**Supplementary Table R10-1 List of manually curated GIAB small variants**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Category** | **Chr** | **Position** | **Variant** | **Line Element** | **Correct**  **Call** | **Notes** |
| AM.INDEL | 3 | 37,083,407 | G/A | No | GIAB | PacBio misaligned in short adjacent homopols |
| AM.INDEL | 2 | 9,591,845 | CT/C | No | GIAB | PB error in long homopol |
| AM.INDEL | 8 | 73,675,279 | TAAAA/T | No | GIAB | PB misses 1bp ins in compound het in homopolymer |
| AM.INDEL | 13 | 76,646,445 | G/GA | No | GIAB | PB misses 1bp ins in compound het in homopolymer |
| AM.INDEL | 19 | 1,586,670 | A/ATTT | No | GIAB | PB calls 3bp and 4bp ins in homopol |
| AM.INDEL | 15 | 44,350,983 | C/CA | No | GIAB | PB calls 1bp and 2bp ins in homopol |
| AM.SNP | 5 | 16,287,108 | A/C | No | GIAB | PB miscalls hom as het due to homopol |
| AM.SNP | 13 | 71,512,745 | A/T | No | GIAB | PB miscalls hom as het due to homopol |
| AM.SNP | 13 | 34,840,815 | G/T | No | GIAB | PB miscalls hom as het due to homopol |
| AM.SNP | 11 | 41,384,344 | C/T | No | GIAB | PB miscalls hom as het due to homopol, and nearby het ins as hom |
| AM.SNP | 12 | 51,793,781 | A/C | No | GIAB | PB miscalls hom as het due to homopol |
| FN.INDEL | 10 | 29,087,199 | T/TCC | No | GIAB | PB calls true ins in 6bp homopol as hom ref and 1bp instead of 2 |
| FN.INDEL | 3 | 107,982,543 | AT/A | No | GIAB | PB calls true indel as hom ref due to homopol |
| FN.INDEL | 17 | 5,198,683 | C/CA | No | GIAB | PB misses true indel due to homopol |
| FN.INDEL | 4 | 149,672,221 | A/ATT | No | GIAB | PB misses true indel due to homopol |
| FN.INDEL | 2 | 152,262,374 | G/GTT | No | GIAB | PB calls true indel as hom ref due to homopol |
| FN.INDEL | 2 | 236,062,930 | G/GTT | No | GIAB | PB calls true indel as hom ref due to homopol |
| FN.SNP | 5 | 55,201,041 | A/G | No | GIAB | PB calls true snp as hom ref due to homopol |
| FN.SNP | 17 | 68,021,050 | T/A | No | GIAB | PB calls true snp as hom ref due to homopol |
| FN.SNP | 9 | 129,471,234 | T/A | No | GIAB | PB calls part of complex var as hom ref due to homopol |
| FN.SNP | 7 | 135,981,582 | T/A | No | GIAB | PB has incorrect candidates as hom ref due to homopol |
| FP.INDEL | 8 | 132,575,025 | C/CAAAAAAAAA | No | GIAB | 10bp expansion of 5bp homopolymer miscalled as 9bp expansion |
| FP.INDEL | 6 | 64,897,720 | A/AT | No | GIAB | PB calls extra ins in homopol |
| FP.INDEL | 4 | 112,819,087 | GA/G | Line Element | PacBio | PB correctly calls del in homopol in LINE as well as other SNPs |
| FP.INDEL | 7 | 38,338,238 | C/CA | No | GIAB | PB calls 1bp ins instead of 2bp ins in homopol |
| FP.INDEL | 11 | 61,993,476 | CA/C | No | GIAB | PB miscalls compound het in homopol |
| FP.INDEL | 3 | 97,014,398 | AT/A | Line Element | GIAB | PB error in long homopol in LINE |
| FP.SNP | 20 | 4,159,335 | C/T | No | GIAB | PB calls del instead of SNP in homopol |
| FP.SNP | 4 | 55,520,593 | G/A | No | GIAB | PB makes incorrect complex call in homopolymer |
| AM.INDEL | 6 | 41,984,320 | ACTAT/A | No | GIAB | PacBio misaligned due to complex variant with nearby homopol |
| FP.SNP | 9 | 6,900,971 | C/T | No | GIAB | PB calls SNP and DEL instead of just DEL due to homopolymer |
| AM.INDEL | 2 | 232,051,483 | G/GCATCATGGAGAATGGGACATCTC | Line Element | GIAB | PacBio data sometimes misaligned as 2bp del + 25bp ins, when it should be a hom 23bp ins; also in LINE |
| AM.INDEL | 2 | 232,051,483 | GCA/G | No | - | - |
