Requirements Document for Illumina Sequencing: Constructed Library Submission

Genome Sciences Centre, BC Cancer Agency

January 2014

Table of Contents

Submission of Constructed Libraries

Illumina sells the oligonucleotides required for library construction separately; therefore, they can be purchased and used in conjunction with other library preparation reagents. Commercial sources of these oligonucleotides are also available and have been successfully used for library construction. The oligonucleotide sequences are available from the sequencing forum SEQanswers. Please note the following when ordering oligonucleotides:

- The adapter starting with GATC, must be phosphorylated (lower strand in Figure 1 below).
- The adapter can be synthesized with a special linkage between the 3' terminal T and the preceding C. This is a phosphorothioate linkage which renders this overhanging T more nuclease resistant (after annealing the top and bottom adapter oligonucleotides). This provides nuclease resistance for this base, diminishing the probability of adapter dimers (Figure 1).
- Information on some of the sequences used by Illumina is available here.

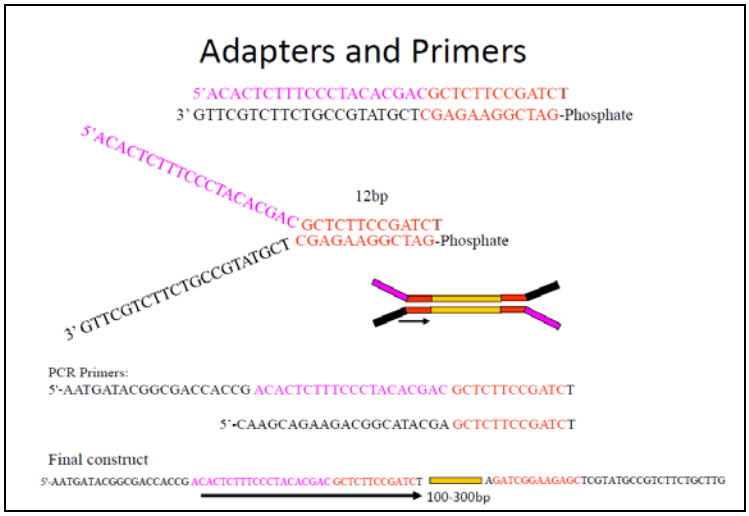

Figure 1. GSC compatible adapters and primers and final construct.

Library Construction Method for DNA

- Libraries constructed using Illumina's TruSeq DNA sample preparation kits are now compatible with the Genome Sciences Centre's (GSC's) internally constructed libraries.
  - TruSeq Universal Adapter:

    5'AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
- Libraries constructed using alternate kits must be compatible with the GSC's standard sequencing primers.

	Name	Sequence
Adapter 5'	Standard GSC	CAAGCAGAAGACGGCATACGAGAT**NNNNNN**CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
Adapter 3'	Standard GSC	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Seq read 1 primer	Standard GSC r1	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Seq read 2 (index) primer	GSC Index	GATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCG
Seq read 3 primer	Standard GSC r3	CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

- GSC compatible primer sequences and/or index sequences must be provided to the GSC prior to sample submission.
- For indexed libraries, index sequences must also be compatible and provided to the GSC prior to sample submission (see below for GSC index sequences).
- ***We encourage you to perform the research and assess your project's compatibility with our DNA pipeline.***

Library Construction Method for RNA

- For mRNA-seq libraries the standard TruSeq Illumina kit is compatible with our pipeline, as well as, our in house chemistries.
- Commercially available products, with more specialized applications for other types of RNA libraries, may also be compatible with our pipeline.
- microRNA libraries constructed using external adaptors may not be compatible with our sequencing pipeline. The GSC has only constructed microRNA libraries using our in house microRNA adapters; therefore, external adapters and library construction protocol(s) would be untested.
- microRNA constructed libraries cannot be sequenced on the Illumina MiSeq® platform as the miSeq runs at a higher temperature, causing the sequencing primers to be stripped off, resulting in the absence of any reads.
- ***We encourage you to perform the research and assess your project's compatibility with our RNA pipeline.***

