GenomEast Platform
CERBM GIE – IGBMC
1 rue Laurent Fries
67404 Illkirch Cedex
France

+33 3 88 65 34 26
thibault@igbmc.fr
jost@igbmc.fr
genomeast.igbmc.fr

PRODUCT SHEET: SMALL RNA-SEQ

Small RNA-seq allows expression analysis and discovery of new miRNA and other small non-coding RNA, using high-throughput sequencing technology.

1 Experimental design

1.1 Biological replicates

It is very important to include replicates in your experimental design (cf. Hansen et al., Nature Biotechnology 29:572-573, 2011). A randomized and balanced experimental design is also important. We also encourage project managers to try to reduce batch effects during sample preparation. Project managers who need advices to define the most appropriate experimental design according to their biological questions are encouraged to contact us before starting their experiments.

1.2 Multiplexing

For a small RNA-seq project, we usually recommend to multiplex up to 20 samples per lane on an Illumina HiSeq 4000 sequencer.

2 Services provided

1. Sample checking:
   - Quantity and quality check using a fluorometer (Qubit or Varioskan) and a capillary electrophoresis machine (Bioanalyzer, Agilent).
2. Library preparation:
   - Preparation of fragmented cDNA libraries and ligation of indexed sequencing adapters to DNA fragments. Indexes are DNA sequences of ≥ 6 nt long used to identify each sample. Usage of indexes allows for pooling multiple samples on a single sequencing lane.
   - Libraries quantification and quality control by capillary electrophoresis (Bioanalyzer from Agilent or Fragment Analyzer from AATI).
3. Sequencing using Illumina HiSeq 4000 technology:
   - Loading of libraries on a flow cell and generation of clusters on the Cbot (Illumina).
   - Single-read sequencing 1x50 bases.
4. Primary data analysis:
   - Demultiplexing and generation of FASTQ files.
   - Adapter dimer removal.
   - Sequencing quality check.
   - Detection of potential contaminations.
   - Generation of a report summarizing the methods used in the pipeline as well as the results obtained (one report per sample and one global report for the project).
5. Downstream data analysis (optional, see section 6 for more information)

3 Sample preparation (done by the project manager)

The project manager prepares total RNA samples. Quality of small RNA-Seq results is closely related to initial samples quality. The project manager should therefore try to avoid any contamination (Phenol, DEPC, genomic DNA, etc.) or degradation.

GenomEast Platform
CERBM GIE – IGBMC
1 rue Laurent Fries
67404 Illkirch Cedex
France

+33 3 88 65 34 26
thibault@igbmc.fr
jost@igbmc.fr
genomeast.igbmc.fr

Most RNA extraction kits using columns do not keep small RNAs. We therefore recommend using RNA purification kits retaining RNAs of small size (e.g. Trizol).

Characteristics of total RNA that should be provided to the platform	
Minimal quantity	1 µg (optimal : 2 µg).
Concentration	From 150 to 300 ng/µl.
Minimal volume	10 µl.
Quality	- OD260/OD280 ≥ 1.8 or no degradation on agarose gel or 28S/18S ≥ 1.6 and/or RIN ≥ 8 on an Agilent Bioanalyzer profile.
Shipping condition	In solution, in water on dry ice. Sample names must be clearly indicated on the tubes as well as in the platform's LIMS.

4 Quality controls

Quality controls listed below are performed and corresponding results are sent to the project manager after each of the following steps. Quality controls performed at steps 1 and 2 are also available through the platform's LIMS (http://ngs-lims.igbmc.fr).

1. Sample checking	
Quantity (Fluorometry)	≥ 1µg.
Quality (capillary electrophoresis)	28S/18S ≥ 1.6 and/or RIN ≥ 8.
2. Library preparation	
Library profile (capillary electrophoresis)	140-150 bp peak(s)
Library purity (capillary electrophoresis)	Limited presence of adapter dimers (120-130 bp band).
3. Sequencing and primary data analysis	
Total number of clusters (i.e. number of reads in Single-read and number of reads ÷ 2 in Paired-end)	≥ 250 x L millions, where L represents the total number of lanes requested by the user
Mean quality score (Phred Score)	≥ 85% of bases for 50 bp reads.

5 Results delivery

For each sample, the following results are available:

- Raw sequencing data (nucleotide sequences in FASTQ format. The files contain reads passing quality filters and do not contain dimer adapter sequences).
- A report with sequences quality controls (in PDF format).

In addition to these sample files, two files are provided for each project:

- A project report (in PDF format) containing the number of raw reads, the percentage of bases with a Phred quality score over 30 and the size of each FASTQ sequence file to be downloaded.

GenomEast Platform
CERBM GIE – IGBMC
1 rue Laurent Fries
67404 Illkirch Cedex
France

+33 3 88 65 34 26
thibault@igbmc.fr
jost@igbmc.fr
genomeast.igbmc.fr

- A text file providing the MD5 strings of each FASTQ file. The project manager can use these MD5 to check the integrity of the files after their download.

The project manager is informed of the availability of the data by email once the sequencing process is done. This email contains a login and a password to be used to retrieve the generated data on the platform FTP server.

According to the "GenomEast Platform terms and conditions of business", the project manager is responsible for his data to be saved and archived on its own. Following their transfer to the Beneficiary, the Platform guarantee the conservation of raw data only for a limited period of six months.

6 Downstream analysis (optional)

Data analysis is not part of the standard service but can be done in collaboration between the project manager and the platform. The following analyses can be performed:

- 3' adapter trimming.
- Alignment on a reference genome.
- Quantification of known miRNA and other small non-coding RNA using public databases (miRBase, Rfam, etc.).
- Normalization and statistical analysis in order to highlight significantly differentially expressed small non-coding RNA between different conditions.
- Prediction of new miRNA.

This list is not exhaustive and we recommend the project manager that would like to collaborate with the platform for data analysis to contact the platform before starting their experiment so that we can define the analyses that best fit to the project manager's needs