

TRANSFAC Getting Started Guide

Version 2016.3

TRANSFAC is a database containing published data on eukaryotic transcription factors and miRNAs, their experimentally-proven binding sites, and regulated genes. The extensive compilation of binding sites forms the basis of derived positional weight matrices which can be used with the included Match tool to search DNA sequences for putative transcription factor binding sites – a method that is crucial for filling in gaps in the published literature. Integrated transcription factor ChIP-Seq, DNase hypersensitivity and histone methylated intervals from the ENCODE project complement the curated and predicted binding site data.

This guide will walk you through some of the most common uses of TRANSFAC. For additional help, please contact our support team at <mailto:support@genexplain.com>

Quick links to sections covered:

[Identify experimentally demonstrated regulators for a gene of interest](#)
[Visualize a gene's promoter and associated regulatory features](#)
[Identify predicted binding sites for a gene of interest](#)
[Identify experimentally demonstrated targets for a transcription factor of interest](#)
[Upload your own sequence and identify the predicted binding sites within](#)
[Import your own custom matrix for use in binding site predictions](#)
[Compare your own custom matrix against TRANSFAC's matrix library](#)
[Analyze a microarray or gene-based RNA-seq data set](#)
[Analyze a ChIP-seq data set](#)
[Analyze a transcript-based RNA-seq data set](#)

Identify experimentally demonstrated regulators for a gene of interest

TRANSFAC Locus Reports provide information about genes, including the transcription factors which have been experimentally demonstrated to bind to the gene's sequence.

To navigate to the Locus Report for a gene of interest:

- Select the "Genes and proteins" radio button
- Select the "Name" radio button
- Enter the name of your gene in the search box (PTEN in this example)
- Click the "search" button

geneXplain Welcome to TRANSFAC® | logout

search tools my data Home taskbar help

BIOBASE search search

▼ Hide search options

Limit search to

☒ Genes and proteins ☐ Transcription factors ☐ Variants

☐ miRNAs ☐ Matrices ☐ Diseases

☐ Pathways ☐ Drugs

Search genes and proteins by

☒ Name ☐ Identifier ☐ BLASTp ☐ TF binding site

Search which species?

Search all species

► Upload a list of genes or proteins in bulk and search for them

Genes and proteins search help

View statistics for: TRANSFAC 2016.2

example searches:

name BRCA1 (optionally specify Human)

identifier 672 (specify EntrezGene)

BLASTp SKAVGIDLGT TYSCVAHFAN DRVDIANDQ
GNRTTPSFVA FTDTERLIGD AAKNQAA...

TF site R24252 or HS\$TLR2_02

When you search Genes and proteins, you have the option to search by Name, Identifier, BLASTp or TF binding site.

Choose **Name** to limit your search to names and synonyms of genes, proteins, and complexes.

Choose **Identifier** to limit your search to BIOBASE and 3rd party identifiers for genes and proteins. Specify the identifier type for a faster and more accurate search.

Choose **BLASTp** to search using up to five fasta-formatted sequences.

Choose **TF binding site** to search for genes linked to TRANSFAC binding sites.

Choose **Upload a list of genes or proteins in bulk** to search by multiple names or identifiers.

The list of matched genes is returned:

Genes and proteins
11 of 11 total (All species)

Select results and view as:

Save these results Export these results Pathfinder Ontology Match FASTA Profiles Binding factors for gene Search

First 1 Last

Mark all on page Filter Hits on page 25

#	Name	Species/Taxon	Description
<input type="checkbox"/>	PTEN	Human	Phosphatase and tensin homolog, inhibits PKB signaling, phosphatidylinositol-3, 4, 5-trisphosphate production, cytokine production, and cell proliferation, downregulated in medulloblastoma, cutaneous melanoma, bladder and lung cancer, and ganglioneuroma
<input type="checkbox"/>	Pten	Mouse	Phosphatase and tensin homolog, inhibits PKB signaling, phosphatidylinositol-3, 4, 5-trisphosphate production, cytokine production, and cell proliferation; human PTEN correlates with medulloblastoma, melanoma, bladder and lung cancer, and ganglioneuroma
<input type="checkbox"/>	Pten	Rat	Phosphatase and tensin homolog, inhibits PKB signaling, phosphatidylinositol-3, 4, 5-trisphosphate production, cytokine production, and cell proliferation; human PTEN correlates with medulloblastoma, melanoma, bladder and lung cancer, and ganglioneuroma

Click the human “PTEN” link to navigate to the Locus Report.

Scroll down the page or use the “Table of Contents” section to jump to the “Transcriptional Regulation” section:

A summary of all curated, experimental binding sites is given in the table. For more detailed information about a specific binding site, click the link in the “Identifier” column to navigate to the Site Report which contains detailed supporting information about the method(s) used to measure binding, the experimental system and more, including links to the abstract of the reference from which the information was curated.

Locus Report - Human
PTEN (PTEN)

[logout](#)
[help](#)

Transcriptional Regulation what is this?

Regulation of PTEN gene expression

Predicted promoter sequences : Match →

Best supported : [PM000631593](#)

All promoters for the gene : [PM000631593](#), [PM000631594](#), [PM000631595](#)

Viewing nucleotides -54 to 53

* Note: Only binding sites whose location is relative to the TSS are graphically displayed in this graphical display.

Transcription factor binding sites within the PTEN gene (31 entries)

Show 5 entries

Search:

Identifier	Relative Location	Genomic Location	Binding Factor(s)	DNA Binding Reaction	Effect
HS\$PTEN_10	-2200 to -2167 *	Chr10 87862269 87862302 +	Max(m.s.)	Max(m.s.) --> PTEN(h)	DNA binding
HS\$PTEN_09	-2200 to -2167 *	Chr10 87862269 87862302 +	USF1(h):USF1(h)	USF1(h):USF1(h) --> PTEN(h)	DNA binding
HS\$PTEN_09	-2200 to -2167 *	Chr10 87862269 87862302 +	USF2(h):USF2(h)	USF2(h):USF2(h) --> PTEN(h)	DNA binding
HS\$PTEN_09	-2200 to -2167 *	Chr10 87862269 87862302 +	USF2(h):USF1(h)	USF2(h):USF1(h) --> PTEN(h)	DNA binding
HS\$PTEN_09	-2200 to -2167 *	Chr10 87862269 87862302 +	(usf1(m.s.))2	(usf1(m.s.))2 --> PTEN(h)	DNA binding

Showing 1 to 5 of 31 entries

[First](#)
[Previous](#)
[1](#)
[2](#)
[3](#)
[4](#)
[5](#)
[Next](#)
[Last](#)

Table of Contents

- Introduction**
 - Description
 - Synonyms
- Functional Attributes**
 - Add a subscription to PROTEOME and this report will display detailed information about:
- Gene Ontology**
 - Molecular function
 - Biological process
 - Cellular component
- Expression**
 - Tissue expression
- Transcriptional Regulation**
 - Regulation of PTEN gene expression
 - Regulation of PTEN mRNA expression
- RNA Features**
 - Overview of RNA sequence
- Protein Features**
 - Overview of protein sequence and structure

Visualize a gene's promoter and associated regulatory features

TRANSFAC contains promoter sequences for human, mouse, rat, Arabidopsis, rice and soybean species. Promoter sequences are 11,000 nucleotides in length, spanning from 10,000 nucleotides upstream of the transcription start site (TSS) to 1,000 nucleotides downstream of a virtual TSS that is calculated based on clusters of experimental TSSs. An individual gene may have more than one associated promoter. When this is the case one of the promoters will be selected as the best-supported, which indicates that it is the promoter with the greatest fraction of clustered TSSs.

There are two ways to graphically visualize the promoter sequence plus associated regulatory features which have been mapped to the promoter, including experimentally demonstrated factor binding sites.

The first way is to navigate to the "Transcriptional Regulation" section of the desired gene (see [Identify experimentally demonstrated regulators for a gene of interest](#)) and then click one of the "predicted promoter sequences" links:

Transcriptional Regulation what is this?

Regulation of PTEN gene expression

Predicted promoter sequences: [Match](#)

Best supported: PM000631593

All promoters for the gene: PM000631593, PM000631594, PM000631595

The second way is to select the desired gene in the search results and perform a second search for its associated promoter(s):

Genes and proteins
11 of 11 total (All species)

Select results and view as: [Pathfinder](#) [Ontology](#) [Match](#) [FASTA](#) [Profiles](#)

[Save these results](#) [Export these results](#)

First [1](#) Last

☐ Mark all on page [Filter](#)

#	Name	Species/Taxon	Description
<input checked="" type="checkbox"/>	PTEN	Human	Phosphatase and tensin homolog, inhibits PKB signaling, phosphatidylinositol-3-OH kinase, and cell proliferation, downregulated in medulloblastoma, cutaneous melanoma, and prostate cancer
<input type="checkbox"/>	Pten	Mouse	Phosphatase and tensin homolog, inhibits PKB signaling, phosphatidylinositol-3-OH kinase, and cell proliferation; human PTEN correlates with medulloblastoma, melanoma, bladder and lung cancer, and ganglioneuroma

Hits on page 25

Binding factors for gene [Search](#)

- Gene Regulation
 - Binding factors for gene
 - Gene bound by factor
 - Interacting factors
 - Promoter
 - TF binding matrix
- mRNA Regulation
 - Regulating miRNAs

Once the promoter search results load, click the link in the “Accession” column to open the Promoter Report. The Promoter Report provides a graphical view of the sequence:

BIOBASE Promoter Report - Human PTEN (TSS 10:87864163) [logout](#) [help](#)

[Table of Contents](#)

Sequence View what is this?

Sequence [FASTA](#)

GC-rich

83

Churchill

MECP2

SP1

CGGGCGGCGGCTGGCACTCCAGGGAACCGGGCGGGTTTAAACCTCCCGTGGCGCGCGCGCGCAACCCCGTGGCGCGGGCTCCGGAGGCGCGCGCGGAGGCAGCGGTCGGAGGATTATTCGTCT

87864100 87864110 87864120 87864130 87864140 87864150 87864160 87864170 87864180 87864190 87864200 87864210

Viewing nucleotides 87864094 to 87864221, [shows regions of conservation between placental species](#)

Displayed features:

<input checked="" type="checkbox"/> TRANSFAC experimental binding site	<input type="checkbox"/> DNase hypersensitivity sites	<input checked="" type="checkbox"/> SNP
<input type="checkbox"/> ChIP fragments	<input checked="" type="checkbox"/> Predicted transcription factor binding site	<input checked="" type="checkbox"/> Repeat
<input type="checkbox"/> Histone modification fragments	<input checked="" type="checkbox"/> CpG island	<input checked="" type="checkbox"/> TSS

Use the “Displayed features” options shown above to add or remove different types of features. The red orientation point and flanking yellow region in the frequency bar below the nucleotide sequence indicates the portion of the sequence that is displayed. Drag the red orientation point along the frequency bar to inspect different regions of the promoter at the nucleotide level.