GSC Compatible Adaptor & Primer Sequences

	Name	Sequence
Adapter 5'	TruSeq	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC**NNNNNN**ATCTCGTATGCCGTCTTCTGCTTG
Adapter 3'	TruSeq	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
Seq read 1 primer	TruSeq r1	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Seq read 2 (index)	TruSeq Index	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC

primer		
Seq read 3 primer	TruSeq r3	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
Adapter 5'	Direct Seq	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC**NNNNNN**ATCTCGTATGCCGTCTTC TGCTTGC**TG**
Adapter 3’	DirectSeq	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT**G AC**
Seq read 1 primer	Direct Seq r1	ACACTCTTTCCCTACACGACGCTCTTCCGATCTC**TG**
Seq read 2 (index) primer	TruSeq Index	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC
Seq read 3 primer	Direct Seq r3	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT**GAC**

GSC Index Sequences

Sequence Name	Sequence
IIA_000200	CAAGCAGAAGACGGCATACGAGAT***CGTGAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000201	CAAGCAGAAGACGGCATACGAGAT***CTGATC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000202	CAAGCAGAAGACGGCATACGAGAT***GGGGTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000203	CAAGCAGAAGACGGCATACGAGAT***CTGGGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000204	CAAGCAGAAGACGGCATACGAGAT***AGCGCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000205	CAAGCAGAAGACGGCATACGAGAT***CTTTTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000206	CAAGCAGAAGACGGCATACGAGAT***TGTTGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000207	CAAGCAGAAGACGGCATACGAGAT***AGCTAG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000208	CAAGCAGAAGACGGCATACGAGAT***AGCATC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000209	CAAGCAGAAGACGGCATACGAGAT***CGATTA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000210	CAAGCAGAAGACGGCATACGAGAT***CATTCA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000211	CAAGCAGAAGACGGCATACGAGAT***GGAACT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000212	CAAGCAGAAGACGGCATACGAGAT***ACATCG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000213	CAAGCAGAAGACGGCATACGAGAT***AAGCTA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000214	CAAGCAGAAGACGGCATACGAGAT***CAAGTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000215	CAAGCAGAAGACGGCATACGAGAT***GCCGGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000216	CAAGCAGAAGACGGCATACGAGAT***CGGCCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000217	CAAGCAGAAGACGGCATACGAGAT***TAGTTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000218	CAAGCAGAAGACGGCATACGAGAT***GCGTGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000219	CAAGCAGAAGACGGCATACGAGAT***GTATAG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000220	CAAGCAGAAGACGGCATACGAGAT***CCTTGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000221	CAAGCAGAAGACGGCATACGAGAT***GCTGTA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000222	CAAGCAGAAGACGGCATACGAGAT***ATGGCA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000223	CAAGCAGAAGACGGCATACGAGAT***TGACAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000224	CAAGCAGAAGACGGCATACGAGAT***GCCTAA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000225	CAAGCAGAAGACGGCATACGAGAT***GTAGCC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000226	CAAGCAGAAGACGGCATACGAGAT***AGTCTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000227	CAAGCAGAAGACGGCATACGAGAT***TATCGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

