testing

Wes Horton

July 14, 2016

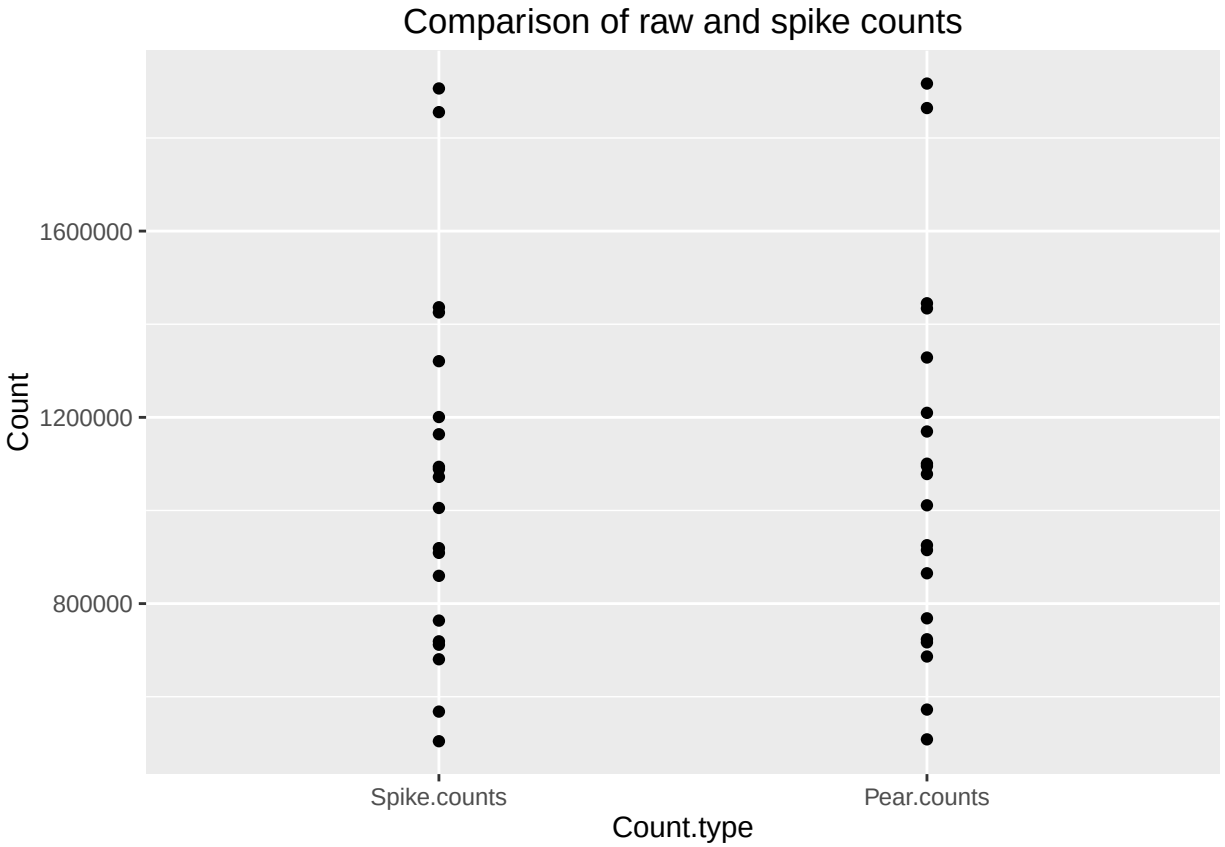
Overview

When testing primer independence using samples 1-20 from DNA160609LC, Burcu noticed a four-fold range in spike counts between samples. For example, the spike corresponding to V1/J1-1 in sample 1 could have a count of 10, whereas the same spike has a count of 40 in sample 2. We usually give the sequencing core unequal amounts of DNA in our samples, but these 20 samples have the same initial concentration of spikes. We would expect them to then have the same final concentration as well, assuming that they amplified similarly.

We can compare the 9-bp spike count totals with the PEAR fastq read counts and determine if the distributions are similar.

Compare distributions

```
##    Sample Spike.counts Pear.counts
## 1       1       680413      686374
## 2       2      1200647     1209673
## 3       3       859639      865117
## 4       4      1088957     1095477
## 5       5      1436461     1445022
## 6       6      1093733     1100437
## 7       7      1425387     1433990
## 8       8      1855464     1864230
## 9       9      1320599     1328501
## 10     10      1072110     1078247
## 11     11      1906327     1916805
## 12     12      1005399     1011097
## 13     13       763504      768413
## 14     14       504436      508421
## 15     15       919145      925440
## 16     16       711713      716742
## 17     17       568063      572435
## 18     18       908939      914924
## 19     19       718989      723413
## 20     20      1163698     1169669
```



We can see that the distributions are almost identical. These results suggest that the variation came from any of the following (or a combination of them):

1. Sample preparation - variable amounts of spikes were pipetted into the samples
2. PCR amplification - samples have same starting material, but individual reactions amplified differently in the thermocycler
3. Sequencing - Samples seeded unevenly on the sequencing machine

Next Steps

Do we have quantitative data of sample concentration prior to PCR (nanodrop or cubit?), if so, we should check that distribution. We can also ask Bob if he saw a difference in seeds.