

| <b>Projection</b> | <b>Expression</b> | <b>TSS count</b> | <b>TSS %</b> | <b>Protein coding %</b> |
|-------------------|-------------------|------------------|--------------|-------------------------|
| Orthologous       | robust equiv      | 33582            | 3.2040       | 74.59                   |
|                   | robust            | 8841             | 0.8435       | 49.36                   |
|                   | non-robust        | 77230            | 7.3684       | 36.35                   |
| Aligned           | robust equiv      | 29085            | 2.7750       | 54.51                   |
|                   | robust            | 33649            | 3.2104       | 26.31                   |
|                   | non-robust        | 383446           | 36.5840      | 20.90                   |
| Gap               | robust equiv      | 8728             | 0.8327       | 34.07                   |
|                   | robust            | 14712            | 1.4037       | 11.32                   |
|                   | non-robust        | 157133           | 14.9918      | 8.44                    |
| Unaligned         | robust equiv      | 32729            | 3.1226       | 60.57                   |
|                   | robust            | 23501            | 2.2422       | 24.85                   |
|                   | non-robust        | 245488           | 23.4217      | 19.37                   |

Projecting human TSS into mouse; when projecting human TSS into mouse, what did we find? Orthologous indicates we identified an orthologous mouse TSS, these are necessarily also aligned. Aligned indicates the human TSS could be projected to an orthologous position in mouse but that there was no mouse TSS at that position. Gap projections suggest insertion of the TSS containing sequence in human or it's deletion from mouse. Unaligned indicates the human TSS position is not alignable with the mouse genome, either through technical limitations or through genomic gains and losses. Robust denotes TSS defined as robust in human, robust equiv are a subset of robust that exhibit  $\geq 3$ TPM in at least one human CAGE library that has an equivalent library from mouse. For each category of projection and expression the percentage of TSS identified as protein coding is given.

| <b>Projection</b> | <b>Expression</b> | <b>TSS count</b> | <b>TSS %</b> | <b>Protein coding %</b> |
|-------------------|-------------------|------------------|--------------|-------------------------|
| Orthologous       | robust equiv      | 30615            | 4.6894       | 70.37                   |
|                   | robust            | 3426             | 0.5248       | 35.78                   |
|                   | non-robust        | 85612            | 13.1134      | 21.92                   |
| Aligned           | robust equiv      | 26118            | 4.0006       | 48.89                   |
|                   | robust            | 10192            | 1.5611       | 20.37                   |
|                   | non-robust        | 223630           | 34.2539      | 14.15                   |
| Gap               | robust equiv      | 6866             | 1.0517       | 25.22                   |
|                   | robust            | 3934             | 0.6026       | 9.17                    |
|                   | non-robust        | 84126            | 12.8858      | 5.15                    |
| Unaligned         | robust equiv      | 27935            | 4.2789       | 57.10                   |
|                   | robust            | 7191             | 1.1015       | 21.69                   |
|                   | non-robust        | 143215           | 21.9366      | 15.01                   |

Projecting mouse TSS into human. As for the table above but relating mouse TSS to their projection into the human genome.