IIA_000228	CAAGCAGAAGACGGCATACGAGAT***AATTAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000229	CAAGCAGAAGACGGCATACGAGAT***CCGGTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000230	CAAGCAGAAGACGGCATACGAGAT***CATGGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000231	CAAGCAGAAGACGGCATACGAGAT***TCTGAG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000232	CAAGCAGAAGACGGCATACGAGAT***AAGTGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000233	CAAGCAGAAGACGGCATACGAGAT***ATTATA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000234	CAAGCAGAAGACGGCATACGAGAT***CCAGCA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000235	CAAGCAGAAGACGGCATACGAGAT***GGACGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000236	CAAGCAGAAGACGGCATACGAGAT***TGGTCA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000237	CAAGCAGAAGACGGCATACGAGAT***TACAAG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000238	CAAGCAGAAGACGGCATACGAGAT***TCGCTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000239	CAAGCAGAAGACGGCATACGAGAT***GAGAGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000240	CAAGCAGAAGACGGCATACGAGAT***CCGTAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000241	CAAGCAGAAGACGGCATACGAGAT***ATCGTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000242	CAAGCAGAAGACGGCATACGAGAT***CCACTC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000243	CAAGCAGAAGACGGCATACGAGAT***CAGCAG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000244	CAAGCAGAAGACGGCATACGAGAT***CGCGGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000245	CAAGCAGAAGACGGCATACGAGAT***GAATGA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000246	CAAGCAGAAGACGGCATACGAGAT***GCGCCA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000247	CAAGCAGAAGACGGCATACGAGAT***CTCTAC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000248	CAAGCAGAAGACGGCATACGAGAT***CACTGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000249	CAAGCAGAAGACGGCATACGAGAT***ATGTTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000250	CAAGCAGAAGACGGCATACGAGAT***GTCCTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000251	CAAGCAGAAGACGGCATACGAGAT***ATCAGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000252	CAAGCAGAAGACGGCATACGAGAT***TAGGAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000253	CAAGCAGAAGACGGCATACGAGAT***TGAGTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000254	CAAGCAGAAGACGGCATACGAGAT***TTGCGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000255	CAAGCAGAAGACGGCATACGAGAT***GGTTTC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000256	CAAGCAGAAGACGGCATACGAGAT***TAAGGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000257	CAAGCAGAAGACGGCATACGAGAT***TCGGGA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000258	CAAGCAGAAGACGGCATACGAGAT***TTCGAA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000259	CAAGCAGAAGACGGCATACGAGAT***GCGGAC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000260	CAAGCAGAAGACGGCATACGAGAT***ATTGGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000261	CAAGCAGAAGACGGCATACGAGAT***TGCTTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000262	CAAGCAGAAGACGGCATACGAGAT***CCTATT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000263	CAAGCAGAAGACGGCATACGAGAT***TCTTCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000264	CAAGCAGAAGACGGCATACGAGAT***ATAGAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000265	CAAGCAGAAGACGGCATACGAGAT***CGCCTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000266	CAAGCAGAAGACGGCATACGAGAT***CTAAGG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000267	CAAGCAGAAGACGGCATACGAGAT***TTATTC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000268	CAAGCAGAAGACGGCATACGAGAT***TGGAGC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000269	CAAGCAGAAGACGGCATACGAGAT***CTTCGA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000270	CAAGCAGAAGACGGCATACGAGAT***GGAGAA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000271	CAAGCAGAAGACGGCATACGAGAT***TTTCAC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000272	CAAGCAGAAGACGGCATACGAGAT***GATCTG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000273	CAAGCAGAAGACGGCATACGAGAT***GCATTT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000274	CAAGCAGAAGACGGCATACGAGAT***GTTTGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

IIA_000275	CAAGCAGAAGACGGCATACGAGAT***CTATCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000276	CAAGCAGAAGACGGCATACGAGAT***GCTCAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000277	CAAGCAGAAGACGGCATACGAGAT***GCCATG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000278	CAAGCAGAAGACGGCATACGAGAT***TTCTCG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000279	CAAGCAGAAGACGGCATACGAGAT***TCCGTC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000280	CAAGCAGAAGACGGCATACGAGAT***TGTGCC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000281	CAAGCAGAAGACGGCATACGAGAT***TGCCGA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000282	CAAGCAGAAGACGGCATACGAGAT***AAACCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000283	CAAGCAGAAGACGGCATACGAGAT***GGCCAC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000284	CAAGCAGAAGACGGCATACGAGAT***TCAAGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000285	CAAGCAGAAGACGGCATACGAGAT***CGTACG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000286	CAAGCAGAAGACGGCATACGAGAT***AGATGT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000287	CAAGCAGAAGACGGCATACGAGAT***GATGCT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000288	CAAGCAGAAGACGGCATACGAGAT***AGGAAT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000289	CAAGCAGAAGACGGCATACGAGAT***AAAATG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000290	CAAGCAGAAGACGGCATACGAGAT***ATTCCG***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000291	CAAGCAGAAGACGGCATACGAGAT***TATATC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000292	CAAGCAGAAGACGGCATACGAGAT***CAGGCC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000293	CAAGCAGAAGACGGCATACGAGAT***GGTAGA***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000294	CAAGCAGAAGACGGCATACGAGAT***TTGACT***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT
IIA_000295	CAAGCAGAAGACGGCATACGAGAT***CGAAAC***CGGTCTCGGCATTCCTGCTGAACCGCTCTTCCGATCT

