Newsletter Spring 2012

TRANSFAC® / ExPlain™

Expanded positional weight matrix content and data provenance

TRANSFAC has long been the gold standard for positional weight matrices and now with this release we've gone beyond the standard quarterly update of new content, introducing a redesigned matrix report that makes it easier to understand how each matrix was compiled and to compare related matrices within TRANSFAC and public domain offerings.

Download an overview of the new matrix report features.

- An expanded site alignment table now allows you to easily visualize the collection of binding sites used to compile many of the TRANSFAC matrices, including information about the experimental methods and materials used to describe the binding site with direct links to supporting references.

- A new method classification scheme for matrices plus a new overview of related factor-specific and family matrices allows you to easily identify how a matrix was constructed and to compare it to other matrices within TRANSFAC.

- Reverse complement consensus sequences and logos make it easier to compare binding sites regardless of strand orientation.

With this release we're also excited to introduce a new method of generating positional weight matrices that uses a unique 3D-structure based computational technique (see Alamanova et al , BMC Bioinformatics 2010 11:225). Using this new algorithm, combined with protein structures from PDB and homology modeling, we have begun introducing matrices for transcription factors that do not have a sufficient number of binding sites reported in the literature to use for compiling a positional weight matrix by more traditional methods.

Easier access to transcription factors, the genes they regulate, and their associated positional weight matrices

With this release we've introduced new search features that make it easier to find the information you're looking for:

- A new Matrix quick search option allows you to easily query for a positional weight matrix by name or by BIOBASE accession.

- The reorganized Search within Results option allows you to easily access connected information for transcription factors and genes. For example, perform a Gene/Protein quick search for human IRF1. Use the Search within Results pull-down menu to automatically search for all genes bound by human IRF1, or to identify its associated positional weight matrices. Interested in what transcription factors bind to and regulate IRF1 itself? Use the Search within Results pull-down menu to automatically search for all transcription factors that bind the IRF1 gene, or to export the IRF1 promoter sequence for further analysis of predicted binding sites in the Match tool.

Download an overview of the new search features.

ChIP-ChIP and ChIP-seq data sets published in the literature, or provided through the ENCODE project, provide a wealth of information about transcription factors and the genes that they regulate. Over the years we have integrated many of these data sets into TRANSFAC, processing and organizing the data into more readily usable formats. With this release we're making it even easier to access and use these data sets by introducing ChIP-ChIP and ChIP-seq experiment reports. Each report provides a summary of the experiment, combined with direct links to pre-processed data that is made available for easy export or import into supplementary tools for further analysis. For example, import the list of associated genes into the Ontology Search tool in order to identify over-represented functions associated with the genes potentially regulated by the transcription factor. Or import the list of associated genes, or the FASTA sequences for the fragments, into the ExPlain analysis platform in order to identify potential co-regulatory factors.

For a complete list of all ChIP-ChIP and ChIP-seq experiments, click the "Browse by ChIP-ChIP or ChIP-seq Experiment" button under the TRANSFAC Module Search option.

View the complete TRANSFAC statistics.