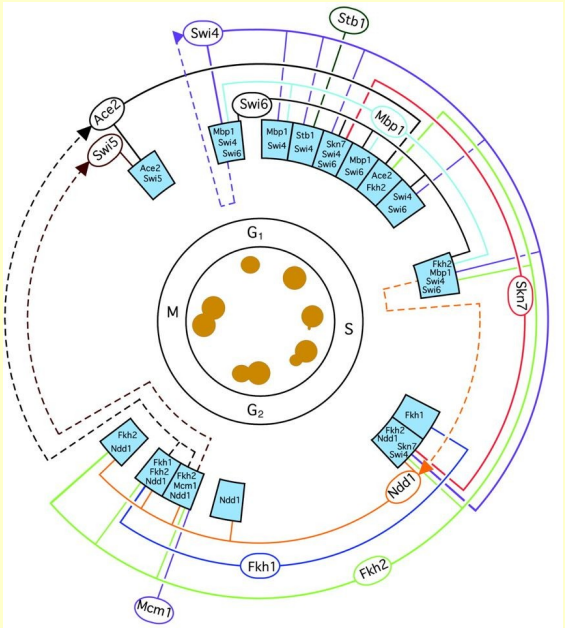
Computational Molecular Biology
Biochem 218 – BioMedical Informatics 231
http://biochem218.stanford.edu/

Gene Regulatory Networks



Doug Brutlag
Professor Emeritus
Biochemistry & Medicine (by courtesy)

Homework 7
Transcription Factor Motifs for Co-expressed Genes

- Choose a group of co-expressed genes from the Yeast Cell Cycle Database by first entering a single gene of interest to you and then clicking on select all co-expressed genes:

  http://genome-www.stanford.edu/cgi-bin/cellcycle/search

- Collect all the names of the genes and enter them in Go Term Finder to discover common biological processes, common molecular mechanisms and common cellular components in the Saccharomyces Genome Database:

  http://www.yeastgenome.org/cgi-bin/GO/goTermFinder.pl

- Using the Batch download link, collect all the gene sequences plus 1 KB upstream and enter the 1 KB upstream sequence from all the genes into MDScan and search for motifs 8, 10 or 12 bases long:

