ILLUMINA LIBRARY STRUCTURE

All Paired-End Format sequencing on the HiSeq2000 and **all sequencing** of any type on the MiSeq or HiSeq4000 **MUST** HAVE FULL-LENGTH P5 and P7 sequences. (Some of the small RNA libraries and alternative genomic library constructions use a partial P7, this is not supported by the HiSeq2000 PE, HiSeq4000, and MiSeq.)

P5: 5' AAT GAT ACG GCG ACC ACC GA 3'

P7: 5' CAA GCA GAA GAC GGC ATA CGA GAT 3'

When checking your libraries by cloning and standard Sanger sequencing, this is the arrangement :

P5---Read1primer—INSERT—Index/Read2(rc)primer--index—P7(rc)	TruSeq (single index) library structure
P5---Read1primer—INSERT—Read2primer(rc)—P7(rc)	Old PE library structure
P5---i5 index---Read1primer—INSERT—Index/Read2(rc)primer--i7 index—P7(rc)	Nextera or TruSeq HT (dual index) library

Note: the Read2primer and the IndexReadprimer are the same location, opposite strands

CUSTOM SEQUENCING PRIMERS

You may use a custom sequencing primer for the Read1 (side 1) sequence on either the HiSeq or the MiSeq. IF you need a custom sequencing primer on side 2, then you must use the MiSeq instrument. The only exception is if you have 7 lanes, then you can run them together on the HiSeq with the custom primer.

SIDE1 vs SIDE2

You must read side1. Even if you don't want the sequence from side1, it is required to set the location coordinates, matrix and phasing information for the run. If there is no data collected for side1, then the side2 read will not be captured. If you want a custom sequence read, e.g. 25 bases on side1 and 125 bases on side2, we can do that on the MiSeq. Please ask. **SEQUENCE DIVERSITY**

Libraries which begin with a linker, barcode, or other "non-random" sequence will not perform well unless they are base-balanced. This is particularly important on the MiSeq which has only 1 lane. If your sample has the same sequence in the first

6 positions, then we must add a balancer DNA, e.g. PhiX, so that the instrument can be calibrated to capture each base type throughout the sequencing run. If your sample has barcodes in these first 6 positions, and they are base-balanced, it should be fine. If they are not balanced, then we will need to add a balancer DNA. Base-balanced sequences have 1 of each base at each position. For barcodes, we can go with 3 of 4 bases at each of the first 3 positions if each of the four bases is present at least twice.
For Example: (AGC) (CTG) (ATC)

If you have any questions, please ask. If you are submitting libraries with a custom sequencing primer, please include a diagram and the results of any topo cloning or other documentation. The more we know the better we can help you. Nemo@umassmed.edu