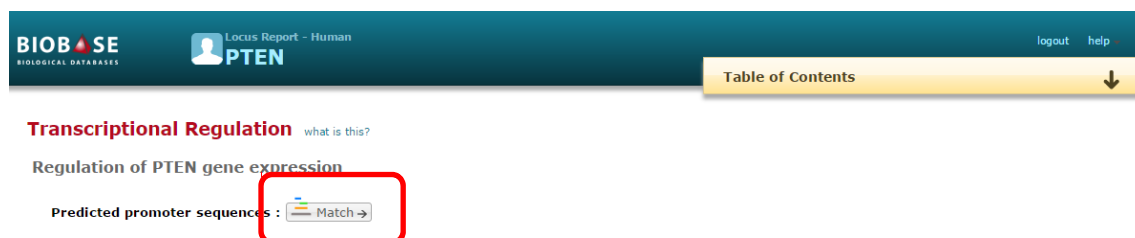
The graphic is followed by a tabular view of the different features

Identify predicted binding sites for a gene of interest

While the Locus Report provides an overview of all curated, experimental binding sites for a gene (see [Identify experimentally demonstrated regulators for a gene of interest](#)) and the Promoter Report provides a graphical view of those experimental binding site plus additional regulatory features located within the gene’s promoter (see [Visualize a gene’s promoter and associated regulatory features](#)), the Predict TF binding sites tool and the Match algorithm that it employs enables you to scan the gene’s promoter for additional predicted binding sites using TRANSFAC’s comprehensive positional weight matrix library.

There are two ways to interrogate the gene’s promoter.

The first way is to navigate to the “Transcriptional Regulation” section of the Locus Report (see [Identify experimentally demonstrated regulators for a gene of interest](#)) and then click the Match icon to launch the Match algorithm:



The second way is to perform a search for the gene of interest (see [Identify experimentally demonstrated regulators for a gene of interest](#)), and then launch the Match algorithm directly from the search results by selecting the gene of interest and clicking the Match icon:



In both cases an intermediate window will open where you are asked to:

- Select the analysis method (use the default “Match”)
- Select the promoter to be analyzed (use the default “best supported promoter” if you want to analyze the promoter with the greatest number of clustered TSSs, or change to “all promoters” if you would like to analyze all promoters for comparison)
- Specify the region of the promoter to analyze (use the default “-10,000 bp to +1,000 bp” to analyze the entire promoter, or select a smaller region within those bounds)
- Specify the data version (use the default, current version)
- Specify the profile (group of matrices) to be used (use the default “vertebrate_non_redundant_minFP” which contains representative matrices for distinctive binding motifs unless you are working with non-vertebrate species or want to limit your analysis to factors that are active in a particular tissue)
- Specify the cut-off criteria (use the default “minimize false positives” which is automatically assigned with the “vertebrate_non_redundant_minFP” profile)

The screenshot shows a web interface titled "Match : Predicting transcription factor binding sites". Below the title is a brief description: "The Match program uses positional weight matrices (PWMs) from the TRANSFAC matrix library to search the stored DNA sequences of the selected gene(s) for potential transcription factor binding sites. [Learn more.](#)". The interface contains several input fields and a "Submit" button. The fields are: "Select analysis method:" with a dropdown menu showing "Match - search for TF binding sites"; "For the selected gene(s), analyze:" with a dropdown menu showing "best supported promoter"; "Within a nucleotide window of:" with two input boxes showing "-10000 bp(max)" and "+1000 bp(max)"; "Data version:" with a dropdown menu showing "2015.3"; "Using the profile (group of matrices):" with a dropdown menu showing "vertebrate_non_redundant_minFP"; a checkbox labeled "Use only high quality matrices" which is checked; and "With cut-off criteria:" with a dropdown menu showing "minimize false positives". A "Submit" button is located at the bottom left of the form.

Match : Predicting transcription factor binding sites

The Match program uses positional weight matrices (PWMs) from the TRANSFAC matrix library to search the stored DNA sequences of the selected gene(s) for potential transcription factor binding sites. [Learn more.](#)

? Select analysis method: Match - search for TF binding sites ▼

? For the selected gene(s), analyze: best supported promoter ▼

? Within a nucleotide window of: -10000 bp(max) to +1000 bp(max)

? Data version: 2015.3 ▼

? Using the profile (group of matrices): vertebrate_non_redundant_minFP ▼

☒ Use only high quality matrices

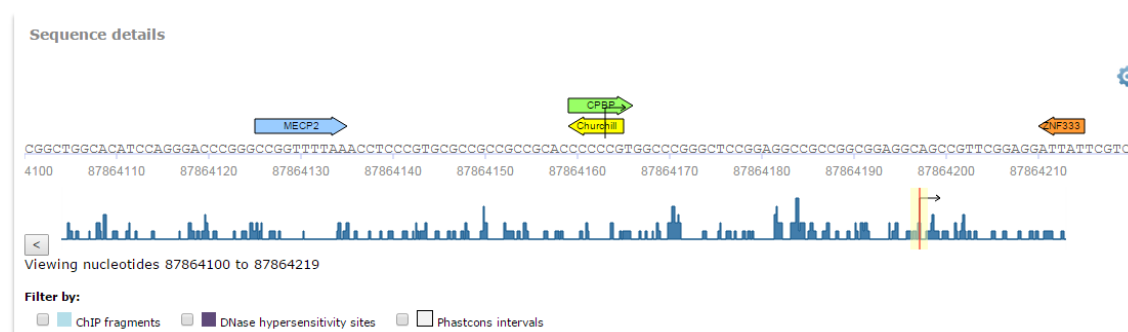
? With cut-off criteria: minimize false positives ▼

Submit

When you click the Submit button your analysis will be forwarded to the taskbar. Once the analysis completes, the Match Analysis Report will automatically load:



The “Matrix summary” section provides a tabular overview of all matrices that were found to have binding sites within the promoter. Scrolling down to the “Sequence details” section a graphical view of the results is provided:



A frequency bar displays the distribution of sites along the entire sequence, with the red orientation point and flanking yellow highlighting indicating the portion of the sequence that is displayed at the nucleotide level above. Slide the red orientation point along the frequency bar to view other portions of the sequence in detail.

Optionally use the ChIP fragments, DNase hypersensitivity sites and Phastcons intervals filters to apply additional biological context to your results. Activating these filters removes any sites that do not overlap with regions of sequence that are suggested to be available to the transcriptional machinery based on being bound by transcription factors (ChIP fragments), being accessible to DNase enzymes (DNase hypersensitivity) or conserved between human and mouse (Phastcons).

Please note that if “all promoters” was selected for the analysis it will be necessary to click the Sequence name in the “Sequence summary” section to display the nucleotide level view.

The table below the graphical view lists all identified binding sites with supporting details:

First

◀

1

2

3

4

5

▶

Last

☐ Mark all on page
 Filter
 Export table
 Hits on page 25 ▼

Showing 25 of 312 entries

#	Matrix	Factor name	Position (strand) ±	Core score	Matrix score	Sequence
<input type="checkbox"/>	V\$YY1_Q6_03	YY1	87854376 (+)	1.000	1.000	CCATTtt
<input type="checkbox"/>	V\$DRI1_01	DRI1	87854613 (-)	1.000	1.000	TTAATt
<input type="checkbox"/>	V\$HMG1Y_Q3	HMG1Y	87854623 (+)	1.000	0.947	tttttAATTtttc
<input type="checkbox"/>	V\$DRI1_01	DRI1	87854626 (-)	1.000	1.000	TTAATt
<input type="checkbox"/>	V\$IK_Q5_01	Ikaros	87854696 (-)	1.000	1.000	ccTCCCA

Identify experimentally demonstrated targets for a transcription factor of interest

There are two ways to identify the experimentally determined targets for a transcription factor. The first way is to navigate to the Locus Report for the factor and view the information provided in the “Transcriptional Regulation” section.

To navigate to the Locus Report for a transcription factor of interest:

- Select the “Transcription factors” radio button
- Select the “Name” radio button
- Enter the name of your gene in the search box (HNF1A in this example)
- Click the “search” button

example searches:

name	FOXA1 (optionally specify Human)
identifier	3169 (specify EntrezGene)
BLASTp	MLGTVKMEGH ETSDWNSYYA DTQEA YSSVP VSNMNSGLGS MNSMNTY...
TF site	R24252 or HS\$TLR2_02

When you search Transcription factors, you have the option to search by Name, Identifier, BLASTp or TF binding site.

Choose **Name** to limit your search to names and synonyms of transcription factors.

Choose **Identifier** to limit your search to BIOBASE and 3rd party identifiers for transcription factors. Specify the identifier type for a faster and more accurate search.

Choose **BLASTp** to search using up to five fasta-formatted sequences.

Choose **TF binding site** to search for transcription factors linked to TRANSFAC binding sites.

When you browse Transcription factors, you have the option to view or export prepared lists of transcription factors directly.

Choose **ChIP-chip** and **ChIP-seq** experiments to access reports presenting high-throughput data sets.

The list of matched factors is returned:

from experiments using closely related mouse and rat Hnf1a factors will only be covered on the respective mouse and rat Locus Reports.

The second way to identify the experimentally determined binding sites for a transcription factor of interest avoids this separation of information by species. In this case perform the same “Transcription factors” search for HNF1A. Then when the results are returned, select the human, mouse and rat entries, select the “Gene bound by factor” option in the pulldown menu in the search results header and click the Search button next to the pulldown menu:

Transcription factors
7 of 7 total

Select results and view as: [Pathfinder](#) [Ontology](#) [Match](#) [FASTA](#) [Profiles](#)

[Save these results](#) [Export these results](#)

First [1](#) [2](#) [3](#) [4](#) [5](#) [Last](#)

☐ Mark all on page [Filter](#)

#	Name	Species/Taxon	Description
<input checked="" type="checkbox"/>	HNF1A	Human	HNF1 homeobox A, a transcriptional activator that functions expressed in kidney and liver cancers; gene mutation causes and Fanconi syndrome gene
<input checked="" type="checkbox"/>	Hnf1a	Mouse	HNF1 homeobox A, a transcriptional activator that functions in insulin secretion and fatty acid transport; genetic knockout is associated with phenylketonuria and Fanconi syndrome, human HNF1A is associated with diabetes and kidney and liver cancers gene
<input checked="" type="checkbox"/>	Hnf1a	Rat	Transcription factor 1, a transcriptional activator that functions in insulin secretion and fatty acid transport; human HNF1A is associated with diabetes and kidney and liver cancers, mouse Hnf1a is associated with phenylketonuria and Fanconi syndrome gene

Hits on page [25](#)

All three factors are used as input to the search. The results are a combined list of the genes that each have been shown to bind experimentally:

Gene bound by factor
25 of 129 total

Select results and view as: [Pathfinder](#) [Ontology](#) [Match](#) [FASTA](#) [Profiles](#)

[Save these results](#) [Export these results](#)

First [1](#) [2](#) [3](#) [4](#) [5](#) [Last](#)

☐ Mark all on page [Filter](#)

