



Questions Frequently Asked about Illumina Sequencing

Table of Contents

Samples, sample submission and sequencing	1
Cost and turnaround time	2
Data and analysis	3
Reference and acknowledgement policy	4

Samples, sample submission and sequencing

What kind of samples can I have sequenced? You can sequence genomic, RNA, microRNA (small RNA) and Chromatin Immunoprecipitated (ChIP'd) samples. Sequencing of other specific input material may also be available. Please contact us with your specific project needs and we will be able to advise if we can accept your samples.

For transcriptome projects, can I send total RNA or do I need to make cDNAs?

You can send us tissue/cells, total RNA or DNase-I treated RNA. The costs associated with library production vary depending on what starting material you submit to us. Please contact the Project Manager for more information.

How should I isolate my RNA/DNA or immunoprecipitate my chromatin?

We can advise on production of RNA, isolation of DNA or immunoprecipitation, but cannot provide optimized protocols. All protocols will need to be optimized in your own lab to ensure it works for your sample in your lab. All protocols include biological samples and as such cannot be completely optimized or standardized for use in every lab. Some variability is unavoidable.

What happens if the run/library fails?

This depends on a number of issues. The submitted DNA is first quality checked, and if there are any issues then we would discuss them with you in advance. If library construction fails, then we would contact you and work out a solution - there is no standard solution. For sequencing itself, a single lane is initially run to assess the quality of the library. If the lane fails any one of our quality metrics, the lane is reviewed internally to identify the source of the problem. If there are concerns with the library construction, the customer will be contacted to discuss possible solutions and options. Basically we would work with you at each step to obtain the best possible results.

How many lanes should I run?

The number of lanes of sequencing you need to run will depend on your experiment and your sample. Sequencing requirements will vary between researchers and between samples. In general for human samples, microRNA experiments require one single-end tag lane of sequence while WTSS, ChIP and genomic sequencing require a minimum of 2 paired-end lanes (although more is recommended).



Cost and turnaround time

How much will my sequencing cost?

Sequencing cost is dependent on the type of sample you submit for sequencing and the number of lanes of sequence that are needed. The total cost for sequencing includes the cost for construction of a library suitable for sequencing on an Illumina Genome Analyzer II and the cost for each lane of sequence. To obtain an accurate quote, please contact us with details about your project and we would be able to supply you with an accurate quotation.

Why aren't prices posted on this website?

The Genome Sciences Centre operates on a collaborative, cost-recovery basis. This means that the costs vary between sample types, and are updated regularly to ensure they accurately reflect the actual cost of the sequence.

Can I save money by constructing the libraries myself?

You are welcome to submit pre-constructed libraries for sequencing. In most instances, the construction of the library will cost you more than if you submitted it to us, as the pipeline has been streamlined to offer maximum value. We do caution researchers that libraries are not always straightforward to construct, and inexperience may result in libraries that have a higher failure rate.

If I submit multiple samples, can I obtain a discount?

We offer solely on a cost-recovery basis. Therefore, we are unable to offer discounts.

How long will it take to receive my data?

It is difficult to accurately predict how long it will take before you receive data from submitted samples. Samples are placed in the queue in the order in which they are received, so the time will depend on the pipeline at the time of sample submission. For more information about the queue at any given time, please contact the Project Manager.



Data and analysis

In what format do I receive my sequencing data?

The costs for Illumina sequencing include providing sequence and quality data in the form of a 'qseq' file. When a reference genome is available, read alignment files in Binary Sequence Alignment Map format (bam files) are available on request at no additional charge. If the UCSC genome browser supports the reference genome used for the alignment, additional files for visualizing the alignment results in the genome browser are also available on request at no additional charge. These file types include 'wig' (and/or 'bigwig'), 'bedgraph' (and/or 'bigbedgraph') and 'bed' (and/or 'bigbed') formats. Data processing included in the costs quoted above is limited to the provision of these file types.

Do you send my files on DVD or email them to me?

The amount of data is too large to be emailed or mailed. Data is available via password protected ftp

Do you keep all data (picture files...)?

No, images are not stored. Other data is stored for a minimum of 6 months.

How much pass filter data am I guaranteed?

We don't have minimum data guarantees, as the data yield depends too much on the sample supplied. We would use our internal QC standards to ensure that the best possible data is generated for each sample.

Does the cost of sequencing include alignments? If so, what software is used?

The price includes a preliminary alignment with our default aligner bwa (<http://bio-bwa.sourceforge.net/bwa.shtml>). Additional alignment, with specific client specified parameters or the use of additional aligners such as Maq, Eland, [NovoAlign](#) etc is available upon request and may require additional payment.

Does that provide multiple-matches, none, only 1, or random placement in the case of ambiguity?

The resulting bwa alignment (referred to as a bam file) is a binary alignment/map file which includes all reads. Predefined columns within the file will flag sequence reads which passed or failed chastity filtering, and mark those reads which are unique and have either multiple, or no matches. For more information regarding the bam output format and tools for working with these alignment files, please refer to the samtools <http://samtools.sourceforge.net>

How many mismatches (and at what quality?) are tolerated?

Most alignment software defaults to 2 base mismatches per alignment as it is fast and fairly effective. We routinely run our alignment software with this as the default setting.

Did you align the filtered data or all of the raw data?

All of the raw data is aligned. No filtering is applied directly to the reads. Reads are however flagged using various filtering approaches (including quality) allowing reads to be filtered out by the end user.



Reference and acknowledgement policy

If the GSC sequences my samples, what are my obligations for referencing when I publish my work? The BC Cancer Agency Genome Sciences Centre (GSC) monitors our contributions to the scientific community. This is done as part of our ongoing support for the activities of our collaborators, as well as to ensure we meet the requirements of both our funding partners and our charter as a non-profit agency. In order to achieve this, we require our collaborators to acknowledge the work performed by the GSC in any or all of the following ways:

1. The GSC does not request or require co-authorship on publications when data has been generated through their cost-recovery collaborative service alone, i.e. when no intellectual contribution has been made.
2. Where intellectual contributions have been made by the GSC, collaborators are required to discuss potential and pending publications based on these contributions with the relevant GSC scientists or staff to identify appropriate co-authorship.
3. At a minimum, acknowledgement of the work of the GSC should be included in peer-reviewed publications. The following sentence can be incorporated into the Acknowledgements section of the article: "The authors wish to acknowledge the BC Cancer Agency Genome Sciences Centre, Vancouver, Canada for [activity]".
4. Alternatively, acknowledgements can appear in the text of peer-reviewed publications, for example in the Materials and Methods sections. A suggested sentence for inclusion is: "[Activity] was performed by the BC Cancer Agency Genome Sciences Centre, Vancouver, Canada".