Indexed Libraries

- Index sequences and protocols used for library construction must be confirmed to be compatible with the GSC's sequencing pipeline prior to sample submission.
- Index sequences should have a balanced mix of bases at each position to ensure no focusing issues during the index read.
- Data from libraries submitted as pools will be split by provided index sequences. Splitting by index is performed with a one-base-mismatch tolerance (i.e. index reads with one mismatch from the expected index will still be counted towards that index). Therefore, please ensure that no two indices in the same pool are different by less than 2 bases, as any pair or indices that differ by only one base (e.g. AGTCCA and ATTCCA) will be considered ambiguous in the splitting process and reads with either index will be lost from the split bam files.
- It is important to ensure that when you submit pooled libraries, the indices in each pool contain an equal representation of each base in each position (e.g. do not have all your indices in one pool start with 'A'). The HiSeq instruments do not focus well when all clusters have the same base in the same position.

Library Quality

- Once constructed libraries have been approved and submitted, the GSC will perform a QC check on the samples to assess the quality and quantity of the library. If the libraries do not pass our quality and quantity QC checks or if there are apparent issues with the libraries, we will contact you. However, despite passing QC checks, libraries may not result in satisfactory data. The GSC is not responsible for the quality of submitted constructed libraries or for the generation of data from such libraries.

- It is expected that libraries have been purified using a suitable PCR clean-up kit; have an A260/280 ratio of > 1.8 and an A260/230 ration of > 1.2. These requirements are mandatory.
- The Agilent Bioanalyzer can be used to provide visual examination of the constructed libraries. The "perfect" library electropherogram, (Figure 2), shows a single peak of the expected size. Common additional forms include primer dimers (Figure 3), adapter dimers (Figure 3), and broader bands of higher molecular weight (MW) than the expected peak.

