Integrated GENOMICS Cores

PhiX Additions to Run

Using PhiX as a control when sequencing low diversity libraries on Illumina platforms

When sequencing libraries with low base diversity, unbalanced nucleotide composition can negatively impact cluster template registration on non-patterned flow cells. It can also reduce Q30 scores and data output.

What is nucleotide diversity?

High nucleotide diversity is when a library has roughly equal proportions of all 4 nucleotides (A, C, G, and T) in every cycle of the run.

Why is it important?

Nucleotide diversity is required for effective template generation on Illumina sequencing platforms and is important for the generation of high-quality data. Diversity is especially important during the first 4–7 cycles of the first sequencing read for MiSeq and HiSeq 2500 systems. The sequencing software uses images from these early cycles to identify the location of each cluster in a process called template generation.

Nucleotide diversity is also important for the first 25 cycles in the first sequencing read on all sequencing platforms because this is when phasing/pre-phasing, color matrix corrections, and the pass filter calculations occur. These corrections and calculations are used in base calling and quality score calculations for all cycles in a run for the clusters that pass filter.

To compensate for low base diversity in libraries, Illumina recommends spiking in PhiX Control v3 Library (FC-110-3001, commonly referred to as “PhiX”) for sequencing. The PhiX Control v3 Library has a diverse base composition (45% GC and 55% AT) that provides the balanced fluorescent signals that low diversity sample libraries lack during each sequencing cycle. This, in turn, assists with template registration and improves overall run quality.

How much PhiX is recommended?

Platform	PhiX %
MiSeq	Minimum 5%
HiSeq 2500	Minimum 10%
HiSeq 4000	5-20%
NovaSeq	Minimum 10%

Please note: PhiX is added by volume. Differences in clustering efficiency between PhiX and the sample library can affect the PhiX spike-in percentage required to achieve the above-targeted percent PhiX aligned. For example, more PhiX may be required if the sample library clusters more efficiently than PhiX.

References:

https://support.illumina.com/bulletins/2017/02/how-much-phix-spike-in-is-recommended-when-sequencing-low-divers.html

https://support.illumina.com/bulletins/2016/07/what-is-nucleotide-diversity-and-why-is-it-important.html