## **Extended data Table 1 | The summary statistics of variants discovered in 4,053 individuals of BIGCS.**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variants type | AC=1 | AC=2 | MAF≤0.1% | 0.1%<MAF≤1% | 1%<MAF≤5% | MAF>5% | Total |
| Total variants (SNPs and Indels) | 22,276,474 | 11,919,262 | 9,648,220 | 5,279,786 | 1,907,131 | 5,199,740 | 56,230,613 |
| Known variants | 10,155,543 | 6,931,714 | 8,551,334 | 5,218,588 | 1,892,992 | 5,172,409 | 37,922,580 |
| Novel variants\* | 12,120,931 | 4,987,548 | 1,096,886 | 61,198 | 14,139 | 27,331 | 18,308,033 |
| SNPs | 19,814,479 | 10,975,400 | 8,829,910 | 4,839,893 | 1,749,629 | 4,843,145 | 51,052,456 |
| Ts/Tv (Autosomal) | 1.99 | 2.09 | 2.21 | 2.16 | 2.15 | 2.18 | 2.09 |
| Het/Hom (Autosomal) | - | - | 77,674.12 | 1917.64 | 90.36 | 1.37 | 1.48 |
| Known SNPs | 9,432,687 | 6,487,985 | 7,930,364 | 4,827,755 | 1,749,356 | 4,842,499 | 35,270,646 |
| Novel SNPs\* | 10,381,792 | 4,487,415 | 899,546 | 12,138 | 273 | 646 | 15,781,810 |
| Indels (Insertions and Deletions) | 2,461,995 | 943,862 | 818,310 | 439,893 | 157,502 | 356,595 | 5,178,157 |
| Known Indels | 722,856 | 443,729 | 620,970 | 390,833 | 143,636 | 329,910 | 2,651,934 |
| Novel Indels\* | 1,739,139 | 500,133 | 197,340 | 49,060 | 13,866 | 26,685 | 2,526,223 |
| Insertions | 700,178 | 270,876 | 237,912 | 137,660 | 55,299 | 141,068 | 1,542,993 |
| Known Insertion | 173,805 | 110,741 | 162,461 | 111,025 | 46,470 | 128,634 | 733,136 |
| Novel Insertion\* | 526,373 | 160,135 | 75,451 | 26,635 | 8,829 | 12,434 | 809,857 |
| Deletions | 1,761,817 | 672,986 | 580,398 | 302,233 | 102,203 | 215,527 | 3,635,164 |
| Known Deletion | 549,051 | 332,988 | 458,509 | 279,808 | 97,166 | 201,276 | 1,918,798 |
| Novel Deletion\* | 1,212,766 | 339,998 | 121,889 | 22,425 | 5,037 | 14,251 | 1,716,366 |
|  |  |  |  |  |  |  |  |
| Variants Location | | | | | | | |
| Exon(protein-coding region) | 284,818 | 154,686 | 111,704 | 54,002 | 15,357 | 32,722 | 653,289 |
| Intron | 11,831,432 | 6,342,069 | 5,108,542 | 2,785,876 | 990,029 | 2,659,269 | 29,717,217 |
| Splice-site | 45,168 | 23,608 | 17,272 | 8,522 | 2,773 | 6,816 | 104,159 |
| UTR | 374,429 | 210,279 | 166,708 | 90,089 | 29,620 | 70,796 | 941,921 |
| Upstream | 980,243 | 513,867 | 419,617 | 230,654 | 85,249 | 233,302 | 2,462,932 |
| Downsteam | 801,760 | 423,750 | 347,360 | 191,070 | 70,611 | 193,169 | 2,027,720 |
| Intergenic | 6,595,366 | 3,475,822 | 2,847,084 | 1,573,558 | 587,315 | 1,650,000 | 16,729,145 |
|  |  |  |  |  |  |  |  |
| Annotation function by VEP | | | | | | | |
| Synonymous | 103,377 | 52,301 | 39,370 | 20,442 | 6,837 | 16,385 | 238,712 |
| Missense | 156,572 | 93,223 | 66,250 | 31,203 | 8,023 | 15,492 | 370,763 |
| Stoplost | 510 | 201 | 147 | 87 | 23 | 63 | 1,031 |
| Stopgain | 4,826 | 2,559 | 1,365 | 579 | 117 | 190 | 9,636 |
| Startlost | 646 | 310 | 221 | 96 | 27 | 60 | 1,360 |
| frameshift insertion | 3,238 | 1,109 | 592 | 234 | 66 | 121 | 5,360 |
| frameshift deletion | 8,191 | 2,391 | 1,328 | 567 | 107 | 155 | 12,739 |
| inframeshift insertion | 1,226 | 448 | 356 | 116 | 34 | 67 | 2,247 |
| inframeshift deletion | 4,679 | 1,676 | 1,514 | 593 | 101 | 148 | 8,711 |
|  |  |  |  |  |  |  |  |
| Deleterious variants | | | | | | | |
| SIFT:deleterious | 70,601 | 42,263 | 28,307 | 12,393 | 2,619 | 3,491 | 159,674 |
| Polyphen2:probably damaging | 38,391 | 22,914 | 14,439 | 6,050 | 1,086 | 1,163 | 84,043 |
| Polyphen2:possibly damaging | 25,712 | 15,540 | 10,694 | 4,788 | 991 | 1,416 | 59,141 |
| ClinVar:Pathogenic | 448 | 260 | 198 | 59 | 8 | 18 | 991 |
| ClinVar:Likely pathogenic | 130 | 98 | 50 | 13 | 2 | 3 | 296 |

\*Not present in dbSNP build 154.