#	Search Term	Name	Species/Taxon	Description
<input type="checkbox"/>	GN000004288 - HNF1A GN000002682 - Hnf1a	SERPINA1	Human	Serpin peptidase inhibitor clade A member 1, acts in acute-phase response, aberrantly expressed in rheumatoid arthritis, asthma, pancreatitis, Crohn disease, and Alzheimer disease; gene mutation causes alpha 1-Antitrypsin deficiency and cystic fibrosis
<input type="checkbox"/>	GN000004288 - HNF1A GN000009226 - Hnf1a GN000002682 - Hnf1a	Fgb	Rat	Fibrinogen beta chain, a structural molecule that acts in cell adhesion and acute-phase response, may act in blood coagulation; human FGB is associated with afibrinogenemia, venous thrombosis, coronary artery disease, and recurrent pregnancy loss
<input type="checkbox"/>	GN000004288 - HNF1A GN000009226 - Hnf1a GN000002682 - Hnf1a	HNF4A	Human	Hepatocyte nuclear factor 4 alpha, a transcriptional coactivator that acts in hepatocyte differentiation and glucose metabolism, downregulated in Crohn disease; mRNA is overexpressed in hepatocellular carcinoma and SNP correlates with type 2 diabetes
<input type="checkbox"/>	GN000004288 - HNF1A	SLCO1B1	Human	Solute carrier organic anion transporter family member 1B1, a transmembrane transporter that acts in transport of drug, thyroid hormone, prostaglandin, and bile acid and bile salt

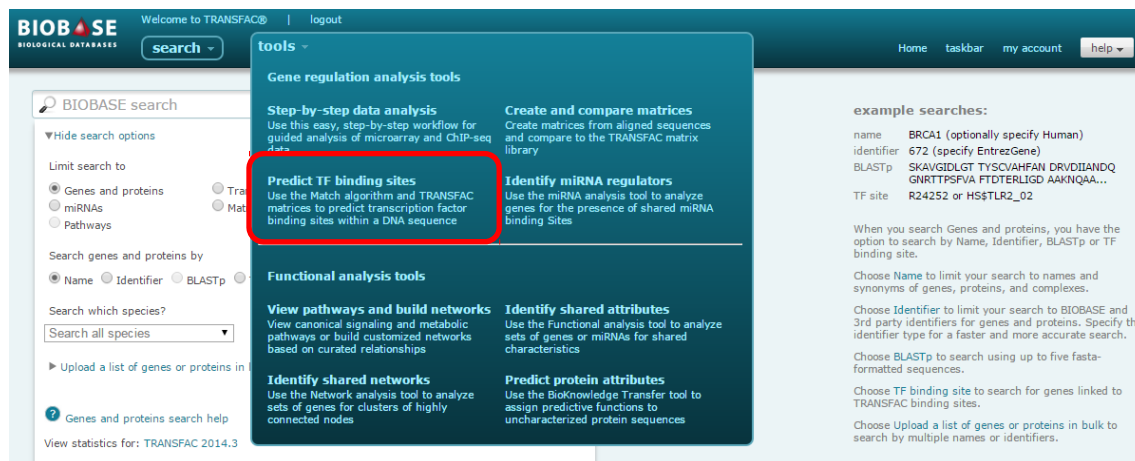
Hits on page [25](#)

Following the link to the Locus Report of the transcription factor or the gene and navigating to the “Transcriptional Regulation” section will allow you to access the same experimental details described above.

Upload your own sequence and identify the predicted binding sites within

TRANSFAC provides preloaded promoter sequences for human, mouse, rat, Arabidopsis, rice and soybean genes but you may upload your own sequence or set of sequences for these or any other species and then analyze the sequence(s) for predicted transcription factor binding sites.

Select the “Predict TF binding sites” option from the tools menu:



When the tool loads, click the “I am analyzing DNA sequences” radio button followed by the “Upload a new sequence” radio button:

Predict Transcription Factor Binding Sites [what is this?](#)

- ☒ I am analyzing DNA sequences
- ☐ I am analyzing a gene or miRNA set

- ☐ Example sequence
- ☐ Select a previously uploaded sequence:

3000_seqs_250_bp ▼

- ☒ Upload a new sequence

Name:

I am uploading

- ☒ DNA sequences
- ☐ Genomic intervals for automatic sequence retrieval

Copy and paste DNA sequences up to 10,000,000 nucleotides in total length in FASTA, EMBL, Genbank or RAW format

or

[reset form](#)

You will be presented with the option to upload your data as a set of DNA sequences in FASTA, EMBL, Genbank or RAW format or as a set of genomic intervals in .bed format (note that the genomic intervals option is only supported for human, mouse and rat species).

Once you have specified the sequence(s) to be uploaded, accept or adapt in the remaining parameters:

- Select the analysis method (use the default "Match")
- Specify the profile (group of matrices) to be used (use the default "vertebrate_non_redundant_minFP" which contains representative matrices for distinctive binding motifs unless you are working with non-vertebrate species or want to limit your analysis to factors that are active in a particular tissue)
- Keep the "Use default parameters" option enabled

Select analysis method:
 Match - search for TF binding sites ▼

Select a profile (group of matrices):
 vertebrate_non_redundant_minFP ▼

☒ Use default parameters

Data version: 2015.3 ▼

☒ Use only high-quality matrices

Set cut-offs: Minimize false positives ▼

[start search](#) [reset form](#)

When you click the start search button your analysis will be forwarded to the taskbar. Once the analysis completes, the Match Analysis Report will automatically load:

Match Analysis Report
[Return to Predict TF binding sites page](#)

[Save this report](#) [Export this report](#)

Analysis summary

Match analysis identified:

Total sequence length : 11001	Total number of sequences : 1
Total number of sites : 312	Number of sequences with sites : 1
Frequency of sites : 0.02836	Average number of sites per sequence: 312.00

[Show parameters used](#)

Matrix summary

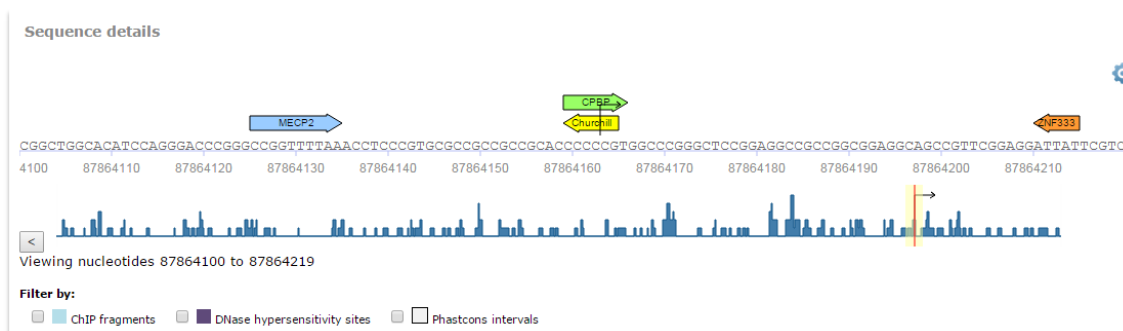
First ◀ 1 2 3 4 5 ▶ Last

☐ Mark all on page (selecting matrices will filter sequence summary) [Filter](#) [Profiles](#) [Export table](#) Hits on page 5 ▼

Showing 5 of 69 entries

#	Matrix	Factor	Consensus sequence	Classification	Category	Sites	Sequences	Sites/Sequences
<input type="checkbox"/>	V\$CPBP_Q6	CPBP		ZFC2H2	matrix compiled from individual genomic sites	28	1	28.00
<input type="checkbox"/>	V\$ZNF333_01	ZNF333		ZFC2H2	SELEX (CASTing, SAAB, TDA, Target detection)	27	1	27.00

The “Matrix summary” section provides a tabular overview of all matrices that were found to have binding sites within the uploaded sequence. Scrolling down to the “Sequence details” section a graphical view of the results is provided:




A frequency bar displays the distribution of sites along the entire sequence, with the red orientation point and flanking yellow highlighting indicating the portion of the sequence that is displayed at the nucleotide level above. Slide the red orientation point along the frequency bar to view other portions of the sequence in detail.

If you uploaded your sequence in genomic coordinates, you may optionally use the ChIP fragments, DNase hypersensitivity sites and Phastcons intervals filters to apply additional biological context to your results. Activating these filters removes any sites that do not overlap with regions of sequence that are suggested to be available to the transcriptional machinery based on being bound by transcription factors (ChIP fragments), being accessible to DNase enzymes (DNase hypersensitivity) or conserved between human and mouse (Phastcons).

Please note that if more than one sequences was submitted for analysis it will be necessary to click the Sequence name in the “Sequence summary” section to display the nucleotide level view.

The table below the graphical view lists all identified binding sites with supporting details:

First	◀	1	2	3	4	5	▶	Last
<input type="checkbox"/>	Mark all on page	Filter	Export table					Hits on page 25 ▼
Showing 25 of 312 entries								
#	Matrix	Factor name	Position (strand) 	Core score	Matrix score	Sequence		
<input type="checkbox"/>	V\$YY1_Q6_03	YY1	87854376 (+)	1.000	1.000	CCATTtt		
<input type="checkbox"/>	V\$DRI1_01	DRI1	87854613 (-)	1.000	1.000	TTAATt		
<input type="checkbox"/>	V\$HMG1Y_Q3	HMG1Y	87854623 (+)	1.000	0.947	tttttAATTTtttc		
<input type="checkbox"/>	V\$DRI1_01	DRI1	87854626 (-)	1.000	1.000	TTAATt		
<input type="checkbox"/>	V\$IK_Q5_01	Ikaros	87854696 (-)	1.000	1.000	ccTCCCA		

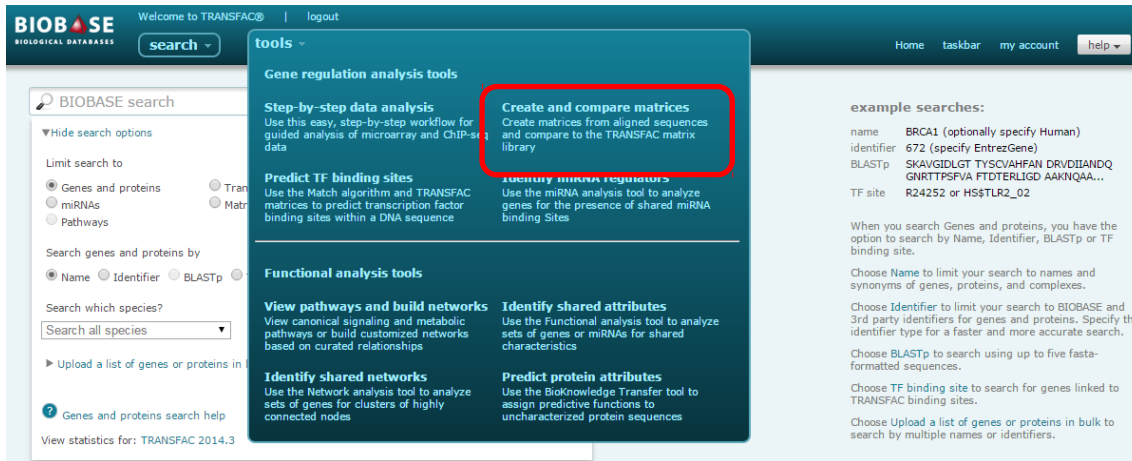
Your analysis will be preserved in the taskbar for 7 days. To preserve it indefinitely, click the “Save this report” link at the top of the report.

