geneXplain platform release 4.8

The geneXplain platform toolbox for bioinformatics data analysis contains these new functional features in the current release:

- **New analysis methods**

Eight new methods of data analysis are now available at the geneXplain platform. These methods comprise the following options:

Site analysis:

 **Construct composite modules on track (correlation)** – method predicts composite module using the result of the "Site search on gene set" analysis.

 **Cluster track** – method clusters sites in a track, what is useful for merging of closely spaced sites into one big cluster.

 **Compute profile thresholds** – method computes profile thresholds minimizing either false negatives(minFN) or false positive(minFP) on the random DNA sequence.

 **Create miRNA promoters** – method extracts miRNA promoters from mirprom database for a given list of miRNAs

 **Get transcripts track** – method extracts track from a database by a transcript ID

 **Recalculate composite module score on new track** – method takes best composite model from the given CMA result and calculates its scores on all sites of a given track.

 **Continue CMA** – method continues prediction of composite module using results of the previous prediction as a start point. Prediction parameters are customizable.

Statistics:

 **Table Imputation** – method replaces missing data in the given input table with row means.

- **New HTML report for site search analysis**

You can now create a summary of your site search analysis including visualization of input promoters together with identified enriched transcription factor binding sites (TFBSs) in HTML format, which can be exported to your local computer. These results can be easily used for presentations or publications.

REPORT

Data analysis is done with the geneXplain platform release 4.8
Project: data/Projects/Neetu
Date: Mon Jan 22 2018 14:49:32 GMT+0100 (CET)

Workflow: Identify enriched motifs in promoters (TRANSFAC(R))

Workflow path: analyses/Workflows/TRANSFAC/Identify enriched motifs in promoters (TRANSFAC(R))

Enriched TFBSs

In this study, 440 transcription factor binding sites were identified as enriched, with filter >1 by TFBS enrichment fold.

Folder: Sites
Number of rows: 440
Complete name: data/Examples/Brain Tumor GSE1825, Affymetrix HG-U133A microarray/Data/Ewing Family Tumor versus Neu...

TRANSFAC ID	Matrix logo	Transcription factor	Yes-No ratio	P-value
V$ZFP46_01		Zfp46	1.3350	1.9686e-35
V$ZNF436_01		ZNF436	1.3264	1.4705e-32
V$LRF_Q3		LRF	1.2942	4.9723e-26
V$CTCF_16		CTCF	1.3048	1.0631e-24
V$EGR1_06		Egr1	1.4378	1.7720e-43

Figure 1: HTML report with summary of all identified enriched transcription factor binding sites and the name of the corresponding transcription factor.

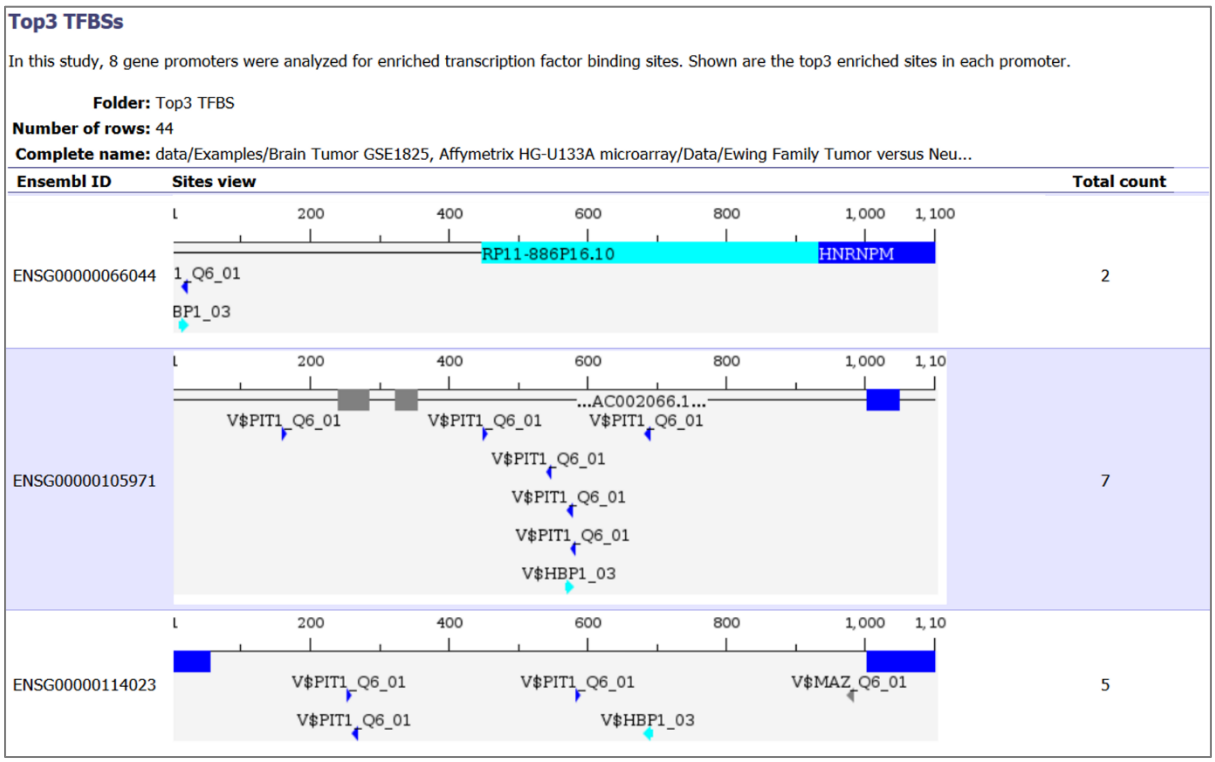

Figure 2: HTML report contains information and visualization for the most three enriched transcription factor binding sites (example here: V$PITI, V$HBP1 and V$MAZ) in all 44 input promoters.

- **New toolbar buttons**

Yes, yes, yes! The geneXplain platform now is not only one of the smartest tools in bioinformatics, it has also become one of the prettiest ones! Check out our new toolbar icons which will lead you to remarkable results in your research simply by a couple of clicks.

- **Integration with updated TRANSFAC®, TRANSPATH® and HumanPSD™ databases in release 2018.1**

The TRANSFAC® database of transcription factors, their genomic binding sites and DNA-binding motifs (PWMs), TRANSPATH® database of mammalian signal transduction and metabolic pathways and Human Proteome Survey Database (HumanPSD™) with focus on human proteins as disease biomarkers and drug targets in their 2018.1 release versions are currently integrated with the geneXplain platform.