# Assessing the Differences in 16S rRNA Gene Sequencing Due to High Fidelity DNA Polymerase

Marc A Sze and Patrick D Schloss

Supplemental

**Table S1: Wilcoxson Rank Sum Results of High versus Low GC groups based on the median percent of the V4 hypervariable region or the whole genome**

| Number of Cycles | Polymerase | Rank Sum Statistic | P-value | BH | Test Type |
|---|---|---|---|---|---|
| 30x | PL | 404.0 | 9.7e-05 | 2.4e-03 | V4 Region |
| 35x | PL | 56.0 | 5.8e-03 | 5.4e-02 | V4 Region |
| 20x | PHU | 59.0 | 8.5e-03 | 5.4e-02 | V4 Region |
| 30x | Q5 | 532.0 | 9.5e-03 | 5.4e-02 | V4 Region |
| 25x | PL | 61.0 | 1.1e-02 | 5.4e-02 | V4 Region |
| 30x | K | 546.0 | 1.5e-02 | 5.4e-02 | V4 Region |
| 25x | PHU | 64.0 | 1.5e-02 | 5.4e-02 | V4 Region |
| 20x | PL | 69.0 | 2.6e-02 | 8.1e-02 | V4 Region |
| 30x | PHU | 74.0 | 4.3e-02 | 1.1e-01 | V4 Region |
| 35x | PHU | 77.0 | 5.6e-02 | 1.3e-01 | V4 Region |
| 25x | ACC | 82.5 | 9.0e-02 | 1.9e-01 | V4 Region |
| 25x | K | 85.0 | 1.1e-01 | 1.9e-01 | V4 Region |
| 35x | ACC | 88.0 | 1.4e-01 | 2.3e-01 | V4 Region |
| 20x | ACC | 90.0 | 1.6e-01 | 2.5e-01 | V4 Region |
| 25x | Q5 | 96.0 | 2.4e-01 | 3.3e-01 | V4 Region |
| 20x | Q5 | 98.0 | 2.7e-01 | 3.6e-01 | V4 Region |
| 35x | Q5 | 99.0 | 2.9e-01 | 3.6e-01 | V4 Region |
| 35x | K | 102.5 | 3.5e-01 | 4.1e-01 | V4 Region |
| 20x | K | 108.0 | 4.6e-01 | 5.3e-01 | V4 Region |
| 30x | ACC | 139.0 | 7.0e-01 | 7.6e-01 | V4 Region |
| 30x | Q5 | 1340.0 | 4.1e-08 | 1.0e-06 | Whole Genome |
| 30x | K | 1354.5 | 9.8e-08 | 1.2e-06 | Whole Genome |
| 20x | PL | 249.0 | 1.5e-07 | 1.2e-06 | Whole Genome |
| 25x | PHU | 237.0 | 6.9e-06 | 2.9e-05 | Whole Genome |
| 25x | PL | 237.0 | 6.9e-06 | 2.9e-05 | Whole Genome |
| 25x | Q5 | 227.0 | 7.1e-05 | 2.5e-04 | Whole Genome |

| Number of Cycles | Polymerase | Rank Sum Statistic | P-value | BH | Test Type |
|:---:|:---:|:---:|:---:|:---:|:---:|
| 20x | Q5 | 224.0 | 1.3e-04 | 4.0e-04 | Whole Genome |
| 20x | PHU | 220.0 | 2.7e-04 | 7.6e-04 | Whole Genome |
| 25x | ACC | 223.5 | 3.4e-04 | 7.8e-04 | Whole Genome |
| 35x | K | 223.5 | 3.4e-04 | 7.8e-04 | Whole Genome |
| 30x | ACC | 217.0 | 4.7e-04 | 9.7e-04 | Whole Genome |
| 25x | K | 210.0 | 2.1e-03 | 4.1e-03 | Whole Genome |
| 30x | PHU | 206.0 | 2.6e-03 | 4.4e-03 | Whole Genome |
| 20x | K | 201.0 | 6.3e-03 | 9.8e-03 | Whole Genome |
| 20x | ACC | 194.0 | 1.4e-02 | 2.0e-02 | Whole Genome |
| 35x | PHU | 186.0 | 2.9e-02 | 4.0e-02 | Whole Genome |
| 30x | PL | 961.0 | 1.2e-01 | 1.5e-01 | Whole Genome |
| 35x | Q5 | 168.0 | 1.4e-01 | 1.6e-01 | Whole Genome |
| 35x | ACC | 152.0 | 3.8e-01 | 4.1e-01 | Whole Genome |
| 35x | PL | 115.0 | 6.4e-01 | 6.7e-01 | Whole Genome |

