# Comparison of MiSeq, iSeq and NovaSeq

Marcel Brun, Matthew A. Stull, Emma Howard, Joshua Hill, Richard Metz, and Charles D. Johnson.

## Methods

Study based on 200 plant DNA samples. 12 samples per 96 well plate were QCed using Fragment Analyzer® automated capillary gel electrophoresis system and Denovix® spectrophotometer and found to be high molecular weight, intact, pure, DNA. Samples were quantified by Picogreen® assay on a Perkin Elmer Janus NGS Express® automated liquid handler and normalized to 100ng input to a Perkin Elmer custom NEXTflex® DNA Library preparation with enzymatic fragmentation method. Library preparation was performed on a Sciclone NGSx® automated liquid handler. Protocol inputs were a 13-minute fragmentation time and size selection of 520-720 bp. Samples were indexed with Nextflex HT® barcoded adaptors. Libraries were amplified with 10 cycles of PCR and then QCed on the fragment analyzer, quantified by PG, normalized, and pooled. Pooled samples were quantified using qPCR. The libraries were loaded on three Illumina platforms: MiSeq® Nano 2X150 bp loaded at 15pM, Illumina iSeq™ 2X150 loaded at 50pM, and Illumina NovaSeq™ 6000 S2 2X150 loaded at 2.60nM. The expected yield for the NovaSeq System was 6 Gb data per sample using a whole flow cell. Demultiplexing and generation of FASTQ files was done using bcl2fastq, filtering the reads that did not pass qc requirements.

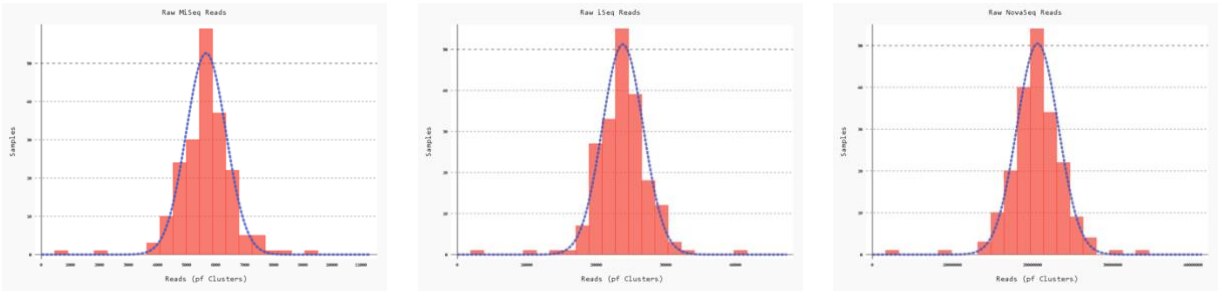
## Results

The resulting output was of 1.1M, 4.7M and 4.1G clusters for the MiSeq, iSeq and NovaSeq systems, respectively. For iSeq the number of clusters was above specs, which may be the results of overloading it (this was our first run of the iSeq System). Table 1 below shows the statistics over the 200 samples for the three platforms. The results for iSeq show d a low percentage of reads passing filters, which may result from the overloading.

*Table 1 – Reads statistics for the three platforms*

| | MiSeq | iSeq | NovaSeq |
|---|---|---|---|
| **Total Reads** | 1,132,166 | 4,733,905 | 4,105,344,212 |
| **Average Per Sample** | 5,661 | 23,670 | 20,526,721 |
| **Standard Deviation** | 923 | 3,781 | 3,237,837 |
| **Median** | 5,699 | 23,730 | 20,382,510 |
| **Minimum** | 530 | 2,465 | 2,602,220 |
| **Maximum** | 9,082 | 39,821 | 33,725,054 |
| **Pass Filter** | 94.9 % | 67.6 % | 77.3 % |

We computed the histogram of number of reads per sample for the three platforms. We expect to have a good agreement between these three histograms if they produce similar output. Figure 1 show the three histograms for the 3 platforms. The blue lines show a gaussian density based on the estimated median/variance. In these figures we can see a good agreement between them.



*MiSeq* *iSeq* *NovaSeq*

*Figure 1. Histogram of Number of Reads per Sample*

To be able to compare the histograms in the same scale, we computed normalized histograms, where instead of number of reads, we use the proportion of number of reads (# Reads / Total number of reads), and instead of counting the number of samples in each bin, we use the proportion too (# Samples / 200). Figure 2 shows the three normalized histograms together, were we can see a very good agreement between the three platforms.