Import your own custom matrix for use in binding site predictions

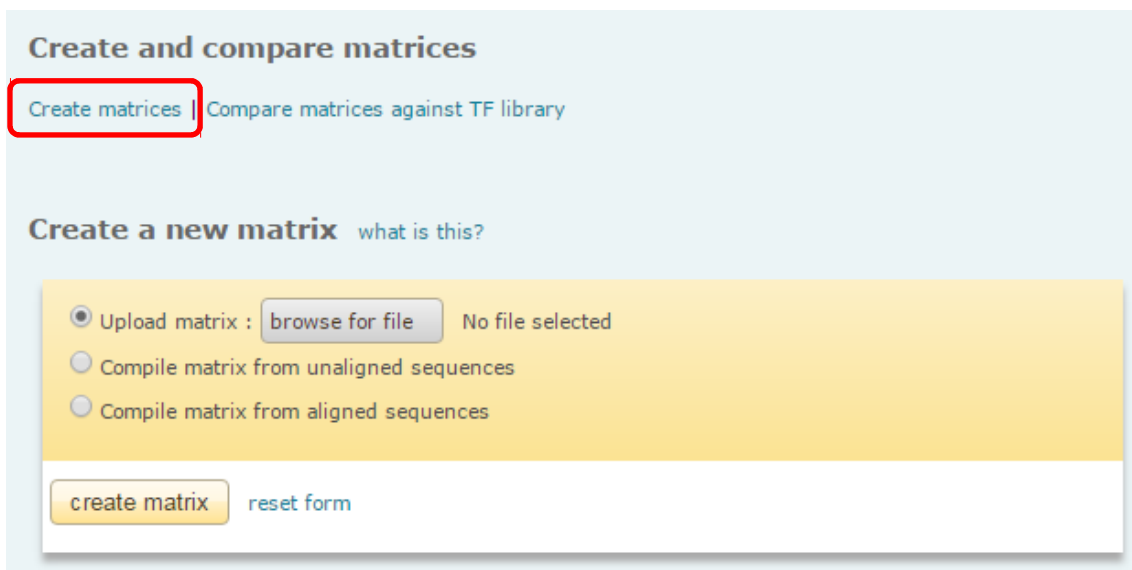
TRANSFAC's matrix library is based on the published literature and published data sets. If

you have your own matrix, or set of sequences that you would like to compile into a matrix, you can upload it into TRANSFAC which will make it available to use in all analyses.

Select the “Create and compare matrices” option from the tools menu:



When the tool loads, click the “Create matrices” option to refresh the view if needed:



You have the option to:

- Upload a matrix in TRANSFAC file format
- Compile a matrix from unaligned sequences
- Compile a matrix from aligned sequences

To compile a matrix from unaligned sequences, you need to provide a set of positive sequences (those that are expected to contain the motif) and a set of negative sequences (those that are not expected to contain the motif).

To compile a matrix from aligned sequences, you need to provide a minimum of five sequences that are expected to contain the motif and are of the exact same length.


Once you have chosen the relevant method and supplied the requested information, click the “create matrix” button. The consensus sequence and nucleotide position frequency is displayed:

Create matrices

Matrix name : Example

Date created : Sat Sep 19 15:59:15 2015

Identifier : V\$Example_30313



Nucleotide position frequency

	A	C	G	T	Consensus
1	3	0	0	2	W
2	5	0	0	0	A
3	0	0	1	4	T
4	0	5	0	0	C
5	0	1	4	0	G
6	5	0	0	0	A
7	0	5	0	0	C
8	2	2	1	0	M

Save matrix and specify cut-off values

(Note that this operation may take a few minutes)

Click the “Save matrix and specify cut-off values” button to complete the process of calculating the cut-offs. It may take a few minutes, but once the calculation is complete the cut-offs will be displayed and you will be notified that your matrix has been saved:

Cut-offs :					
		Core similarity	Matrix similarity		
Minimize false negative matches		0.748	0.840		
Minimize false positive matches		0.748	0.964		
Minimize sum of both error-rates		0.748	0.908		
Matrix quality : high					
False positive frequency when tolerating a false negative rate :					
FN	10%	30%	50%	70%	90%
MSS	(0.840)	(0.923)	(0.977)	(0.987)	(0.996)
FP frequency	1.455	0.118	0.014	0.007	0.006

In order to use the matrix in an analysis you will need to take the additional step of creating a profile that contains the matrix. From the help menu on the right hand side of the screen, click the “Create a profile” link:

Creating and comparing matrices

A positional weight matrix is a motif which describes the nucleotide frequency at each position of a set of sequences bound by a transcription factor. The Match tool uses matrices to search DNA sequences for predicted transcription factor binding sites. The TRANSFAC matrix library provides a large collection of matrices which have been compiled from individual genomic binding sites as well as ChIP Seq, SELEX and other types of experiments. Alternatively, you may use the tools provided here to create your own matrix and compare against the TRANSFAC matrix library.

Creating matrices: Use the [Create matrices](#) tool to create a custom matrix. Start by providing five or more aligned sequences in FASTA, ClustalW or Gibbs format. After specifying the cut-off values and saving the matrix it can be used to [Create a profile](#) for use in Match analysis.

Comparing matrices: Use the [Compare matrices](#) tool to compare a created matrix against the TRANSFAC matrix library to identify similar, characterized motifs.

When the Create profile tool loads you can quickly find your matrix by clicking the “Show only user matrices” option:

Select matrices to be included in a profile from the list below:

☐ Exclude low quality matrices ☒ Show only user matrices

Show 5 entries

Search:

Accession Number	Matrix Quality	Matrix Identifier	Matrix Name	Matrix Category	Matrix Classification	Sequence Logo
U00002	high	V\$CREB_1105_01	CREB_1105_01			
U00003	high	V\$Example_30313	Example			

Showing 1 to 2 of 2 entries (filtered from 5,505 total entries)

First Previous **1** Next Last

Click on your matrix to select it and then click the “Select matrices” button. It will be

displayed below with the calculated cut-offs:

Select matrices to include in the profile.		FP frequency (per 1kb nucleotides) when false negative rate of:					Proceed to cut-off selection	
<input checked="" type="checkbox"/> Mark all								
Matrix	Quality	10%	30%	50%	70%	90%		
<input checked="" type="checkbox"/> Example V\$Example_30313 U00003	high	1.455	0.118	0.014	0.007	0.006		

Click the “Proceed to cut-off selection” button. A new window loads where you are able to specify a name for your profile as well as set the default cut-off to be used in analyses:

Select matrices and cut-offs for your profile:

Please enter a name for your profile:

Select the following cut-off for all matrices: [what is this?](#)

☐ Unmark low quality matrices

Core / Matrix similarity (and FP frequency) at the selected cut-offs

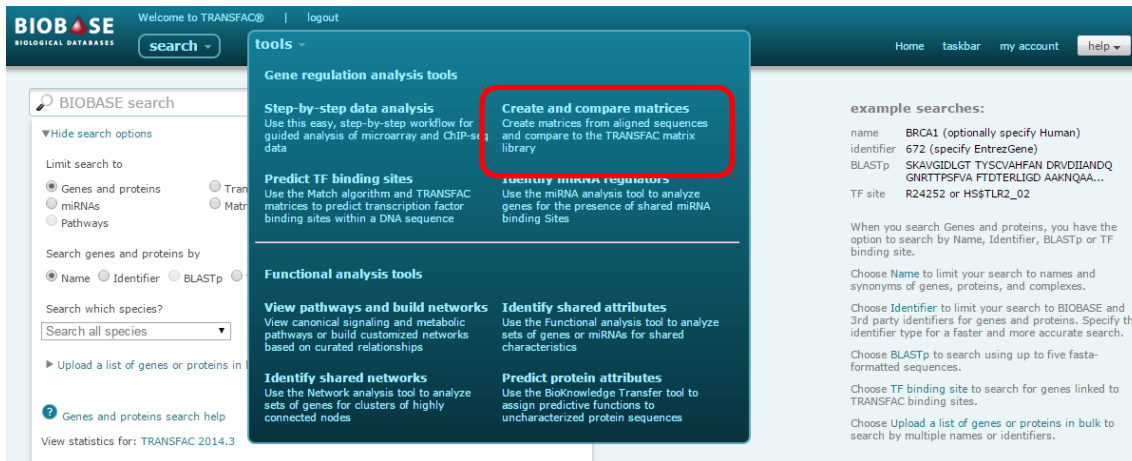
Matrix	Quality	FN10 / minFN	FN30	FN50	FN70	FN90	minFP	minSUM	Core / Matrix similarity
<input checked="" type="checkbox"/> Example V\$Example_30313 U00003	high	<input type="radio"/> 0.748 / 0.840 : 1.455	<input type="radio"/> 0.748 / 0.923 : 0.118	<input type="radio"/> 0.748 / 0.977 : 0.014	<input type="radio"/> 0.748 / 0.987 : 0.007	<input type="radio"/> 0.748 / 0.996 : 0.006	<input checked="" type="radio"/> 0.748 / 0.964	<input type="radio"/> 0.748 / 0.908	<input type="text"/>

The suggested default is minFP. Click the “Save profile” button to save the profile and make it available in the matrices pull-down menu of all analyses.

Compare your own custom matrix against TRANFAC's matrix library

If you have uploaded a custom matrix (see [Import your own custom matrix for use in binding site predictions](#)) you can use the Compare matrices tool to investigate whether your matrix is similar to other matrices within the TRANFAC matrix library.

Select the “Create and compare matrices” option from the tools menu:



When the tool loads, click the “Compare matrices against TF library” option to refresh the view if needed:

The screenshot shows the 'Create and compare matrices' tool interface. The 'Compare matrices against TF library' option is highlighted with a red box. Below this, the 'Compare matrices against the TRANSFAC matrix library' section is visible. It includes a 'Select a matrix' section with three options: 'Use our example: (U00001;V\$CREB)', 'Select a previously uploaded matrix: Select one', and 'Upload a file: browse for file No file selected'. There is also a 'Data version' dropdown set to '2015.3' and a 'Limit comparison to:' section with checkboxes for 'Vertebrate', 'Plant', 'Insect', and 'Fungi'. At the bottom, there are 'compare matrices' and 'reset form' buttons.

