TRANSFAC® release 2021.3

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

- **New silencer data**

  31,933 human and mouse silencer sequences have been imported from the SilencerDB database and lifted over to the GRCh38 and GRCm39 genome assemblies. Like enhancer reports, silencer reports display genes with which promoters the silencer interacts, tissues and cell types/lines the silencer is active in, and genomic regions such as histone modification sequences, DNase I hypersensitivity sites, and transcription factor binding sites that overlap with the silencer.

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

  83 new human transcription factor binding site ChIP-Seq experiments released by the ENCODE phase 4 project have been integrated. The data sets comprise 588,473 fragments bound by 79 distinct transcription factors, of which 51 factors were not yet covered by ChIP-Seq data in TRANSFAC.
  For 47 of the sets, an existing positional weight matrix for the respective transcription factor was used together with the MATCH tool to predict altogether 388,362 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.

- **Enhanced human SNP content**

  The 2021 dbSNP release 155 data for human has been integrated and increases the number of SNPs mapped to promoter, enhancer, and silencer sequences by more than 72,000,000 new single nucleotide variations compared to the previously used version 154.

- **Ensembl version update**

  Genomic information for genes, promoters, and ChIP fragments for the species human, mouse, rat, pig, macaque, Drosophila, and Arabidopsis is now based on Ensembl release 104.

- **Mouse genome assembly update**

  TRANSFAC now uses the latest mouse genome assembly GRCm39 (mm39).
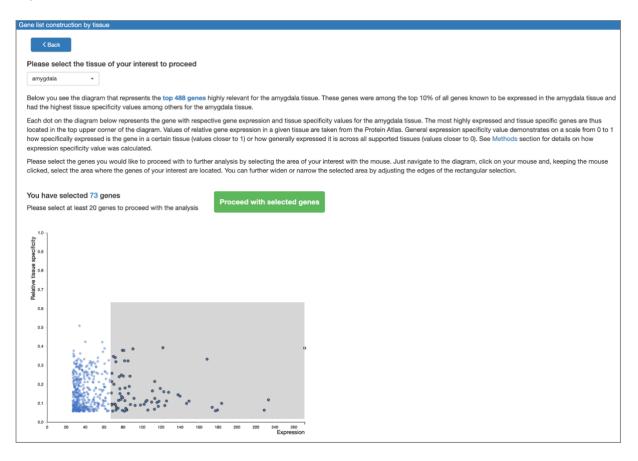
MATCH Suite toolbox of TRANSFAC 2.0 contains the following new features:

- **Tissue-specific gene list construction**

In its release 1.1 the MATCH Suite tool for gene regulation studies includes a new input gene list selection option: construct the input gene list from scratch based on the genes that are highly and specifically expressed in the selected tissue.

This option provides you with the ability to select any tissue among the 61 supported tissues and visualize the most tissue-specific genes among the top 10% of all genes known to be expressed in the selected tissue according to the information about relative gene expression levels provided by the Protein Atlas. You can further select the genes of your interest on the interactive diagram to either proceed with them to the analysis launch or just save the selected genes into the inputs of your project. Please note that tissue-specific gene list construction option is available only to the HumanPSD™ database subscription holders.