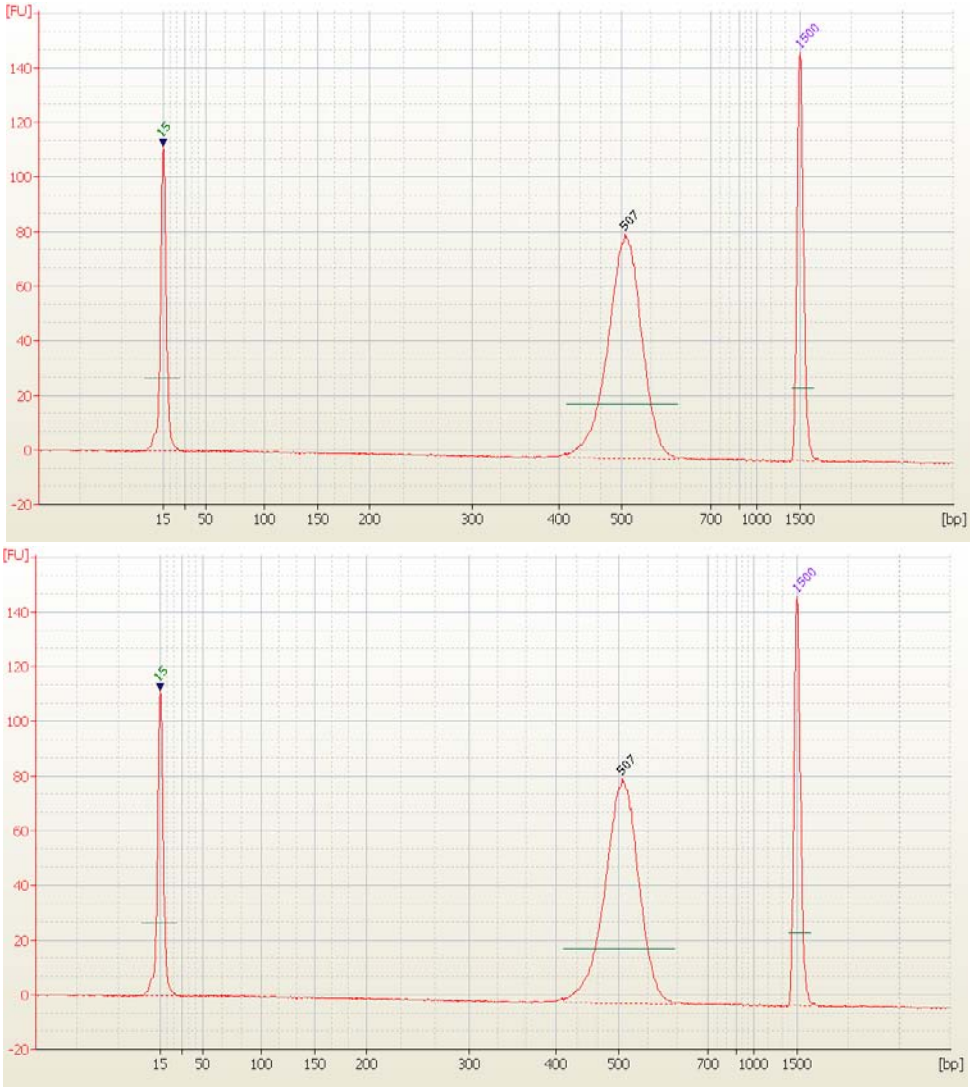

Figure 2: Electropherogram of Constructed Library

- Primer dimers can be minimized by size-selection (e.g. using magnetic beads or gel excision) but may not pose a problem unless they completely dominate the reaction. Useful data have been obtained from libraries despite the presence of 50% primer dimers.
- Adapter dimers can be a problem because they sequence efficiently. As a result, whatever the proportions of adapter dimers present in your library, at least the same or more of the proportion of reads will be seen in your final data files. Because adapter dimers are very efficient at generating clusters on the flow cell, usually, a higher

proportion of reads will be seen in the final data files. Adapter dimers can be minimized by adjusting the adapter:insert ratios during library construction and exercising care in gel extraction or other size selection steps.

- Larger MW Fractions are typically more hump shaped forms when visualized on the Bioanalyzer and are probably a result of excess amplification during the final PCR step. While some amounts of larger MW fractions are tolerable, the library should be re-amplified from the gel extracted material if they are too prominent.