Select your previously uploaded matrix from the pull-down menu, select the data version (use the default current version) and select whether you want to limit your comparison to vertebrate, plant, insect or fungal matrices. Click the “compare matrices” button. The screen will be refreshed and the results of the comparison will be provided in a table:

Result [what is this?](#)

Create and compare matrices

[Create matrices](#) | [Compare matrices against TF library](#)

Query Matrix	Transfac Id	Matrix accession	Matrix class	Alignment	P-value	FDR
U00001	V\$CREB_01	M00039	BZIP	View	1.1913E-10	5.2096E-7
U00001	V\$ATF_B	M00338	BZIP	View	6.6155E-8	1.2175E-4
U00001	V\$CREBP1_Q2	M00179	BZIP	View	8.3561E-8	1.2175E-4
U00001	V\$ATF2_Q6	M07312	BZIP	View	1.5165E-7	1.6576E-4
U00001	V\$CREB1_01	M02279	BZIP	View	4.8909E-7	3.8012E-4
U00001	V\$CREB_Q4	M00178	BZIP	View	5.2214E-7	3.8012E-4
U00001	V\$CREB1_Q6	M03544	BZIP	View	1.1678E-6	5.0606E-4
U00001	V\$CREB_Q4_01	M00917	BZIP	View	1.2472E-6	5.0606E-4
U00001	V\$ATF2_Q5	M01862	BZIP	View	1.2598E-6	5.0606E-4
U00001	V\$ATF1_03	M02738	BZIP	View	1.3240E-6	5.0606E-4

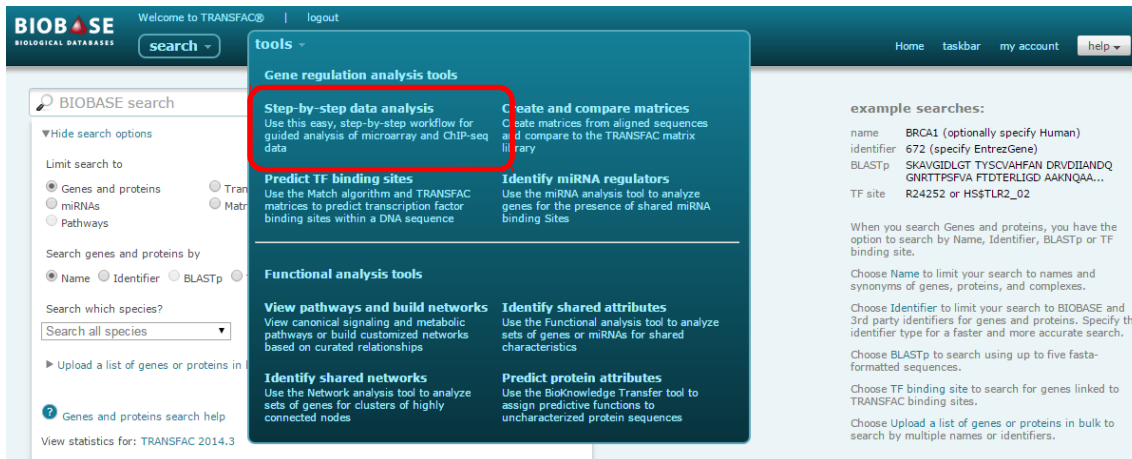
Up to ten best matching matrices will be listed. Click the “View” link in the Alignment column to see a graphical alignment of the motifs:



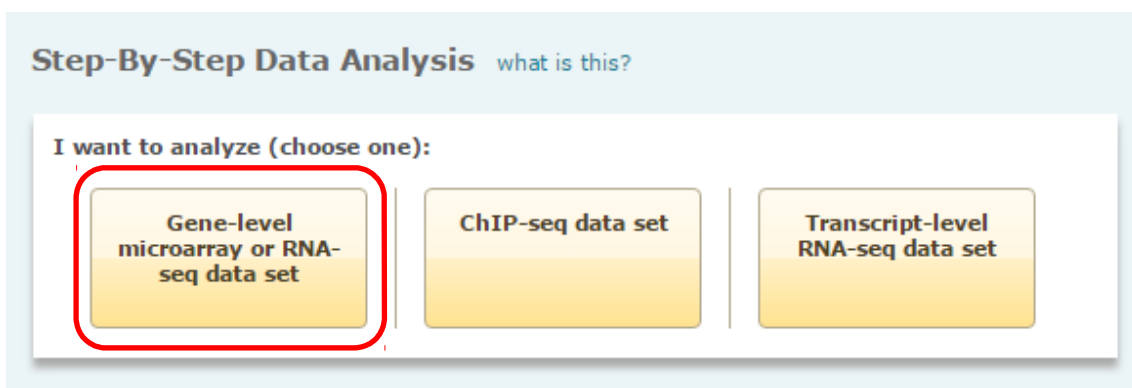
Analyze a microarray or gene-based RNA-seq data set

Use TRANSFAC’s step-by-step analysis workflow for gene-level microarray and RNA-seq data sets to identify the transcription factors that are most likely to be responsible for the differential expression pattern observed in your experimental data as compared to background data.

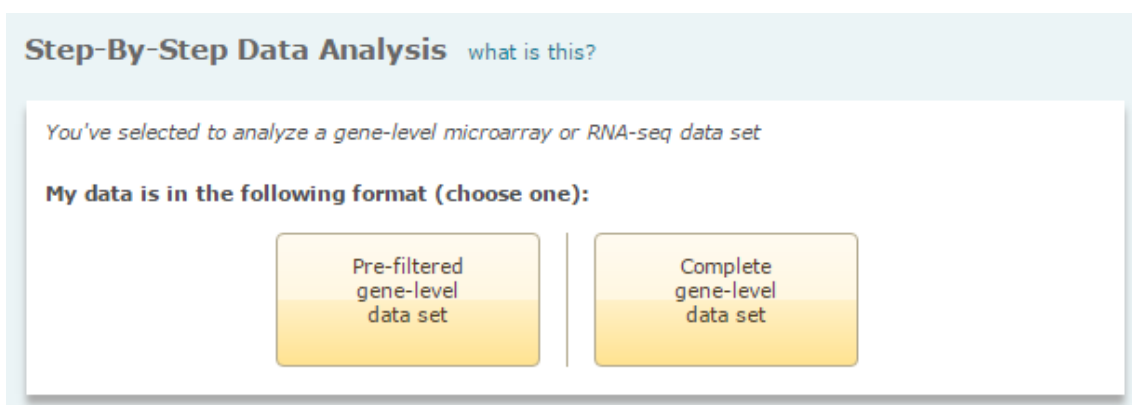
Select the “Step-by-step data analysis” options from the tools menu:



Then select the Gene-level microarray or RNA-seq data set option:



Specify whether your data is in the form of a pre-filtered list (only the subset of genes that were identified as up- or down-regulated is included) or whether you are uploading the complete data set including genes that showed no change in expression:



Upload your data set (in this example we are uploading the complete data set):

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a gene-level microarray or RNA-seq data set

Now select the complete gene-level microarray or RNA-seq data set to be analyzed:

☒ Upload a new data set

[browse for file](#)

☒ Names ☐ Identifiers

Which species?

Synonym handling?

File contains column headers : ☐ Yes ☒ No

[Next](#)

Note that once you've run at least one analysis you will have the additional option to select a previously uploaded data set instead of uploading a new data set.

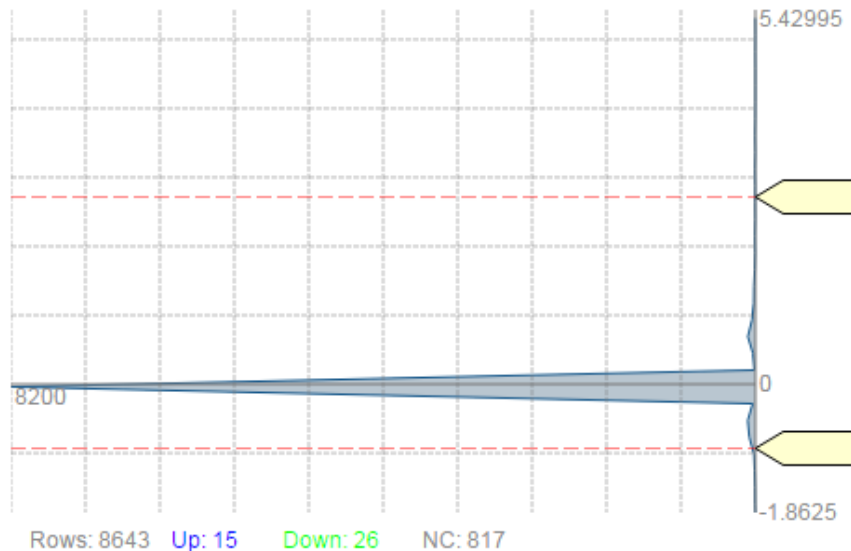
You will be given a preview of the list of matched genes and then will be forwarded to a screen where you will specify the criteria for selecting the up-regulated, down-regulated and non-changed gene sets:

Step-By-Step Data Analysis what is this?

You've selected to analyze a gene-level microarray or RNA-seq data set

You've selected the experimental data set: HUVEC_GSE2639_all_8764_log2FC

Extract Up/Down/Non-change



Select observation column:

☒ Select all upregulated elements: expression >

☐ Select all downregulated elements: expression <

Select non-changed elements: expression = +/-

☐ Analyze Up and Down regulated sets in parallel

In this example, we have chosen to select all up-regulated genes which show a log2 fold change of >1 as our experimental set and to select all genes which show a log2 fold change of 0 as the non-changed set.

Note that if your data is uploaded in the form of a pre-filtered list this step will automatically be skipped.

Next you will specify the background set to use:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a gene-level microarray or RNA-seq data set

You've selected the experimental data set: HUVEC_GSE2639_all_8764_log2FC

Now select the background data set to be used in the analysis :

- ☒ A non-change set from the experiment set
- ☐ Randomly generated gene set
- ☐ Previously uploaded gene set :

Housekeeping_human_90 ▼
- ☐ Upload a new background data set

Next

You have multiple options for the background set including:

- Using the non-change genes from the data set that you just uploaded, (the default, recommended option)
- Using a randomly generated gene set
- Using a previously uploaded gene set
- Uploading a new gene set

Note that if your data is in the form of a pre-filtered list you will not have the option to directly use the non-change genes from the data set as they were not included, you will only have the latter three options.

Next you will accept or change the default analysis parameters:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a gene-level microarray or RNA-seq data set

You've selected the experimental data set: HUVEC_GSE2639_all_8764_log2FC

You've selected the background data set : non-change set

Now select :

A profile (group of matrices):

Data version:

P-value threshold:

Specify the nucleotide region of the gene's promoter to be analyzed :

to A range of -10,000 to +1,000 bp is supported. [Learn more](#)

You will:


- Specify the profile (group of matrices) to be used (use the default “vertebrate_non_redundant_minFP” which contains representative matrices for distinctive binding motifs unless you are working with non-vertebrate species or want to limit your analysis to factors that are active in a particular tissue)
- Specify the data version (use the default, current version)
- Specify the P-value threshold (0.01 is the default)
- Specify the region of the gene's promoter to be considered in the analysis (a range of -10,000 bp to +1,000 bp relative to the TSS is supported, but the default range of -500 bp to +100 bp relative to the TSS is recommended for most analyses)

When you click the “Run FMatch” button your analysis will be forwarded to the taskbar. Once the analysis completes, the FMatch Analysis Report will automatically load.




