

Table S2. Orientation Bias for Genic Integrations Pre-ART^a.

| Chromosome | Gene name ^b | Unique integrations with the gene (CHER) ^c | Unique integrations against the gene (CHER) ^c | Unique integrations with the gene (<i>ex vivo</i>) ^c | Unique integrations against the gene (<i>ex vivo</i>) ^c | Adjusted p-value ^d |
|--------------|------------------------|---|--|---|--|-------------------------------|
| <i>chr6</i> | <i>BACH2</i> | 26 | 5 | 60 | 72 | 0.0019 |
| <i>chr17</i> | <i>STAT5B</i> | 24 | 5 | 284 | 278 | 0.0078 |
| <i>chr16</i> | <i>ANKRD11</i> | 3 | 16 | 380 | 403 | 0.0281 |
| <i>chr22</i> | <i>HORMAD2</i> | 11 | 4 | 192 | 241 | 0.15 |
| <i>chrX</i> | <i>MECP2</i> | 4 | 12 | 228 | 267 | 0.45 |
| <i>chr17</i> | <i>VMP1</i> | 6 | 12 | 376 | 374 | 0.52 |
| <i>chr17</i> | <i>GRB2</i> | 11 | 5 | 339 | 336 | 0.52 |
| <i>chr17</i> | <i>NPLOC4</i> | 5 | 10 | 418 | 408 | 0.52 |
| <i>chr17</i> | <i>POLR2A</i> | 6 | 10 | 161 | 177 | 0.82 |
| <i>chr11</i> | <i>MALAT1</i> | 6 | 9 | 113 | 96 | 0.82 |
| <i>chr11</i> | <i>PACSI</i> | 18 | 21 | 852 | 821 | 0.84 |
| <i>chr11</i> | <i>KDM2A</i> | 9 | 10 | 806 | 724 | 0.84 |
| <i>chr19</i> | <i>CARD8</i> | 10 | 7 | 331 | 315 | 0.84 |
| <i>chr19</i> | <i>VAV1</i> | 7 | 8 | 255 | 222 | 0.84 |
| <i>chr17</i> | <i>RPTOR</i> | 14 | 12 | 820 | 807 | 0.89 |
| <i>chr17</i> | <i>CYTH1</i> | 10 | 9 | 352 | 363 | 0.89 |
| <i>chr22</i> | <i>TNRC6B</i> | 9 | 10 | 444 | 435 | 0.89 |
| <i>chr1</i> | <i>ASH1L</i> | 8 | 9 | 378 | 385 | >0.99 |

^a Data shown only for integrations into genes for which at least 15 unique integrations were detected

^b Genic coordinates mapped to hg19

^c “With” gene and “Against” gene defined as orientation of integrated provirus compared with the sense of the host gene it’s integrated into

^d Adjusted p-value determined by Fisher Test with post-hoc Benjamini-Hochberg Correction