**Table S2: ANOVA results of differences in number of OTUs between polymerases in fecal and mock samples**

| DF | Sum Squares | Mean Squares | F value | P-value | BH | Cycle | Sample Type |
|---|---|---|---|---|---|---|---|
| 4 | 3.66 | 0.91 | 1.25 | 3.5e-01 | 3.5e-01 | 20x | Fecal Sample |
| 4 | 6.90 | 1.73 | 2.84 | 6.1e-02 | 1.2e-01 | 25x | Fecal Sample |
| 3 | 4.67 | 1.56 | 2.55 | 1.0e-01 | 1.4e-01 | 30x | Fecal Sample |
| 4 | 13.01 | 3.25 | 16.35 | 2.4e-05 | 9.7e-05 | 35x | Fecal Sample |
| 2 | 364.56 | 182.28 | 15.82 | 1.7e-03 | 2.2e-03 | 20x | Mock Sample |
| 4 | 3478.53 | 869.63 | 75.15 | 2.7e-09 | 1.1e-08 | 25x | Mock Sample |
| 4 | 2206.58 | 551.64 | 10.28 | 1.8e-05 | 3.6e-05 | 30x | Mock Sample |
| 4 | 1971.92 | 492.98 | 2.31 | 1.0e-01 | 1.0e-01 | 35x | Mock Sample |

**Table S3: Tukey post-hoc results of differences in number of OTUs between polymerases in fecal and mock samples**

| Difference | Lower | Upper | P Adjusted | Comparison | Cycle | Sample Type |
|---|---|---|---|---|---|---|
| 25.00 | -17.32 | 67.32 | 0.40 | PL-ACC | 35x | Fecal Sample |
| 24.25 | -18.07 | 66.57 | 0.43 | PL-K | 35x | Fecal Sample |
| 20.50 | -21.82 | 62.82 | 0.58 | Q5-ACC | 35x | Fecal Sample |
| 19.75 | -22.57 | 62.07 | 0.61 | Q5-K | 35x | Fecal Sample |
| 13.50 | -28.82 | 55.82 | 0.86 | PHU-ACC | 35x | Fecal Sample |
| 12.75 | -29.57 | 55.07 | 0.88 | PHU-K | 35x | Fecal Sample |
| 11.50 | -30.82 | 53.82 | 0.91 | PL-PHU | 35x | Fecal Sample |
| 7.00 | -35.32 | 49.32 | 0.98 | Q5-PHU | 35x | Fecal Sample |
| -4.50 | -46.82 | 37.82 | 1.00 | Q5-PL | 35x | Fecal Sample |
| 0.75 | -41.57 | 43.07 | 1.00 | K-ACC | 35x | Fecal Sample |
| -13.50 | -20.36 | -6.64 | 1.3e-03 | PL-PHU | 20x | Mock Sample |
| -38.50 | -46.60 | -30.40 | 5.1e-09 | Q5-K | 25x | Mock Sample |
| -37.50 | -45.60 | -29.40 | 7.3e-09 | PHU-K | 25x | Mock Sample |
| 36.50 | 28.40 | 44.60 | 1.1e-08 | K-ACC | 25x | Mock Sample |
| -35.50 | -43.60 | -27.40 | 1.5e-08 | PL-K | 25x | Mock Sample |
| 22.03 | 9.31 | 34.74 | 1.8e-04 | K-ACC | 30x | Mock Sample |
| -15.48 | -25.20 | -5.75 | 5.7e-04 | Q5-K | 30x | Mock Sample |
| -18.78 | -31.49 | -6.06 | 1.5e-03 | PHU-K | 30x | Mock Sample |
| 15.55 | 3.03 | 28.07 | 9.0e-03 | PL-ACC | 30x | Mock Sample |