The “Matrix summary” section provides a tabular overview of all matrices that were found to have significantly more binding sites within your experimental data set than were found in the background data set:

Matrix summary

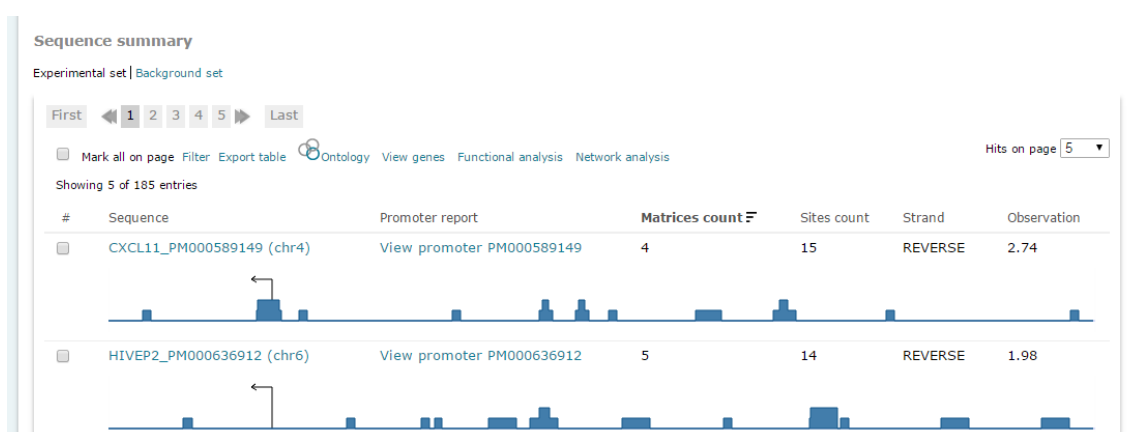
First ◀ 1 ▶ Last

☐ Mark all on page (selecting matrices will filter sequence summary) Filter  Profiles Export table View factors Hits on page 5 ▼

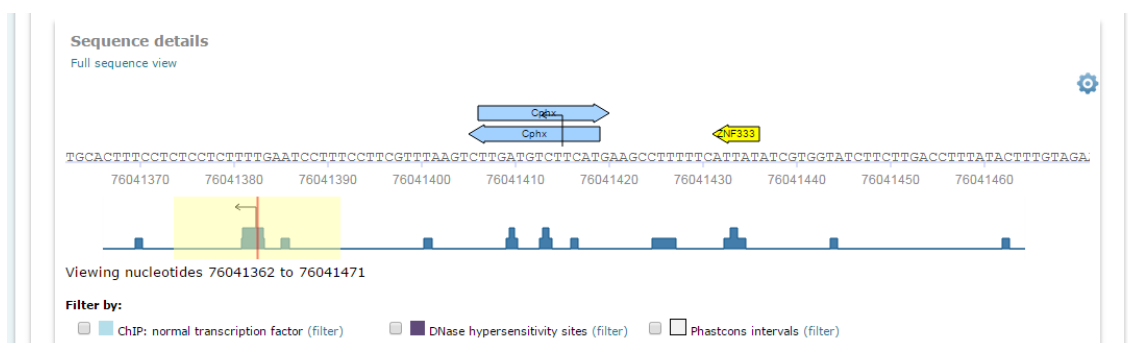
Showing 3 of 3 entries

#	Matrix	Factor name	Graph	Yes	No	Yes/No \neq	Matched promoters in Yes	Matched promoters in No	>>
<input type="checkbox"/>	V\$BBX_03	Bbx		0.0115	0.0023	5.0000	0.0077	0.0023	
<input type="checkbox"/>	V\$GLI_Q3	GLI		0.0107	0.0023	4.6667	0.0092	0.0023	
<input type="checkbox"/>	V\$CEBPA_Q6	C/EBPalpha		0.0667	0.0437	1.5263	0.0636	0.0429	

Scrolling down to the “Sequence details” section a graphical view of the results is provided:



Click the Sequence name to display the nucleotide level view for a specific promoter:



A frequency bar displays the distribution of sites along the entire sequence, with the red orientation point and flanking yellow highlighting indicating the portion of the sequence that is displayed at the nucleotide level above. Slide the red orientation point along the frequency bar to view other portions of the sequence in detail.

Optionally use the ChIP fragments, DNase hypersensitivity sites and Phastcons intervals filters to apply additional biological context to your results. Activating these filters removes any sites that do not overlap with regions of sequence that are suggested to be available

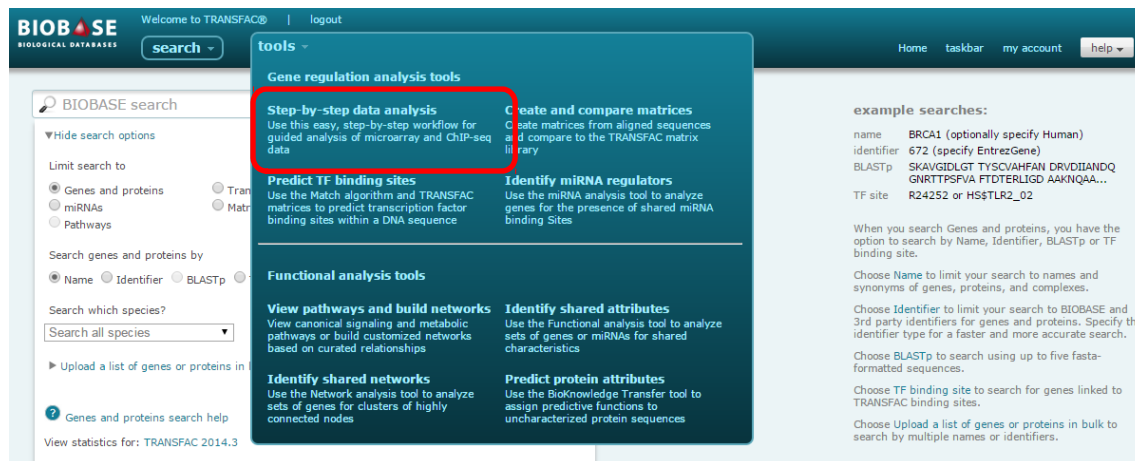
to the transcriptional machinery based on being bound by transcription factors (ChIP fragments), being accessible to DNase enzymes (DNase hypersensitivity) or conserved between human and mouse (Phastcons).

Your analysis will be preserved in the taskbar for 7 days. To preserve it indefinitely, click the “Save this report” link at the top of the report.

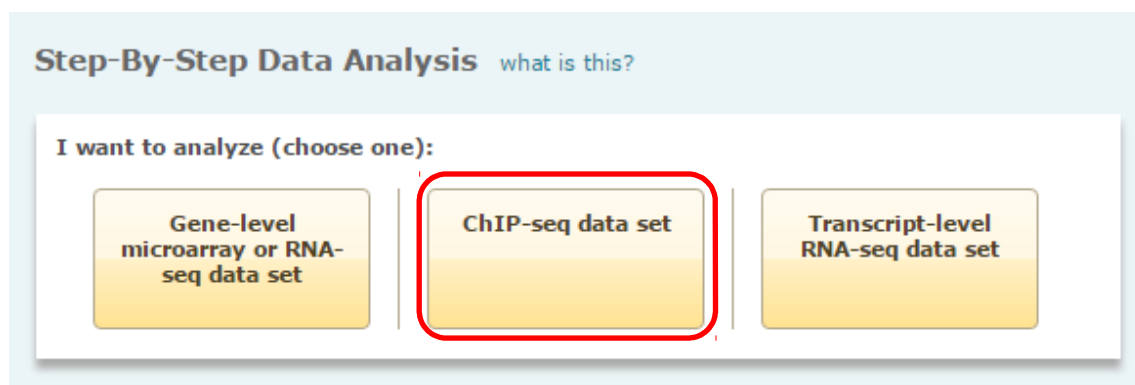
Analyze a ChIP-seq data set

Use TRANSFAC’s step-by-step analysis workflow for ChIP-seq data sets to identify the transcription factors that are most likely to be responsible for the differential expression pattern observed in your experimental data as compared to background data.

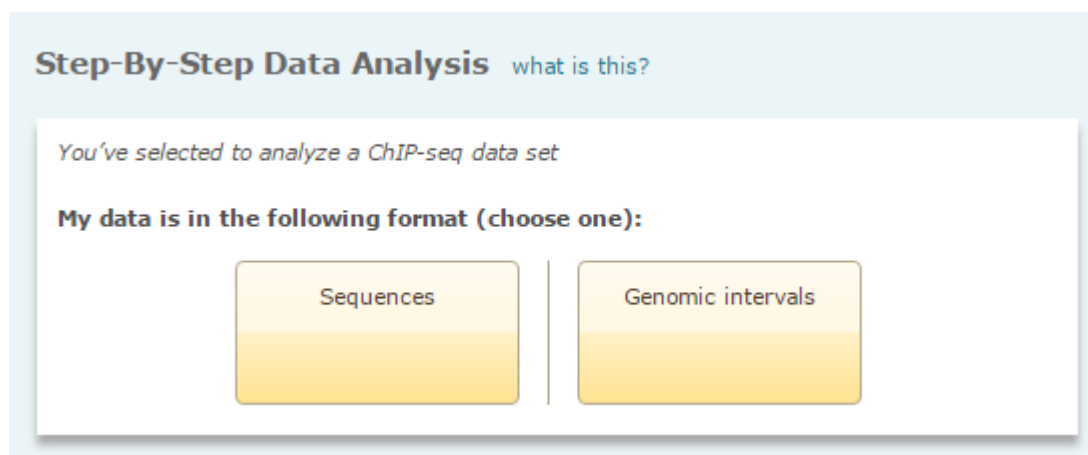
Select the “Step-by-step data analysis” options from the tools menu:



Then select the ChIP-seq data set option:



Specify whether your data is in the format of sequences (FASTA, EMBL, Genbank and RAW formats are supported) or .bed formatted genomic intervals:



Step-By-Step Data Analysis [what is this?](#)

