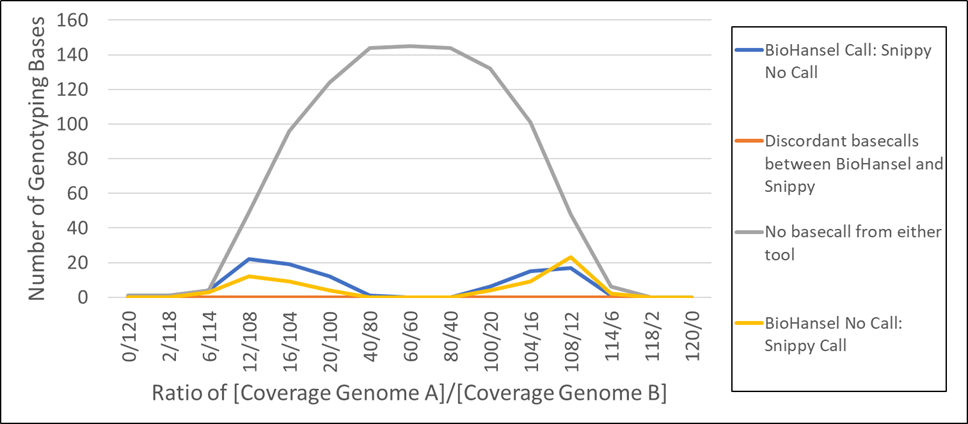
**Table S19.**  Comparison of base detection results in synthetically contaminated datasets between BioHansel and Snippy using defined genotyping bases from five schemes: SE, SH, ST, Typhi, and MTB.  The two most divergent genotypes for each scheme were artificially mixed at different fold coverage to make up a total of 120X coverage for each read set. The comparison was restricted to the BioHansel genotyping bases for each scheme (SH=202 bases; SE=317 bases; ST=430 bases; Typhi=68 bases; MTB=62 bases), resulting in 1079 bases compared for each of the 15 coverage ratios, with a total of 16,185 basecalling comparisons.  The results are categorized as described in the column headers.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ratio of [Coverage genome A]/[Coverage Genome B] | Fully concordant basecalls between tools | Discordant basecalls between BioHansel and Snippy | No basecall from either tool | BioHansel Call: Snippy No Call | BioHansel No Call: Snippy Call |
| 0/120 | 1078 | 0 | 1 | 0 | 0 |
| 2/118 | 1078 | 0 | 1 | 0 | 0 |
| 6/114 | 1068 | 0 | 4 | 4 | 3 |
| 12/108 | 996 | 0 | 49 | 22 | 12 |
| 16/104 | 955 | 0 | 96 | 19 | 9 |
| 20/100 | 939 | 0 | 124 | 12 | 4 |
| 40/80 | 934 | 0 | 144 | 1 | 0 |
| 60/60 | 934 | 0 | 145 | 0 | 0 |
| 80/40 | 935 | 0 | 144 | 0 | 0 |
| 100/20 | 937 | 0 | 132 | 6 | 4 |
| 104/16 | 954 | 0 | 101 | 15 | 9 |
| 108/12 | 991 | 0 | 48 | 17 | 23 |
| 114/6 | 1070 | 0 | 6 | 1 | 2 |
| 118/2 | 1079 | 0 | 0 | 0 | 0 |
| 120/0 | 1079 | 0 | 0 | 0 | 0 |
| Total | 15027 | 0 | 995 | 97 | 66 |
| Percentage (n=16,185) | 92.85% | 0.00% | 6.15% | 0.60% | 0.41% |

****

**Figure S1.** Comparison of base detection results in synthetically contaminated datasets between BioHansel and Snippy using defined genotyping bases from five schemes: SE, SH, ST, Typhi, and MTB.  The two most divergent genotypes for each scheme were artificially mixed at different fold coverage to make up a total of 120X coverage for each read set, as shown in the X axis. The comparison was restricted to the BioHansel genotyping bases for each scheme (SH=202 bases; SE=317 bases; ST=430 bases; Typhi=68 bases; MTB=62 bases), resulting in 1079 bases compared for each of the 15 coverage ratios, with a total of 16,185 basecalling comparisons.  The two tools gave a fully concordant basecall for 15,027 (92.85%) of the target genotyping bases (Table S3). This graph is only showing the 1,158 results (equivalent to 7.15% of the comparisons) where there was not a fully concordant and unambiguous basecall between the 2 tools. Results are categorized as shown in the figure legend on the right.