**Table S4: Kruskal-Wallis results of differences in error rate, sequences with an error, and chimera prevalence between polymerases in mock samples**

| Rank Sum Statistic | DF | P-value | BH | Cycle | Test Type |
| --- | --- | --- | --- | --- | --- |
| 7.85 | 2 | 2.0e-02 | 2.0e-02 | 20x | Error Rate |
| 12.95 | 4 | 1.2e-02 | 1.5e-02 | 25x | Error Rate |
| 28.31 | 4 | 1.1e-05 | 4.3e-05 | 30x | Error Rate |
| 15.00 | 4 | 4.7e-03 | 9.4e-03 | 35x | Error Rate |
| 7.85 | 2 | 2.0e-02 | 2.0e-02 | 20x | Sequences with Error |
| 15.98 | 4 | 3.0e-03 | 5.2e-03 | 25x | Sequences with Error |
| 29.22 | 4 | 7.0e-06 | 2.8e-05 | 30x | Sequences with Error |
| 15.44 | 4 | 3.9e-03 | 5.2e-03 | 35x | Sequences with Error |
| 8.91 | 2 | 1.2e-02 | 1.2e-02 | 20x | Chimera Prevalence |
| 15.54 | 4 | 3.7e-03 | 4.9e-03 | 25x | Chimera Prevalence |
| 21.41 | 4 | 2.6e-04 | 1.1e-03 | 30x | Chimera Prevalence |
| 16.28 | 4 | 2.7e-03 | 4.9e-03 | 35x | Chimera Prevalence |

**Table S5: Dunn's Post-Hoc results of differences in error rate, sequences with an error, and chimera prevalence between polymerases in mock samples**

| Chi Squared | Z Statistic | P Adjusted | Comparison | Cycle | Test Type |
|:---:|:---:|:---:|:---:|:---:|:---:|
| 7.85 | 2.77 | 2.8e-03 | PHU - PL | 20x | Error Rate |
| 7.85 | -1.64 | 5.0e-02 | PL - Q5 | 20x | Error Rate |
| 12.95 | 3.26 | 5.6e-04 | K - Q5 | 25x | Error Rate |
| 12.95 | 2.68 | 3.7e-03 | K - PHU | 25x | Error Rate |
| 12.95 | 2.01 | 2.2e-02 | ACC - Q5 | 25x | Error Rate |
| 12.95 | 1.88 | 3.0e-02 | PL - Q5 | 25x | Error Rate |
| 28.31 | -4.46 | 4.1e-06 | ACC - K | 30x | Error Rate |
| 28.31 | 4.32 | 7.7e-06 | K - PL | 30x | Error Rate |
| 28.31 | -2.69 | 3.6e-03 | ACC - Q5 | 30x | Error Rate |
| 28.31 | 2.65 | 4.0e-03 | K - PHU | 30x | Error Rate |
| 28.31 | 2.37 | 8.8e-03 | K - Q5 | 30x | Error Rate |
| 28.31 | -2.00 | 2.3e-02 | PL - Q5 | 30x | Error Rate |
| 15.00 | 3.14 | 8.6e-04 | K - PHU | 35x | Error Rate |
| 15.00 | -3.05 | 1.2e-03 | PHU - PL | 35x | Error Rate |
| 15.00 | -2.33 | 1.0e-02 | ACC - K | 35x | Error Rate |
| 15.00 | -2.05 | 2.0e-02 | ACC - PL | 35x | Error Rate |
| 15.00 | -1.79 | 3.7e-02 | PHU - Q5 | 35x | Error Rate |
| 15.00 | 1.68 | 4.7e-02 | K - Q5 | 35x | Error Rate |
| 7.85 | 2.77 | 2.8e-03 | PHU - PL | 20x | Sequences with Error |
| 7.85 | -1.64 | 5.0e-02 | PL - Q5 | 20x | Sequences with Error |
| 15.98 | 3.37 | 3.7e-04 | K - Q5 | 25x | Sequences with Error |
| 15.98 | 2.79 | 2.6e-03 | K - PHU | 25x | Sequences with Error |
| 15.98 | 2.76 | 2.9e-03 | ACC - Q5 | 25x | Sequences with Error |
| 15.98 | 2.14 | 1.6e-02 | ACC - PHU | 25x | Sequences with Error |
| 15.98 | 1.86 | 3.1e-02 | K - PL | 25x | Sequences with Error |
| 29.22 | 4.17 | 1.5e-05 | K - Q5 | 30x | Sequences with Error |