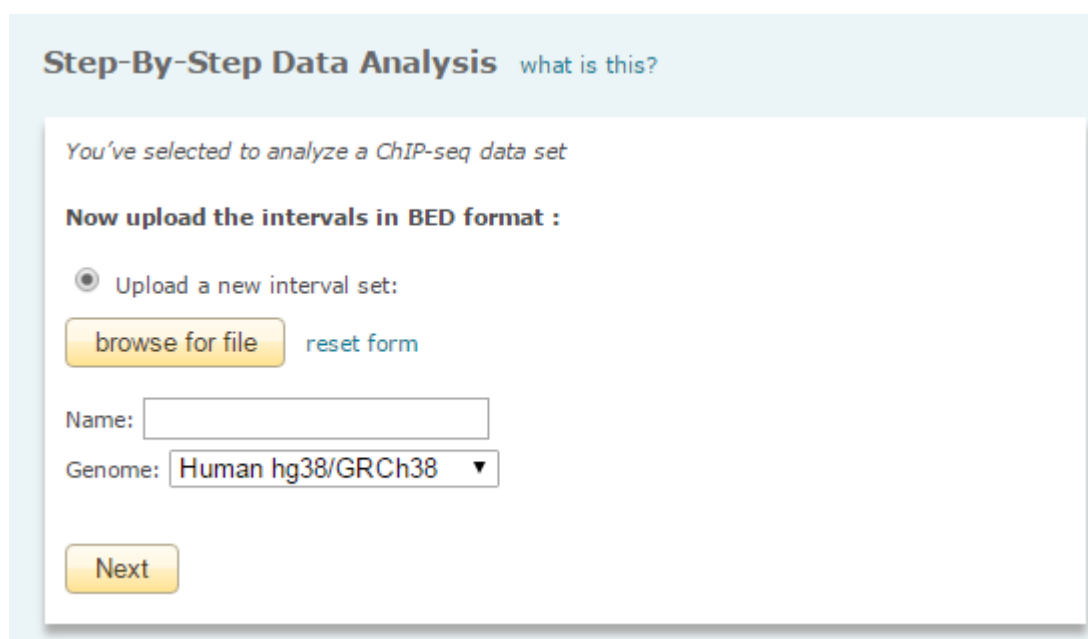
You've selected to analyze a ChIP-seq data set

My data is in the following format (choose one):

Sequences

Genomic intervals

Upload your data set (in this example we are uploading a set of genomic intervals):



Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a ChIP-seq data set

Now upload the intervals in BED format :

☒ Upload a new interval set:

browse for file

[reset form](#)

Name:

Genome: ▼

Next

Note that once you've run at least one analysis you will have the additional option to select a previously uploaded data set instead of uploading a new data set.

Next you will specify the background set to use:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a ChIP-seq data set

You've selected the experimental genomic intervals: ChIP_JUN_exp595_K562_chr1

Now select the background sequence set to be used in the analysis :

- ☒ Experimental genomic intervals shifted by 1000 bp
- ☐ Randomly generated sequences
- ☐ Use a previously uploaded interval set:
- ☐ Upload a new interval set:

Next

You have multiple options for the background set including:

- Using the data set that you just uploaded, shifted by 1,000 bp (the default, recommended option)
- Using a randomly generated sequence set
- Using a previously uploaded interval set
- Uploading a new interval set

Note that if your data is uploaded as sequences you will not have the option to shift the data set by 1,000 bp, you will only the latter three options.

Next you will accept or change the default analysis parameters:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a ChIP-seq data set

You've selected the experimental genomic intervals: ChIP_ETS1_exp405_K562_hg19_chr1

You've selected the background sequence set : shifted genomic intervals

Now select :

A profile (group of matrices): vertebrate_non_redundant_minFP ▼

Data version: 2015.2 ▼

P-value threshold: 0.01

[Run FMatch](#)

You will:

- Specify the profile (group of matrices) to be used (use the default “vertebrate_non_redundant_minFP” which contains representative matrices for distinctive binding motifs unless you are working with non-vertebrate species or want to limit your analysis to factors that are active in a particular tissue)
- Specify the data version (use the default, current version)
- Specify the P-value threshold (0.01 is the default)

When you click the “Run analysis” button your analysis will be forwarded to the taskbar. Once the analysis completes, the FMatch Analysis Report will automatically load.

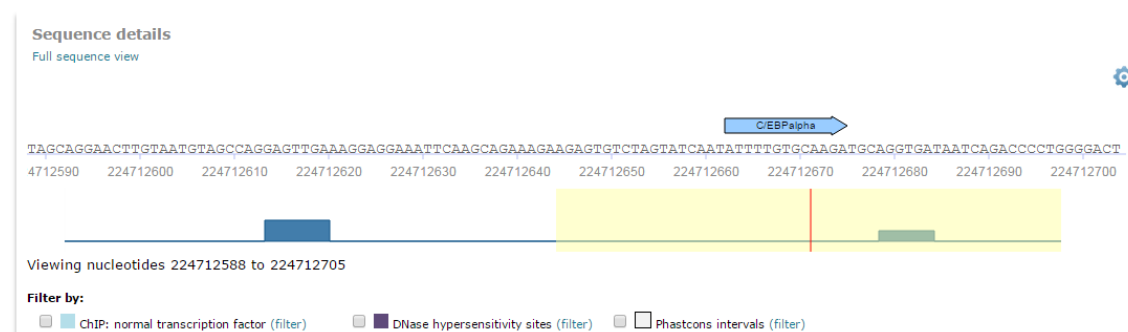
The “Matrix summary” section provides a tabular overview of all matrices that were found to have significantly more binding sites within your uploaded data set than were found in the background data set:

Matrix summary								
First ◀ 1 ▶ Last								
<input type="checkbox"/> Mark all on page (selecting matrices will filter sequence summary) Filter Profiles Export table View factors Hits on page 5 ▼								
Showing 3 of 3 entries								
#	Matrix	Factor name	Graph	Yes	No	Yes/No $\frac{Y}{N}$	Matched promoters in Yes	Matched promoters in No
<input type="checkbox"/>	V\$BBX_Q3	Bbx		0.0115	0.0023	5.0000	0.0077	0.0023
<input type="checkbox"/>	V\$GLI_Q3	GLI		0.0107	0.0023	4.6667	0.0092	0.0023
<input type="checkbox"/>	V\$CEBPA_Q6	C/EBPalpha		0.0667	0.0437	1.5263	0.0636	0.0429

Scrolling down to the “Sequence details” section a graphical view of the results is provided:



Click the Sequence name to display the nucleotide level view for a specific sequence:



A frequency bar displays the distribution of sites along the entire sequence, with the red orientation point and flanking yellow highlighting indicating the portion of the sequence that is displayed at the nucleotide level above. Slide the red orientation point along the frequency bar to view other portions of the sequence in detail.

Optionally use the ChIP fragments, DNase hypersensitivity sites and Phastcons intervals filters to apply additional biological context to your results. Activating these filters removes any sites that do not overlap with regions of sequence that are suggested to be available to the transcriptional machinery based on being bound by transcription factors (ChIP fragments), being accessible to DNase enzymes (DNase hypersensitivity) or conserved between human and mouse (Phastcons). Note that this option will not be available if you uploaded your data as a set of FASTA or other formatted sequences.

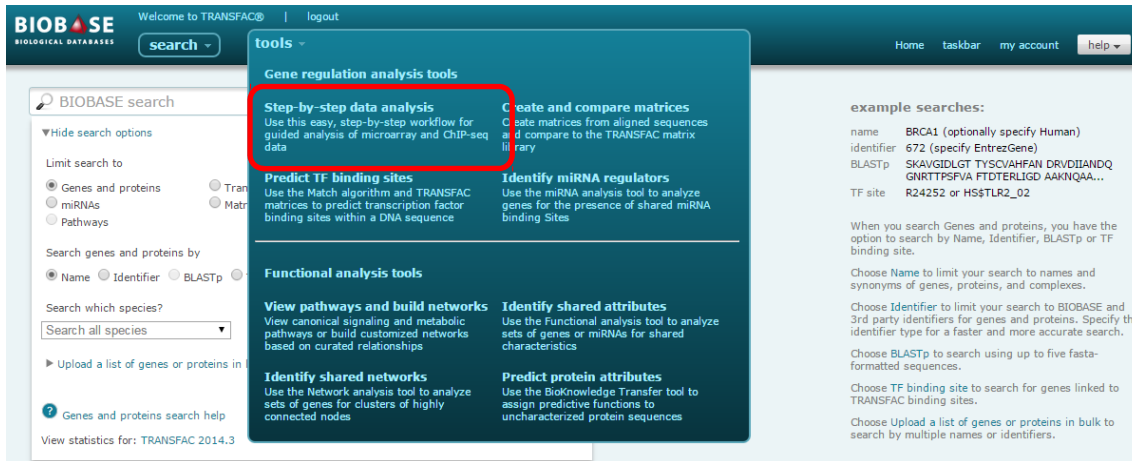
Your analysis will be preserved in the taskbar for 7 days. To preserve it indefinitely, click the “Save this report” link at the top of the report.

Analyze a transcript-based RNA-seq data set

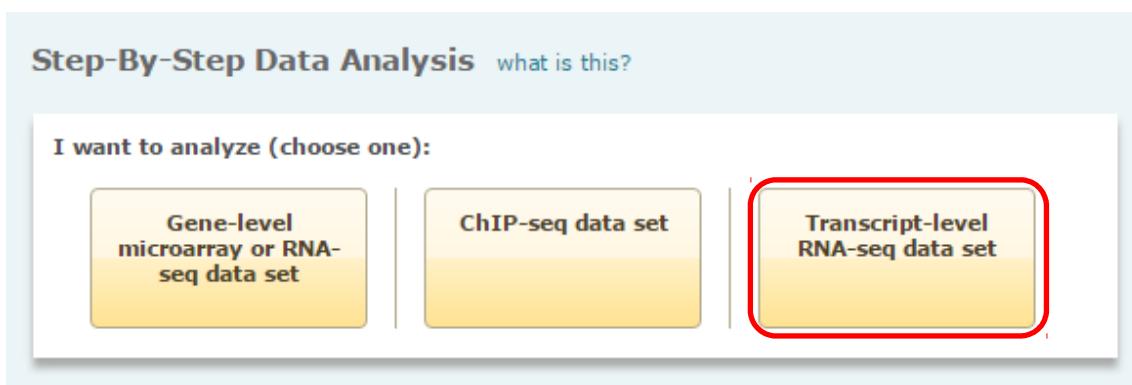
Use TRANSFAC’s step-by-step analysis workflow for transcript-level RNA-seq data sets to

identify the transcription factors that are most likely to be responsible for the differential expression pattern observed in your experimental data as compared to background data.

Select the “Step-by-step data analysis” options from the tools menu:



Then select the Transcript-level RNA-seq data set option:



Specify whether your data is in the form of a pre-filtered list (only the subset of transcripts that were identified as up- or down-regulated is included) or whether you are uploading the complete data set including transcripts that showed no change in expression:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a transcript-level RNA-seq data set

Now upload the data set to be analyzed (choose the appropriate format) :

Pre-filtered
transcript-level
data set

Complete
transcript-level
data set

Upload your data set (in this example we are uploading the complete data set):

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a transcript-level RNA-seq data set

Now select the complete transcript-level RNA-seq data set to be analyzed:

☒ Upload a new data set

browse for file

Enter gene set name

☒ Identifiers

Ensembl ▼

File contains column headers : ☐ Yes ☒ No

Next

Note that once you've run at least one analysis you will have the additional option to select a previously uploaded data set instead of uploading a new data set.

