DeepBlue: Diving into Epigenomic Data

Felipe Albrecht [1,2], Markus List[1], Christoph Bock [1,3,4] and Thomas Lengauer [1]

1.Max Planck Institute for Informatics, 66123 Saarbrücken, Germany
2 Graduate School of Computer Science, Saarland University, 66123 Saarbrücken, Germany
3 CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, 1090 Vienna, Austria
4 Department of Laboratory Medicine, Medical University of Vienna, 1090 Vienna, Austria

Large volumes of data are generated by several epigenomic consortia, including ENCODE, Roadmap Epigenomics, BLUEPRINT, and DEEP. To enable users to utilize these data effectively in the study of epigenetic regulation, we have developed the DeepBlue Epigenomic Data Server [1]. With the DeepBlue Epigenomic Data Server, we provide programmatic access to vast amounts of epigenomic data, in order to facilitate storing, organizing, searching, and retrieving of epigenetic data.

Here, we present a series of tools that build upon the DeepBlue API and enable users not proficient in scripting or programming languages to benefit from our efforts and to analyze epigenomic data in a user-friendly way: (i) an R/Bioconductor package (http://deepblue.mpi-inf.mpg.de/R) integrates DeepBlue into the R analysis workflow. The extracted data are automatically converted to GenomicRanges [2], which are supported by many related packages for analysis and visualization; (ii) a web interface (http://deepblue.mpi-inf.mpg.de) that enables users to search, select, and download the epigenomic data available in DeepBlue; (iii) a web tool for epigenomic data visualization, named DeepBlue Dive (http://dive.mpi-inf.mpg.de), which is inspired by EpiExplorer [3] and helps researchers to visually compare their own epigenomic data to data already available in DeepBlue; (iv) a web tool, named DeepBlue ML, complementary to DeepBlue Dive, which is inspired by EpiGRAPH [4] and uses LOLA [5], reporting the enrichment of epigenomic regions provided by the user among the experiments available in DeepBlue.

DeepBlue and related tools are available at http://deepblue.mpi-inf.mpg.de/. This work has been supported by German Science Ministry Grant No. 01KU1216A (DEEP project) and has been performed in the context of EU grant no. HEALTH-F5-2011-282510 (BLUEPRINT project)

1. Albrecht,F., List,M., Bock,C. and Lengauer,T. (2016) DeepBlue epigenomic data server: programmatic data retrieval and analysis of epigenome region sets. Nucleic Acids Research, doi:10.1093/nar/gkw211
2. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan M and Carey V (2013). “Software for Computing and Annotating Genomic Ranges.” PLoS Computational Biology, 9.
3. Halachev, K., Bast, H., Albrecht, F., Lengauer, T. & Bock, C. EpiExplorer: live exploration and global analysis of large epigenomic datasets. Genome Biology 13, R96 (2012).
4. Bock, C., Halachev, K., Büch, J. & Lengauer, T. EpiGRAPH: user-friendly software for statistical analysis and prediction of (epi)genomic data. Genome Biology 10, R14 (2009).
5. Sheffield N.C., Bock C. LOLA: enrichment analysis for genomic region sets and regulatory elements in R and Bioconductor. Bioinformatics 2016;32:587-589.