Real Data Information

Christian Otto

The app include a real example data set containing 100 (originally 2,500 but reduced for computational time) DNA sequences from a ChIP-chip experiment for the protein USF1 in the human liver, ranked by how likely they are to contain the regulatory motif (Rada-Iglesias et al. 2008). Each DNA sequence was trimmed to be 500 nucleotides long. This trimming was done so the app could show results for real data while not taking several hours to run. If you are interested in the original data, please see the reference at the bottom (Rada-Iglesias et al. 2008).

To load the full data you can use `data <- system.file("Exfiles/USF1_small.fa", package="BCRANK")` (Ameur et al. 2009).

References

Ameur, Adam, Alvaro Rada-Iglesias, Jan Komorowski, and Claes Wadelius. 2009. "Identification of Candidate Regulatory SNPs by Combination of Transcription-Factor-Binding Site Prediction, SNP Genotyping and haploChIP." *Nucleic Acids Research* 37 (12): e85–85. https://doi.org/10.1093/nar/gkp381.

Rada-Iglesias, Alvaro, Adam Ameur, Philipp Kapranov, Stefan Enroth, Jan Komorowski, Thomas R. Gingeras, and Claes Wadelius. 2008. "Whole-Genome Maps of USF1 and USF2 Binding and Histone H3 Acetylation Reveal New Aspects of Promoter Structure and Candidate Genes for Common Human Disorders." *Genome Research* 18 (3): 380–92. https://doi.org/10.1101/gr.6880908.