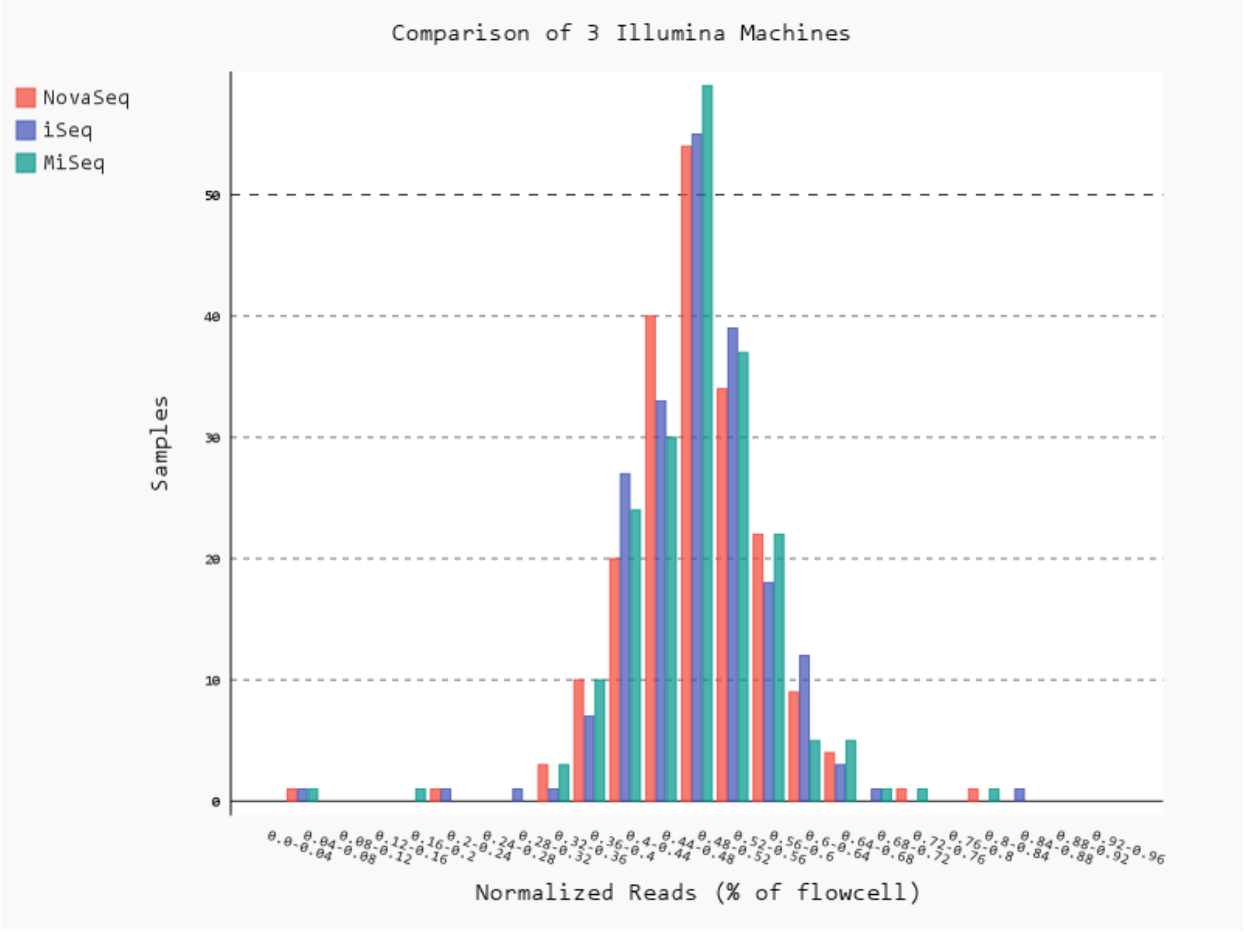
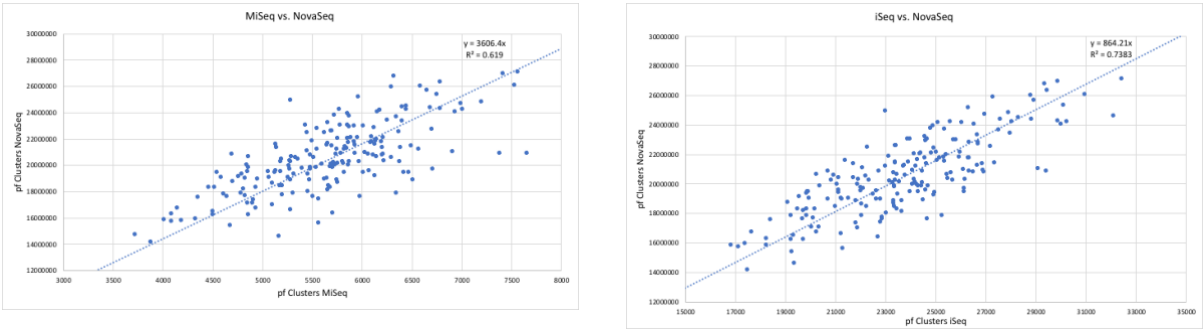



*Figure 2. Combined Normalized Histogram*

Finally, we generated scatter plots and computed the linear correlation between the two small platforms (MiSeq and iSeq) versus NovaSeq. Figure 3 shows these scatter plots, between MiSeq and NovaSeq (left) and between iSeq and NovaSeq (Right). The $r^2$ values obtained were 0.665 and 0.758 for MiSeq and iSeq, respectively.




*(a)* *(b)*

*Figure 3. Scatter Plots of NovaSeq vs MiSeq (a) and iSeq (b)*

# Conclusions

Both platforms (MiSeq and iSeq) show a high level of agreement with NovSeq, besides the fact that the iSeq system was slightly overloaded, with iSeq displaying a better correlation. Future runs of iSeq with better loading should improve these results.