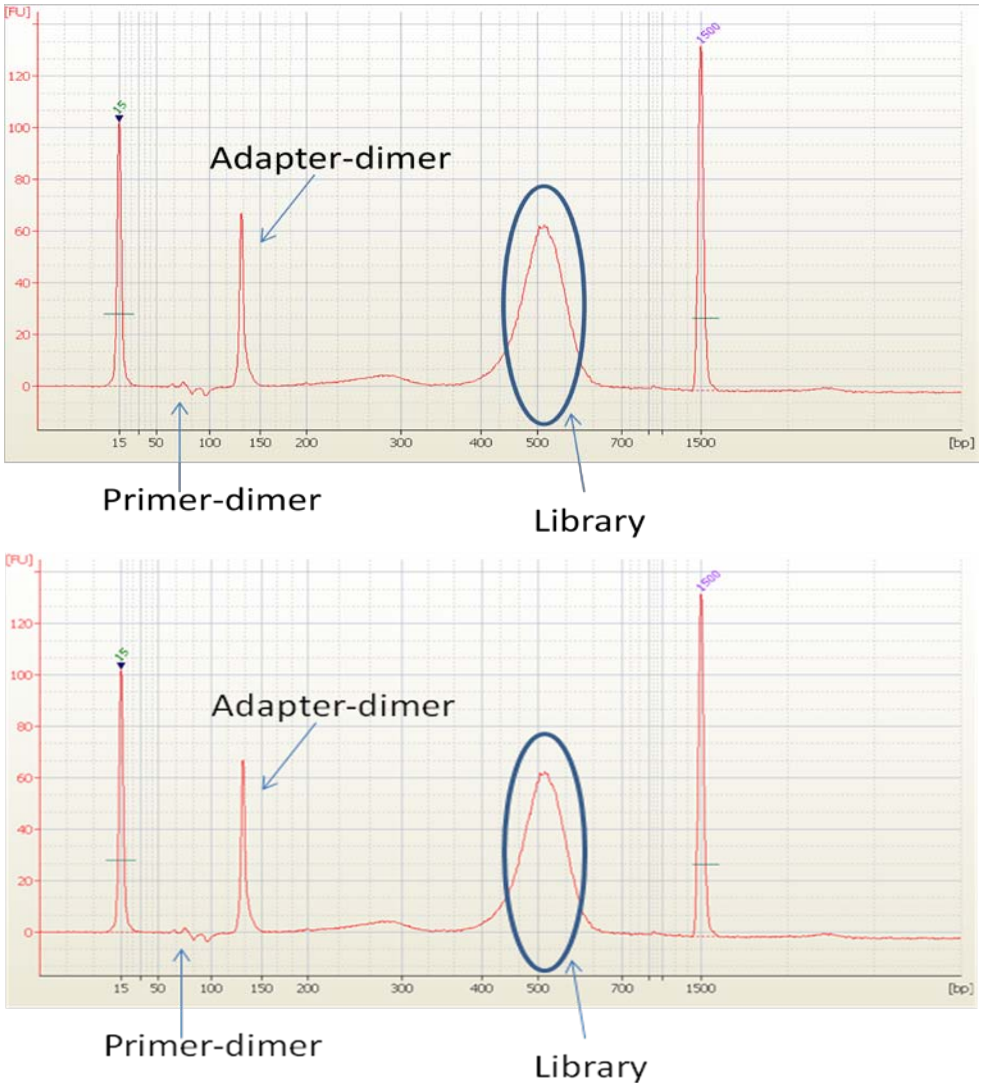

Figure 3: Electropherogram showing primer dimers and adapter dimers

- The proper size of the inserts will depend on the application and the type of sequencing performed. As a general guideline, libraries submitted for sequencing should have inserts no longer than 600bp.
- The GSC cannot guarantee that constructed libraries will be accepted for sequencing even after samples have been submitted and, we cannot guarantee the quality of sequencing data from these libraries.

Library Quantity

- The minimum volume and concentration to submit is 10uL and 3.2 nM.
- For multiplexed libraries, you are required to submit a final pool at > 5 nM per lane of sequencing (i.e. do not submit individual libraries before pooling).
- For maximum yields we strongly recommend that > 5 nM dilutions are quantitated using qPCR (or Qubit). **Nanodrop is not recommended** because readings are not accurate at the very dilute concentrations utilized in next generation sequencing protocols. Additionally, spectrophotometric methods can overestimate library quantities by including unadaptered, or incorrectly adaptered, products.
- Average size should be determined with the Agilent Bioanalyzer 2100 High-Sensitivity DNA kit.

The quantity and concentration of your library is critical. It is not uncommon to require 5 times the minimal amounts given above in order to achieve full sequence output capacity. If your sample concentrations fall just below the minimum required amount, please let us know as we *may* still be able to run your samples if you do not require maximum sequence yields.

References & Acknowledgement Policy

We require our collaborators to acknowledge the work performed by the GSC in the following ways depending on the level of collaborative effort between the GSC and the researcher:

- If the data was generated as a fee for service (cost-recovery collaborative service alone, i.e. when no intellectual contribution has been made), the GSC should be cited using either of the methods below:
    - In peer-reviewed publications incorporate the following sentence into the Acknowledgements section of the article: “The authors wish to acknowledge the Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, Canada for [activity]”.
    - Or alternatively, the GSC can be cited in the Materials and Methods section. A suggested sentence for inclusion is: “[Activity] was performed by the Michael Smith Genome Sciences Centre, BC Cancer Agency, Vancouver, Canada”.
- Where intellectual contributions have been made by researchers at 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. This will ensure that our scientists and staff receive the appropriate credit for their work.

The Michael Smith Genome Sciences Centre (GSC) tracks contributions to the wider scientific community. This is a means to measure 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.