

TRANSFAC® release 2022.1

The TRANSFAC® database on transcription factors, their genomic binding sites and DNA-binding motifs (PWMs), contains these new data features:

- **MATCH Suite: Single gene analysis**

In its release 2.0 MATCH Suite now supports gene regulation analysis for single genes in addition to the previously introduced gene set studies. Specify the gene of your interest using its gene symbol, Entrez or Ensembl ID, and find the transcription factors responsible for regulation of your gene via its promoter and enhancers/silencers.

You can optionally select the tissue of your interest, the promoter type and region to be used for the analysis, or the GO terms to narrow down the identified transcription factors to the ones belonging to the chosen biological processes.

The analysis results will provide you with lists transcription factors (TFs) and respective binding sites that were predicted to regulate the specified gene in the selected conditions. Interactive visualization in genome browser will allow to observe the predicted and experimentally proven TF binding sites in the promoter and enhancers/silencers of your gene in the specified tissue (if any was selected). Analysis report will provide you with comprehensive information on the performed analysis and the obtained results.

- **Extended options to use search results as input for further queries**

Transcription factors that bind to regulatory regions of specific genes in ChIP-Seq experiments, can now be queried directly from a gene/protein search result. This extends the possibility to identify regulatory factors for a gene beyond the ones found through DNA binding sites from low-throughput experiments.

- **Integration of new human ChIP-Seq experiments from ENCODE**

17 new human transcription factor binding site ChIP-Seq experiments released by the ENCODE phase 4 project have been integrated. The data sets comprise 38,319 fragments bound by 17 distinct transcription factors, of which 10 factors were not yet covered by ChIP-Seq data in TRANSFAC.

For 14 of the sets, an existing positional weight matrix for the respective transcription factor was used together with the MATCH tool to predict altogether 28,204 best binding sites inside the fragments.

Predicted best binding sites as well as complete fragments are available in FASTA and BED format via the ChIP Experiment Reports, as are lists of genes in a distance range to the fragments as specified by the user.