

## TRANSFAC® release 2018.1

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

- **Increased transcription factor coverage for 38 mammalian species**

Based on the recently published new version of [TFClass](#), the Classification of Transcription Factors in Mammalia, more than 23,000 factor entries from 38 species such as sheep, dog, or macaque have been added and linked to their human orthologs via the transcription factor classification.

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

76 new human transcription factor binding site ChIP-Seq experiments released by the [ENCODE phase 3 project](#) between June 2017 and September 2017 have been integrated. To increase mouse data coverage, 140 TFBS ChIP-Seq data sets, also from ENCODE, have been added. In total, the experiments comprise 4,893,699 fragments bound by 94 distinct transcription factors, of which 53 factors were not yet covered by ChIP-Seq data. For 128 of the sets, an existing positional weight matrix for the respective transcription factor was used together with the MATCH tool to predict altogether 3,215,450 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.

- **JASPAR 2018 matrix library integration**

New position frequency matrices from the [JASPAR 2018](#) release either added as matrix entries (57 cases) or hyperlinked to existing counterparts in the TRANSFAC matrix library.

- **Ensembl version update**

Genomic information for genes, promoters, and ChIP fragments for the species human, mouse, rat, macaque, and Arabidopsis is now based on Ensembl release 90.



## TRANSFAC® Statistics

### TRANSFAC 2018.1

Factors	46,961
miRNAs	1,279
DNA Sites	49,945
mRNA Sites	21,547
Factor-DNA Site Links	67,618
miRNA-mRNA Site Links	57,769
Genes	82,975
ChIP TFBS	32,124,271
Dnase Hypersensitivity Sites	15,376,241
Histone Modification Fragments	1,071,173
DNA Methylation Fragments	51,926
Matrices	7,454
References	36,944

### TRANSPro 2018.1

Promoter Sequences	
Human	84,900
Mouse	70,194
Rat	33,607
Macaca mulatta	29,108
Arabidopsis	29,221
Rice	29,456
Soybean	68,769

### TRANSCompel 2018.1

Composite Elements	593
Genes	402
Evidence Codes	2,181
References	661