 = 2.0-2.2 |
| **DNA Storage Solution** | 10 mM Tris-HCl, pH 7.5-8.5, 1 mM EDTA |
| **DNA Sample Volume** | no more than 25 µL |

DNA samples should be assessed on the NanoDrop One for purity and quantified on a fluorometric assay such as the Qubit High Sensitivity or Broad Range dsDNA assay. The gDNA size profile should be assessed on a gel or Agilent Femto Pulse. Short DNA samples should be assessed on the Agilent Bioanalyzer, Agilent TapeStation, LabChip GX or similar. These quantification results must be submitted in an Excel spreadsheet or PDF report.

It is recommended to submit more than the minimum amount of required material for circumstances that require repeat preparation, troubleshooting or optimisation. 10 mM Tris-HCl, pH 7.5-8.5 and $H_2O$ are also acceptable DNA storage solutions, however, $H_2O$ is not an advisable solution for long-term storage. Small DNA short be stored at -20°C; gDNA should be stored at 4°C.

## RNA Library Preparation

For **each sample**, the requirements are:

| | |
|---|---|
| **mRNA Prep** | 10-1000 ng of total RNA in no more than 25 µL |
| **RNA Amplicon Prep** | 80 ng RNA amplicons in no more than 8.5 µL |
| **Small RNA Prep** | 0.1-10 ng in no more than 8 µL |

| | |
|---|---|
| **RIN Score** | ≥ 7 for total RNA |
| **RNA Storage Solution** | Ultrapure Nuclease-Free Water. Must not contain guanidinium salts, EDTA, phenol, ethanol or DNA (samples can be treated with DNase I and be subsequently removed with phenol/choloroform extraction and ethanol precipitation). |

RNA purity can be assessed on the NanoDrop One. Quantification and quality can be assessed on the Agilent Bioanalyzer, Agilent TapeStation, LabChip GX, or similar. A PDF report for this QC must be sent to the BRF when submitting samples, along with a spreadsheet listing sample numbers, IDs, concentrations, RIN scores and NanoDrop results.

RIN scores are calculated from multiple metrics. If some of those metrics are undetermined during the fragment analysis, a RIN score of N/A may be displayed. This does not mean that the sample cannot proceed to library preparation - if there is sufficient material of reasonable quality (and if the client expresses their desire to attempt the library preparation anyway) then it can be processed. It is important to remove DNase from the samples after treatment as it can cause degradation of adapters and primers used in the library preparation. RNA samples should be stored at -80°C, and delivered to the BRF on dry ice.

For researchers interrogating all types of RNA in their samples, please contact the BRF.

## HiC Library Preparation

For tissue types other than those listed below, please contact the BRF. For **each sample**, the requirements are:

| | |
|---|---|
| **Cell Culture** | pellet of 1 million cells |
| **Plant** | 5 g of young seedling leaves |
| **Nucleated Blood** | 100 µL of non-coagulated whole nucleated blood in EDTA-coated tubes, or whole nucleated blood preserved in ethanol (containing 100 µL of original blood volume) |
| **Animal Tissue** | ≥ 0.5 $cm^2$, 50-200 mg (the approximate size of a lentil). The preferred order of specimens are: heart, liver, muscle. |
| **Insects** | whole intact insects. For small insects, please provide multiple specimens to ensure enough DNA can be extracted |
| **Fungi** | 50-200 mg of spores or powder. For filamentous fungi, 100-1000 mg |

| | |
|---|---|
| **Sample Handling** | Samples must be freshly caught/collected/dissected and snap frozen for preservation. Please store samples at -80°C and deliver/ship to the BRF on dry ice. Fresh, non-coagulated blood can be stored at 4°C. |

## ATAC-seq Library Preparation

For **each sample**, the requirements are:

| | |
|---|---|
| **Cells** | 50,000-100,000 cells in 1 mL 1X PBS |
| **Tissue** | 20-30 mg (fresh or flash frozen) |

| | |
|---|---|
| **Cell Handling** | fresh or cryopreserved frozen cells can be used. Cells must be of high viability. Frozen cells should be frozen at a controlled rate with media formulated against ice crystal formation. |
| **Tissue** | fresh or flash-frozen tissue at -80°C can be used |

## 10x Genomics Single-Cell RNA-seq

Single cell sample submission needs to be coordinated with the BRF. Please actively inform the BRF on the progression of cell sample preparation and whether you are able to meet your appointment time for submission. We require the cells to be submitted at the earliest possible time on an agreed day. If circumstances change and you are not able to submit by this time or day, please inform the BRF.

The BRF offers 10x Genomics Chromium Single-Cell RNA-seq library preparation and sequencing. The BRF can generate 3' gene expression libraries with feature barcoding; 5' gene expression libaries with feature barcoding and TCR/BCR (mouse and human) immune profiling; and fixed RNA profiling gene expression libraries with feature barcoding. Cell samples should be free from aggregates or debris, FACS-sorted and manually counted under a microscope for best quantification results. For **each sample**, the requirements are:

| | |
|---|---|
| **3' scRNA-seq** | 10,000 - 29,000 cells |
| **5' scRNA-seq** | 10,000 - 29,000 cells |
| **Fixed RNA profiling** | At least 300,000 cells/nuclei |

The 10x Genomics webpage contains numerous demonstrated protocols for cell preparation, tissue dissociation, nuclei isolation and cell hashing for feature barcoding. Please consult the documents relevant to your sample type for more information on how to prepare your samples.

Please aim to supply your final cell samples at a concentration of ~1000 cells/µL. The recovery rate after cell barcoding is ~60 - 80% of the cell input. If you are looking to capture 10,000 cells, aim to supply at least twice this amound of cells. After FACS sorting your samples, use manual cell counting under a microscope for the best quantification results. View the samples under an automated cell counter to ensure there are no clumps or debris present as these can clog the microfluidics of the cell barcoding chip. Cells should have a high viability as ambient RNA (leaking from dead cells) will be captured in the cell barcoding reaction and take sequencing reads away from your cells of interest.

If you are using BioLegend hashtag antibodies for feature barcoding and cell sample multiplexing, please check if your antibodies are compatible with 10x Genomics protocols.

After sequencing, the BRF will run the cellranger analysis pipeline on the dataset, and provide the raw .fastq files and the output files of the pipeline.