| AM.INDEL | 4 | 11,468,804 | CACACATATAT/C | No | PacBio | Appears that both correctly call complex var on one haplotype but GIAB missed 2bp ins on other haplotype |
| AM.INDEL | 5 | 42,740,225 | CT/C | Line Element | PacBio | GIAB calls 1bp del in LINE, but misses SNP phased with other SNPs on other haplotype |
| AM.SNP | 4 | 165,276,021 | T/C | Line Element | PacBio | Not completely sure, but AJ trio appears to be homozygous in this region including a LINE with many SNPs, and illumina probably has mismapped reads here |
| AM.SNP | 2 | 230,174,543 | A/G | No | GIAB | PB calls this and phased del as hom instead of het |
| AM.SNP | 13 | 48,291,499 | A/C | Line Element | PacBio | short reads miss one HP in LINE with many SNPs |
| AM.SNP | 2 | 5,143,996 | G/A | Line Element | PacBio | Not completely sure, but AJ trio appears to be homozygous in this region including a LINE with many SNPs, and illumina/10x probably has mismapped reads here |
| AM.SNP | 21 | 25,668,597 | G/A | No | GIAB | PB miscalls hom as het due to 1 ref read |
| FN.INDEL | 1 | 162,491,859 | A/ATGTCTAG | No | GIAB | PB calls true complex variant hom ref |
| FN.INDEL | 15 | 26,120,981 | C/CTTACACTGGGCTTTTTGTAAGGA | No | GIAB | PB calls true complex variant hom ref |
| FN.INDEL | 15 | 41,943,823 | T/TCCTCTTCTCTCCTCTCC | No | PacBio | GIAB calls half of complex var, and other half is outside bed |
| FN.INDEL | 8 | 5,930,728 | TACAC/T | No | PacBio | PB correctly calls this a 32bp ins in long TR not in GA4GH TRs |
| FN.SNP | 6 | 57,283,620 | T/C | No | GIAB | PB misses series of nearby SNPs |
| FN.SNP | 7 | 157,385,671 | A/G | No | PacBio | error in short reads likely due to long TR not in GA4GH TRs |
| FN.SNP | 17 | 32,064,214 | A/G | Line Element | PacBio | error in short reads likely due to LINE |
| FN.SNP | 9 | 117,917,190 | A/C | No | GIAB | no reads support this in PB 15kb CCS, but good support in PB, 10kb CCS, ONT and 10x, so I suspect it's missed due to few PB HP1 reads |
| FN.SNP | 6 | 9,737,425 | T/C | No | PacBio | error in short reads likely due to long TR not in GA4GH TRs |
| FN.SNP | 6 | 8,353,625 | C/T | No | GIAB | PB calls parts of complex var in adjacent TRs as hom ref |
| FP.INDEL | 1 | 94,256,825 | A/AAC | Line Element | PacBio | ~6kb LINE with many other SNPs inside LINE |
| FP.INDEL | 2 | 153,864,971 | AT/A | Line Element | PacBio | ~6kb+~2kb LINE with many other SNPs inside LINEs |
| FP.INDEL | 11 | 23,338,682 | C/CT | Line Element | PacBio | ~1.5kb LINE but many SNPs inside and around LINE |
| FP.INDEL | 4 | 165,026,074 | A/AG | Line Element | PacBio | ~6kb LINE with many other SNPs inside LINE |
| FP.SNP | 3 | 79,181,734 | C/T | Line Element | PacBio | ~6kb LINE with many other SNPs inside LINE |
| FP.SNP | 1 | 35,034,071 | T/C | Line Element | PacBio | LINE with many other SNPs inside and outside the LINE |
| FP.SNP | 21 | 42,288,851 | C/A | Line Element | PacBio | 2kb LINE with many other SNPs inside LINE |
| FP.SNP | 8 | 46,873,565 | C/T | No | PacBio | 1kb long TR inside satellite - looks good in PB though not certain |
| FP.SNP | 9 | 22,350,168 | A/C | Line Element | PacBio | ~6kb LINE with many other SNPs inside LINE |
| FP.SNP | 4 | 94,532,444 | T/G | Line Element | PacBio | ~6kb LINE with many other SNPs inside LINE |
| FP.SNP | 20 | 1,347,896 | A/G | No | GIAB | PB makes extra paired ins/del that is the same as the SNP that it and GIAB call |
| FP.SNP | 9 | 6,900,971 | C/T | No | GIAB | PB calls SNP and DEL instead of just DEL due to homopolymer |