Technical document for library preparation and sequencing

The general process of sequencing service at BMKGENE consists of six parts, as illustrated below in Chart 1: sample shipment, sample quality assessment, library preparation, sequencing, sequencing data quality control, and data release. During this process, the client would need to provide permission twice to proceed with the project, respectively after the sample quality assessment and after the data quality control. Only if the confirmation of proceeding is received at BMKGENE, will the project move forward to the next stage. A fast-track process without clients' confirmation could be issued as well if it is requested by the client.

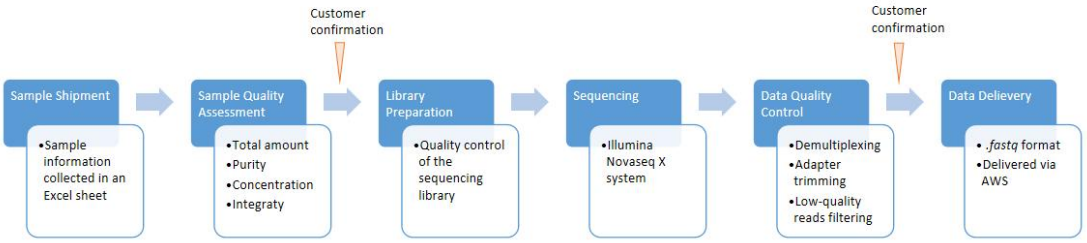

Chart 1. General process of sequencing service at BMKGENE

The detailed information on reagents and methods used in sample quality assessment, library preparation,and sequencing for the ten required sequencing services is described in the rest of part of this document.

Please be informed that, if more flexibility and variability are required in the project, we highly encourage clients to discuss with us before declining the possibility of proceeding. BMKGENE has developed a grading system for handling low-quality samples based on the sequencing experience in the past fifteen years. This system gathers successful and failed experiences in samples with low quality and low concentration, samples with severe pollution, and samples with degraded nucleic acid molecules. With the ongoing gaining of experience, we constantly adjust our sample requirements, aiming to widen the sequencing possibilities for challenging samples and projects.

3. Pre-made library sequencing with Illumina PE150 mode

1) Sample quality assessment

a) Sample quality control

Upon receiving the library samples, sample quality will be assessed by using the listed-below systems.

Measurement	System	Kit
Concentration	Qubit 4.0	Qubit TM dsDNA HS Assay Kit
Molecule integrity, Concentration, Fragment size	LabChip GX Touch HT Nucleic Acid Analyser	HT DNA HS reagent kit (Revvity)
qPCR	-	effective concentration (molarity)

- Our sequencing procedure is compatible with libraries prepared by using kits from the 10x Genomics. If libraries are produced with other kits or brands, please consult the local account manager for library compatibility check.

b) Sample requirements

For ready-made library sequencing, the library requirement is shown below.

- Here, we assume the sequencing library is produced by using kits from 10x Genomics.
- For other libraries, please follow the official recommendation from the kit company.

Measurement	Requirement
qPCR concentration	≥ 2 nM
Volume	≥ 20 μl
Peak range	300-800 bp (scRNA)
Average Fragment Size	420-580 bp (scRNA), 170-1000 bp (scATAC). Other library following the official manual.

c) Sample quality reports will be sent to our customers via email for their ***confirmation***.

2) Sequencing

- If client choose to purchase by lane or by flow cell, then certain percentage of Phix library will be pooled together with the library samples for sequencing. Sequencing will be performed on the Illumina NovaSeq X system with the PE150 mode.
  - For 10x Genomics snRNA, scRNA libraries, and multiome libraries, adding 1% of phix per lane is recommended by the 10X Genomic protocol.
- If client choose to purchase by data volume, qualified libraries will be pooled with other samples and sequenced in a 25B flow cell on the Illumina NovaSeq X system with the PE150 mode.
- If special lengths of read 1 and read 2 for different types of libraries are required, BMKGENE can trim the reads to the desired length free of charge.

3) Data quality control

For all the samples matching our sample quality standards, BMKGENE guarantees high-quality sequencing data to meet the agreed data volume in the contract. Here, the high quality refers to the raw data in which more than 85% of the bases are above Q30.

The resulting raw reads will go through quality control before sharing with the client, and the detailed process is listed below.

a) Raw sequencing data quality control.
   i. Assessing total data volume, read length, and read numbers.
   ii. Counting the percentage of reads with a Phred quality score above Q20 and Q30.
   iii. Adapter trimming and demultiplexing.

4) Data delivery and storage

a) Once data quality control is completed, data quality report will be sent to our customers.

b) Upon receiving the report, our clients would need to ***confirm*** whether they are satisfied with and will accept the sequencing data.

c) Once the confirmation is received by BMKGENE, the project is considered completed, and the data release process starts.

d) By default, the sequencing data will be delivered by AWS (Amazon Web Services) in *.fastq* format. It is also flexible to choose data delivery by hard drive or other methods.

e) Md5 files for each batch of data and reports will be provided together in data release.

f) Sequencing data will be stored free of charge for six months upon receiving the project completion confirmation at BMKGENE. **This length of data storage period should be stated in the final contract.** Sequencing data will be automatically deleted before the first week of the 7th month if no further notice is received .

g) The AWS link for data downloading will be available for one month after project completion, due to technical restriction. However, upon request the data will be uploaded again to the AWS after the first link expires.

h) The RNA samples will be stored for four months after project completion. **This length of sample storage period should be stated in the final contract.**

i) The data delivery or turnaround time is shown as below, which starts to count when samples are received at BMKGENE. If a shorter TAT is required, please consult with the local sales manager.

Purchase mode	Working days
By Gb	30
By lanes	25