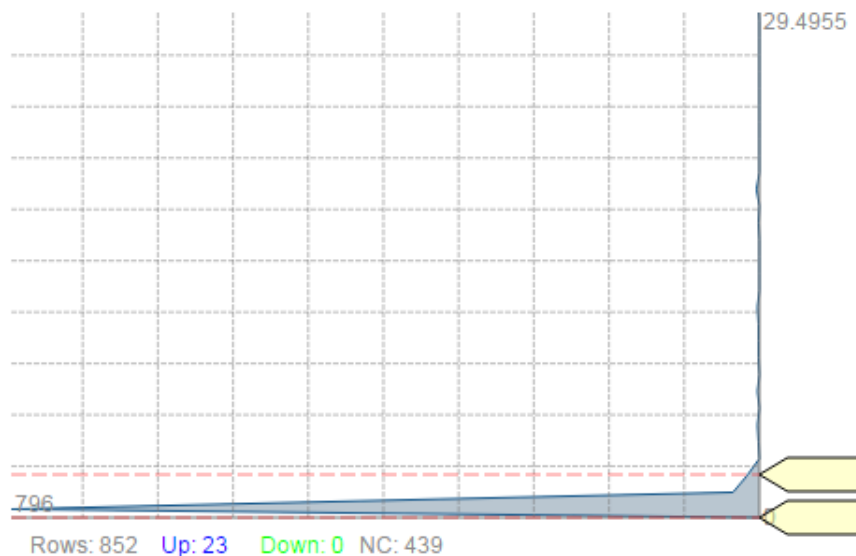
Note that with the initial release only Ensembl transcript identifiers are supported. You will be given a preview of the list of matched transcripts and then will be forwarded to a screen where you will specify the criteria for selecting the up-regulated, down-regulated and non-changed gene sets:

Step-By-Step Data Analysis what is this?

You've selected to analyze a transcript-level RNA-seq data set

You've selected the experimental data set: HEK gkq041

Extract Up/Down/Non-change



Select observation column: **normalized expression (Fold change)**

☒ Select all upregulated elements: expression > 2.512575

☐ Select all downregulated elements: expression < 0

Select non-changed elements: expression = 0.0 +/- 0.15

☐ Analyze Up and Down regulated sets in parallel

Select

In this example, we have chosen to select all up-regulated transcripts which show a normalized expression value of >2.51 as our experimental set and the set of transcripts which show a normalized expression value of 0 ± 0.15 as our non-changed transcript set.

Note that if your data is uploaded in the form of a pre-filtered list this step will automatically be skipped.

Next you will specify the background set to use:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a transcript-level RNA-seq data set

You've selected the experimental data set: HEK gkq041

Now select the background data set to be used in the analysis :

- ☒ A non-change set from the experiment set
- ☐ Randomly generated gene set
- ☐ Previously uploaded gene set :

Housekeeping_human_90 ▼
- ☐ Upload a new background data set

Next

You have multiple options for the background set including:

- Using the non-change transcripts from the data set that you just uploaded, (the default, recommended option)
- Using a randomly generated transcript set
- Using a previously uploaded transcript set
- Uploading a new transcript set

Note that if your data is in the form of a pre-filtered list you will not have the option to directly use the non-change transcripts from the data set as they were not included, you will only have the latter three options.

Next you will accept or change the default analysis parameters:

Step-By-Step Data Analysis [what is this?](#)

You've selected to analyze a transcript-level RNA-seq data set

You've selected the experimental data set: HEK gkq041

You've selected the background data set : non-change set

Now select :

A profile (group of matrices):

Data version:

P-value threshold:

Specify the nucleotide region around the start of the transcript to be analyzed:

to A range of -10,000 to +1,000 bp is supported. [Learn more](#)

You will:

- Specify the profile (group of matrices) to be used (use the default "vertebrate_non_redundant_minFP" which contains representative matrices for distinctive binding motifs unless you are working with non-vertebrate species or want to limit your analysis to factors that are active in a particular tissue)
- Specify the data version (use the default, current version)
- Specify the P-value threshold (0.01 is the default)
- Specify the region surrounding the TSSs (determined by the Ensembl transcripts your data was mapped to) to be considered in the analysis (a range of -10,000 bp to +1,000 bp relative to the TSS is supported, but the default range of -500 bp to +100 bp relative to the TSS is recommended for most analyses)

When you click the "Run FMatch" button your analysis will be forwarded to the taskbar. Once the analysis completes, the FMatch Analysis Report will automatically load.

The "Matrix summary" section provides a tabular overview of all matrices that were found to have significantly more binding sites within your experimental data set than were found in the background data set:

Matrix summary

First 1 Last

☐ Mark all on page (selecting matrices will filter sequence summary) [Filter](#) [Profiles](#) [Export table](#) [View factors](#) Hits on page 5

Showing 3 of 3 entries

#	Matrix	Factor name	Graph	Yes	No	Yes/No	Matched promoters in Yes	Matched promoters in No	>>
<input type="checkbox"/>	V\$BBX_Q3	Bbx		0.0115	0.0023	5.0000	0.0077	0.0023	
<input type="checkbox"/>	V\$GLI_Q3	GLI		0.0107	0.0023	4.6667	0.0092	0.0023	
<input type="checkbox"/>	V\$CEBPA_Q6	C/EBPalpha		0.0667	0.0437	1.5263	0.0636	0.0429	

First 1 Last

☐ Mark all on page (selecting matrices will filter sequence summary) [Filter](#) [Profiles](#) [Export table](#) [View factors](#)

Hits on page 5 ▼

Showing 3 of 3 entries

#	Matrix	Factor name	Graph	Yes	No	Yes/No \neq	Matched promoters in Yes	Matched promoters in No	>>
<input type="checkbox"/>	V\$BBX_Q3	Bbx		0.0115	0.0023	5.0000	0.0077	0.0023	
<input type="checkbox"/>	V\$GLI_Q3	GLI		0.0107	0.0023	4.6667	0.0092	0.0023	
<input type="checkbox"/>	V\$CEBPA_Q6	C/EBPalpha		0.0667	0.0437	1.5263	0.0636	0.0429	

Scrolling down to the “Sequence details” section a graphical view of the results is provided, along with a summary of promoters and genes that are linked to the transcript:

Sequence summary

Experimental set | Background set

First 1 2 3 4 5 Last

☐ Mark all on page [Filter](#) [Export table](#) [Ontology](#) [View genes](#) [Functional analysis](#) [Network analysis](#) Hits on page 5

Showing 5 of 852 entries

#	Sequence	Promoter report	Gene	Matrices count	Sites count	Strand	normalized expression (Fold change)
<input type="checkbox"/>	ENST00000216122 (chr1)	View promoter PM000590233	C1orf110	6	9	REVERSE	0.743883
<input type="checkbox"/>	ENST00000361874 (chr19)	View promoter PM000611093 View promoter PM000611094 View promoter PM000611092	BBC3	3	9	REVERSE	0.636014


Experimental set | Background set

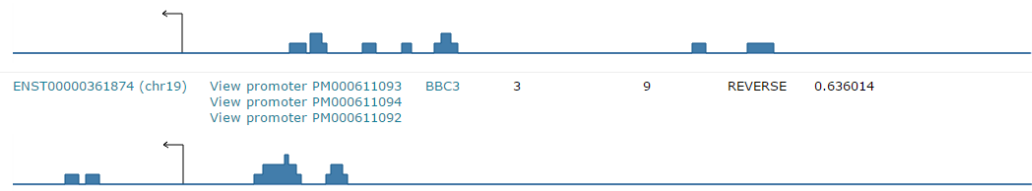
First ◀ 1 2 3 4 5 ▶ Last

☐ Mark all on page [Filter](#) [Export table](#) [Ontology](#) [View genes](#) [Functional analysis](#) [Network analysis](#)

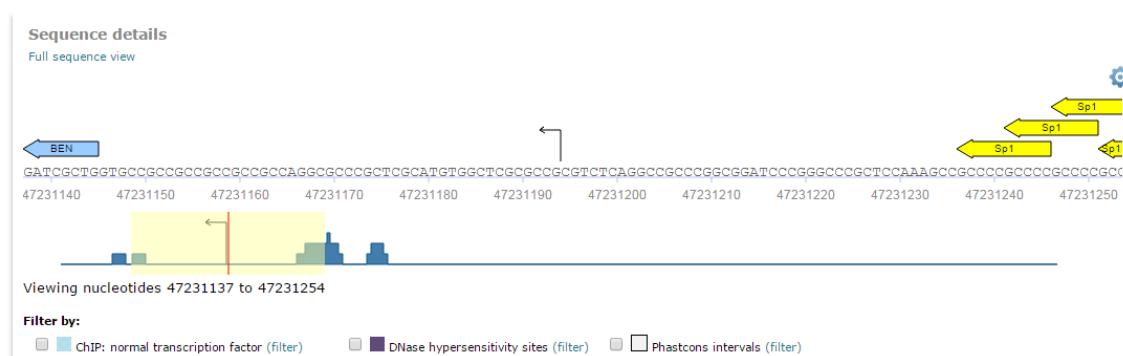
Hits on page

Showing 5 of 852 entries

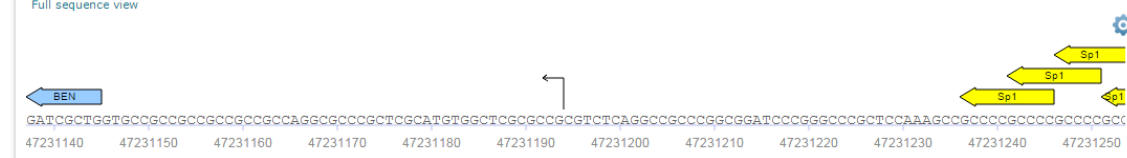
#	Sequence	Promoter report	Gene	Matrices count	Sites count	Strand	normalized expression (Fold change)
	ENST00000216122 (chr1)	View promoter PM000590233	C1orf110	6	9	REVERSE	0.743883



Click the Sequence name to display the nucleotide level view for a specific transcript:



Sequence data



Histogram of the number of children per family. The x-axis is labeled 'Number of children' with values 0, 1, 2, 3, 4, 5. The y-axis is labeled 'Frequency' with values 0, 1, 2, 3, 4, 5. The bars show frequencies: 0 children (1), 1 child (2), 2 children (4), 3 children (3), 4 children (2), 5 children (1). A vertical red line is at x=2.5, and a yellow shaded region is between x=2 and x=3.

Viewing nucleotides 47231137 to 47231254

Filter by:

☐ ☒ ChIP: normal transcription factor (filter) ☐ ☒ DNase hypersensitivity sites (filter) ☐ ☐ Phastcons intervals (filter)

A frequency bar displays the distribution of sites along the entire sequence, with the red orientation point and flanking yellow highlighting indicating the portion of the sequence that is displayed at the nucleotide level above. Slide the red orientation point along the frequency bar to view other portions of the sequence in detail.

Optionally use the ChIP fragments, DNase hypersensitivity sites and Phastcons intervals filters to apply additional biological context to your results. Activating these filters removes

any sites that do not overlap with regions of sequence that are suggested to be available to the transcriptional machinery based on being bound by transcription factors (ChIP fragments), being accessible to DNase enzymes (DNase hypersensitivity) or conserved between human and mouse (Phastcons).

Your analysis will be preserved in the taskbar for 7 days. To preserve it indefinitely, click the "Save this report" link at the top of the report.