| Chi Squared | Z Statistic | P Adjusted | Comparison | Cycle | Test Type |
| --- | --- | --- | --- | --- | --- |
| 29.22 | 3.89 | 5.0e-05 | K - PHU | 30x | Sequences with Error |
| 29.22 | 3.57 | 1.8e-04 | PL - Q5 | 30x | Sequences with Error |
| 29.22 | -3.41 | 3.2e-04 | PHU - PL | 30x | Sequences with Error |
| 29.22 | -1.81 | 3.5e-02 | ACC - K | 30x | Sequences with Error |
| 29.22 | 1.76 | 3.9e-02 | ACC - PHU | 30x | Sequences with Error |
| 15.44 | -3.71 | 1.0e-04 | PHU - PL | 35x | Sequences with Error |
| 15.44 | 2.38 | 8.6e-03 | PL - Q5 | 35x | Sequences with Error |
| 15.44 | 2.38 | 8.7e-03 | K - PHU | 35x | Sequences with Error |
| 15.44 | 1.85 | 3.2e-02 | ACC - PHU | 35x | Sequences with Error |
| 15.44 | -1.85 | 3.2e-02 | ACC - PL | 35x | Sequences with Error |
| 8.91 | 2.98 | 1.4e-03 | PHU - PL | 20x | Chimera Prevalence |
| 15.54 | 3.37 | 3.7e-04 | K - PHU | 25x | Chimera Prevalence |
| 15.54 | -2.76 | 2.9e-03 | PHU - PL | 25x | Chimera Prevalence |
| 15.54 | -2.68 | 3.7e-03 | ACC - K | 25x | Chimera Prevalence |
| 15.54 | -2.01 | 2.2e-02 | ACC - PL | 25x | Chimera Prevalence |
| 15.54 | 1.98 | 2.4e-02 | K - Q5 | 25x | Chimera Prevalence |
| 21.41 | -4.23 | 1.2e-05 | ACC - PL | 30x | Chimera Prevalence |
| 21.41 | 2.89 | 1.9e-03 | PL - Q5 | 30x | Chimera Prevalence |
| 21.41 | -2.84 | 2.2e-03 | ACC - K | 30x | Chimera Prevalence |
| 21.41 | -2.59 | 4.8e-03 | PHU - PL | 30x | Chimera Prevalence |
| 21.41 | -2.05 | 2.0e-02 | ACC - Q5 | 30x | Chimera Prevalence |
| 21.41 | -1.73 | 4.2e-02 | K - PL | 30x | Chimera Prevalence |
| 16.28 | -3.71 | 1.0e-04 | ACC - PL | 35x | Chimera Prevalence |
| 16.28 | -2.65 | 4.0e-03 | PHU - PL | 35x | Chimera Prevalence |
| 16.28 | -2.38 | 8.7e-03 | ACC - K | 35x | Chimera Prevalence |
| 16.28 | -2.12 | 1.7e-02 | ACC - Q5 | 35x | Chimera Prevalence |