

Frequently Asked Questions

(FAQs)

How do I submit a sample?

If this is your first time using the service please familiarize yourself with our requirements for sample submission. It is also requested that you come and speak to Prof. Petersen regarding your project ideas, so that we can answer any questions that you may have, and best advise you on what technique would suit your needs. After this please fill out our registration form, (we hope to have in the future a LIMS for this).

What criteria must my samples meet for submission?

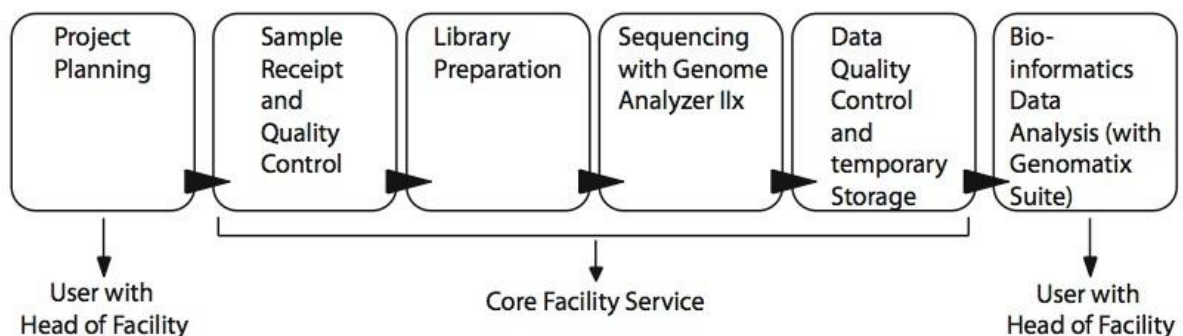
Please view the Sample Submission Form

What run/read length options do I have?

All libraries prepared are by default Paired End compatible, (with the exception of smallRNA), you just need to decide if Single read or Paired End option is to answer your question. Current read lengths available are 36, 72 and 101 bases (either single or PE). Due to the introduction of v5 chemistry (when validated) we hope in the future to offer 150 base reads.

What is the Illumina / Your workflow?

The diagram below outlines the workflow we follow:



For a detailed library preparation guide please view the individual protocols.

How are libraries checked prior to loading?

All libraries are firstly checked for their concentration using the Qubit fluorometer, then ran on the Agilent BioAnalyzer using either the DNA1000 chip or the HighSensitivity DNA chip, (determined by concentration outcome), to check for fragment size. Only libraries passing these two checks will be loaded. Currently no qPCR is performed on libraries due to the problem of cost as each library type for each genome would need to be titrated to produce the standards for the qPCR this is not realistic for a service lab.

What about storage of my Sample?

Unfortunately as with many labs freezer space is of a premium. We will not be able to store your samples or prepared libraries indefinitely. At the moment we will retain samples for a maximum of 2 months, it is your responsibility to collect any remaining sample or their prepared libraries, from us once we have released the data, or in the worst case if there is a preparation problem and we cannot continue. A week before the 2 months is up you will receive an e-mail informing you that the samples will be discarded if not picked up. There will be no exceptions.

How many reads will I get?

We cannot guarantee a set number of reads as this does depend on the library type and base composition of the organism sequenced. However it is the norm to obtain in the region of 200 to 300 K cluster per tile (1 cluster = 1 read; 120 tiles per lane).

What are your quality checks for data being returned?

We have on each flowcell a PhiX control which is our standard to check the instrument is working within manufacturers' standards. After first base, there are certain calibration scores which must be within a certain range for the run to proceed. Throughout the run we monitor the quality score of the bases at each cycle, pass filter cluster density, focus quality and the images being produced. Post run via the pipeline the IVC plots which are generated are checked for "evenness" with these plots we can see any base bias etc. (base bias is normal for certain protocols eg mRNAseq). For those samples which can be aligned to a known genome it is possible to check the alignment score, and the error rate for those samples. If things are looking good we proceed to send the data. In the cases where we deem data to be of bad quality we reserve the right not to release and will repeat the affected sample(s) free of charge.

What data do I receive?

You will receive both a sequence.txt file and an export.txt file. The sequence file contains only those sequences which have passed the chastity filter (this is an Illumina defined filter in an attempt to remove low quality sequences), and contains one read per line in fastq format. The export file contains all sequences including those which have failed chastity, in a tab-delimited fashion, with various columns to describe the read (eg, position on flow cell, passed filter which is noted in column 21/22 depending on the run type performed). The export file also contains the alignment position of each sequence. Both sequence and export files have information as to the quality of each base.

What is the size of the files I will receive?

For 36 base reads in the region of 1-1,5 GB of compressed files. The file is compressed using the linux Tar function. Roughly double this for each 36 base increment, and double for each PE read as you will receive two files, (read one and read 2)

What is the name of my file?

flowcellID_Sample name_ SampleID_UserSurname_lanenumbe_rsequence/export.txt.tar.gz

Samples name is that which you registered your sample with.

Sample ID is the tracking number we provide; in the future we hope to have a LIMS which will automatically generate a bar-code.

Lanenumbe_r is the lane your sample ran on and denoted s_# (for PE reads there is a second number telling if it is read 1 or 2 s_#_1/2).

How long do you store my data?

All data will be stored for 2 months. It is your responsibility to save and back up your data, we cannot do this for you. After 2 months your data will be automatically deleted. You should receive an e-mail one week before the data is removed to remind you to download you data.