**Table S31.  Runtime on HPC computer cluster for BioHansel and Snippy for 1 synthetic Illumina MiSeq dataset with 10X-1000X genome coverage (n=10).**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Tool** | **Scheme** | **Runtime on 1 core for 1 synthetic FASTQ dataset with varying genome coverage (MM:SS)** | | | |
| **10X (Std Dev)** | **50X (Std Dev)** | **100X (Std Dev)** | **1000X (Std Dev)** |
| **BioHansel** | Typhi  (136 k-mers) | 00:13 (00:04) | 00:19 (00:04) | 00:31 (00:05) | 03:20 (00:11) |
| Heidelberg (404 k-mers) | 00:11 (00:04) | 00:20 (00:04) | 00:31 (00:03) | 03:37 (00:12) |
| Enteritidis (634 k-mers) | 00:12 (00:04) | 00:23 (00:04) | 00:34 (00:05) | 03:48 (00:10) |
| Typhimurium (860 k-mers) | 00:14 (00:05) | 00:25 (00:03) | 00:35 (00:03) | 04:13 (00:09) |
| **Snippy** | Typhi | 01:10 (00:06) | 04:32 (00:24) | 09:26 (00:54) | 50:10 (06:36) |
| Heidelberg | 01:06 (00:03) | 04:44 (00:47) | 09:15 (01:47) | 48:10 (03:59) |
| Enteritidis | 01:05 (00:03) | 04:20 (00:11) | 08:26 (00:22) | 46:16 (01:12) |
| Typhimurium | 01:12 (00:14) | 04:28 (00:17) | 08:43 (00:34) | 49:39 (04:17) |

The performance of BioHansel on genome assemblies and synthetic Illumina reads was also measured on a desktop with the following specifications:  CPU : Intel i5-6500 (4) @ 3.20GHz, GPU: Intel HD Graphics 530; RAM: 16Gb; OS: Ubuntu 18.04.2 LTS x86\_64. BioHansel was run on either one core or 4 cores by specifying the number of threads (-t) at runtime, and with either only one output (Simple results) or all 3 outputs (Simple results, Results, and K-mer results).  Each run consisted of 250-1000 genome assemblies of SE, SH, ST, and Typhi (supplementary Tables X-Y).

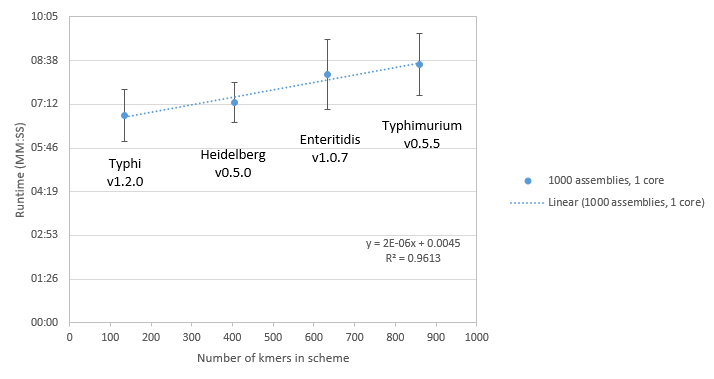


Figure S2. BioHansel runtime per 1000 genome assemblies on one core using different schemes in a high performance computer cluster environment.  Each data point represents an average of at least 8 experiments for each serovar.

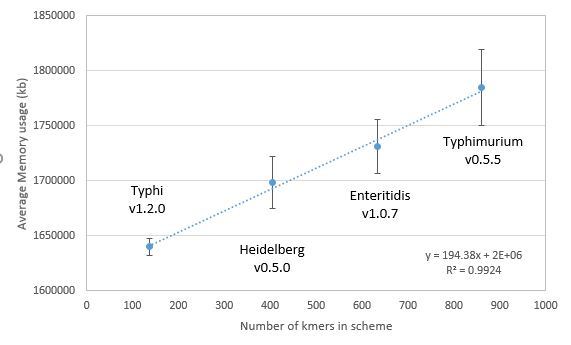


Figure S3.  Average BioHansel memory usage on a HPC cluster.  The datasets included 1000 assemblies of Typhi, SH, SE, and ST.  Each data point represents the average of at least 8 different experiments.

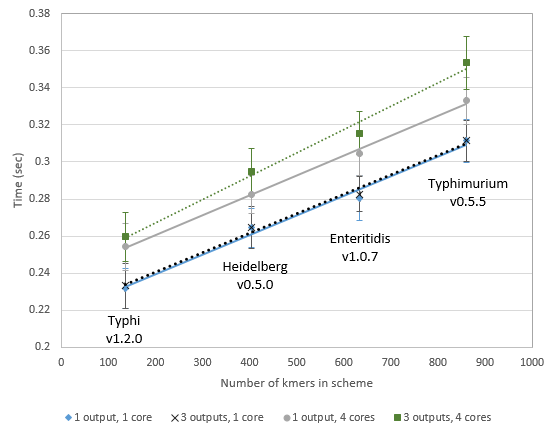


Figure S4.  Runtime per assembly per core for different schemes, using 1 or 4 cores on a personal computer, specifying either 1 or 3 outputs.   Each point represents an average runtime calculated from runs done on 250-1000 assemblies (n=40).