

TRANSFAC® release 2020.3

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

- **Extended range of species for Match / FMatch upload**

Genomic intervals in BED format from the species *Drosophila melanogaster*, rhesus macaque (*Macaca mulatta*), and pig (*Sus scrofa*) can now be uploaded as input for the analysis tools Match, Fmatch, and Composite model search, allowing to search for (overrepresented) transcription factor binding sites or binding site pairs in the promoter sequences of these species.

- **Additional files in the TRANSFAC® and associated TRANSPro™ download packages**

For TRANSFAC files in JSON format, there is a short guide and example scripts how to import them into a PostgreSQL database and have the advantage to join and query data from separate files.

The TRANSPro package now contains promoter data in Gene Transfer Format (GTF) for the species human, mouse, rat, pig, rhesus macaque, *Drosophila melanogaster*, and *Arabidopsis thaliana*.

- **Enhanced human SNP content**

The 2020 dbSNP release 154 data for human has been integrated and increases the number of SNPs mapped to human promoter sequences by more than 10,000,000 new single nucleotide variations compared to the previously used version 151.

- **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 100.