Microarray and sequencing platform

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: SEQUENCING

The platform offers to sequence libraries prepared by project managers.

1 Services provided

1. Library checking:
   - Library quantification and quality control by capillary electrophoresis (Bioanalyzer from Agilent or Fragment Analyzer from AATI).
2. Sequencing using Illumina HiSeq 4000 technology:
   - Loading of libraries in a flow cell and cluster generation on the Cbot (Illumina).
   - Single-read or paired-end sequencing of 50 or 100 bp according to the read length chosen by the project manager. We can realize other read lengths upon demand but only for complete flow cells (8 lanes).
   - Libraries submitted by the project manager will be sequenced on entire lane(s) of flow cell (e.g. libraries will not be multiplexed with libraries prepared by the platform or by other users). We run into a sequencing process as soon as the 8 channels of the flow cell are fulfilled with samples to be sequenced with the same read length.
3. 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).
4. Downstream data analysis (optional, see section 4 for more information)

2 Quality controls

Libraries are checked according to quality criteria (see table below) dependent on the sequencing technology in use on the platform. After library validation, the platform commits to use the Illumina sequencing technology following Illumina's recommendations. The platform will not be responsible for the quality of the results.

Library checking	
Library profile (capillary electrophoresis)	Average size ranging from 200 to 600 bp.
Library purity (capillary electrophoresis)	Limited presence of adapter dimers (120-130 pb band).

Quality control results are sent to the project manager. These results are also available through the platform's LIMS (http://ngs-lims.igbmc.fr).

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

3 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 without 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.
- 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 guarantees the conservation of raw data only for a limited period of six months.

4 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. We recommend the project managers who 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 their needs.