

Cerberus BED file specifications

Prepared by Fairlie Reese
January 25, 2022
Mortazavi Lab, University of California, Irvine

Contact Information

Fairlie Reese
2300 Biological Sciences III
University of California Irvine
Irvine, CA 92697-2300
Telephone: (949) 824-8393
Email: freese@uci.edu

Software used

- cerberus (<https://github.com/fairliereese/cerberus/releases/tag/v1>) v1

File specifications

The transcript end site (TES) and transcript start site (TSS) regions that are output from cerberus have the following bed6+ format:

- Chromosome
- Region start
- Region end
- Region name
 - Formed from the gene ID that the region is associated with and a number to make it unique. For example, ENSG00000000460_1 or ENSG00000000460_2
- Score
- Strand
- Comma-separated list of sources that support the use of each region as a TSS or TES that were used as input into cerberus
- Novelty of region with respect to the sources used as references in the cerberus run. Will be either 'Known' or 'Novel'