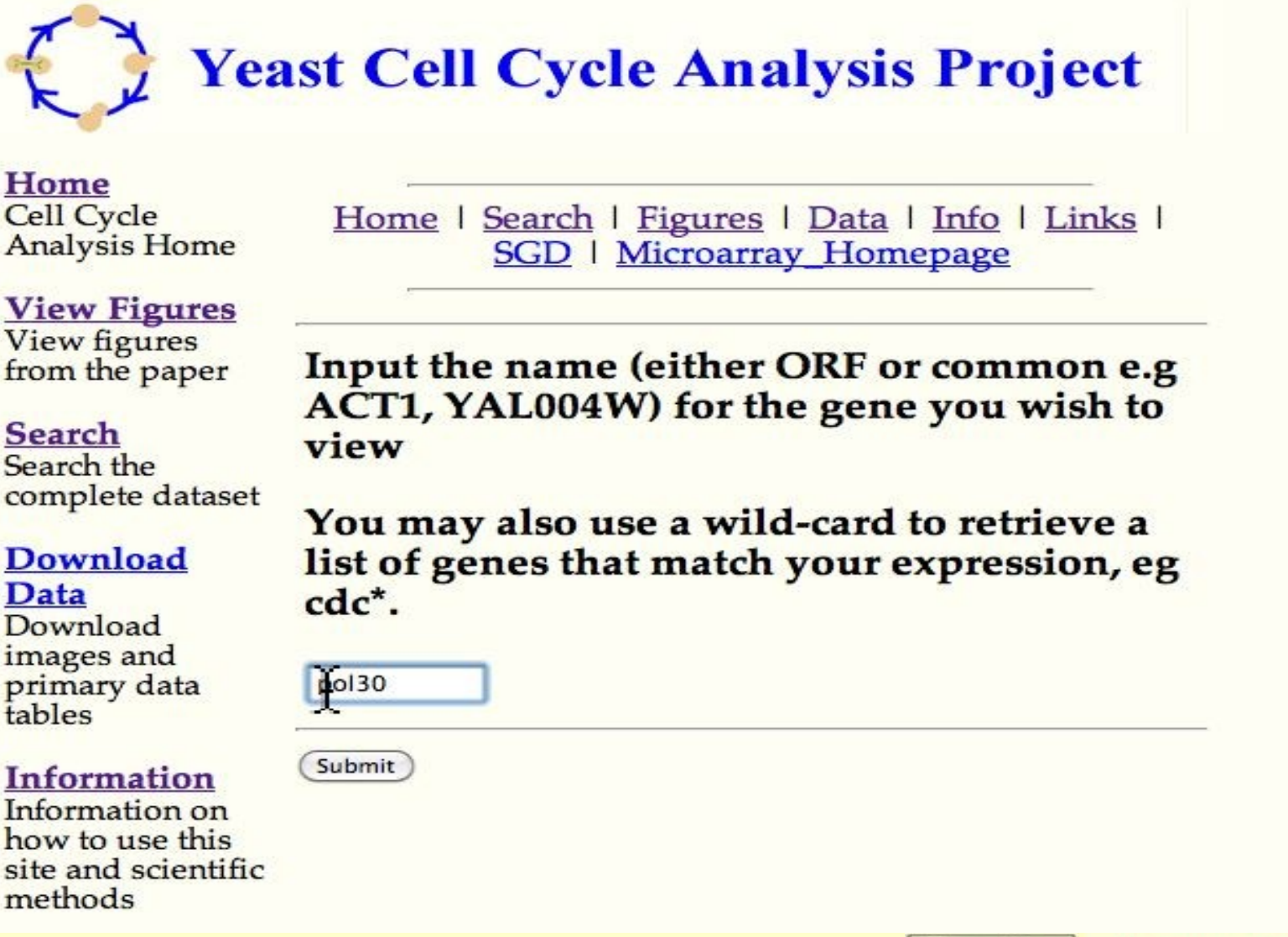
  http://mdscan.stanford.edu/

- You may have to remove simple sequences such as (tatatatata or ttttttttt or cacacacaca) by globally replacing them with an X character in order to ignore simple sequence repeats from the list of motifs found.
- Enter your list of co-expressed gene names in YeasTract web site and see if MDScan has discovered any known transcription binding sites:

  http://www.yeastract.com/

Yeast Cell Cycle Database

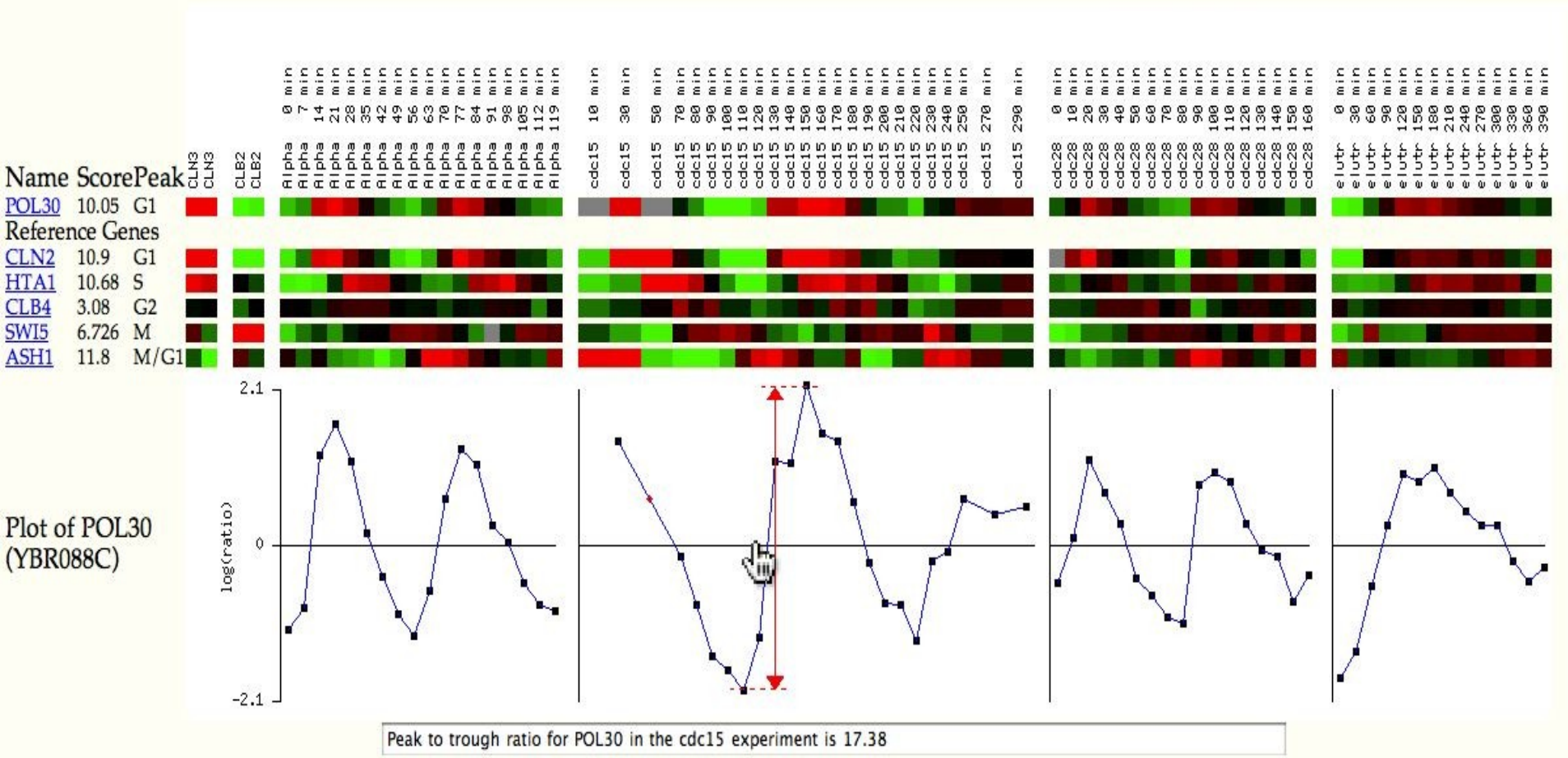
http://genome-www.stanford.edu/cgi-bin/cellcycle/search



Yeast Cell Cycle Database

Yeast Cell Cycle Analysis Project

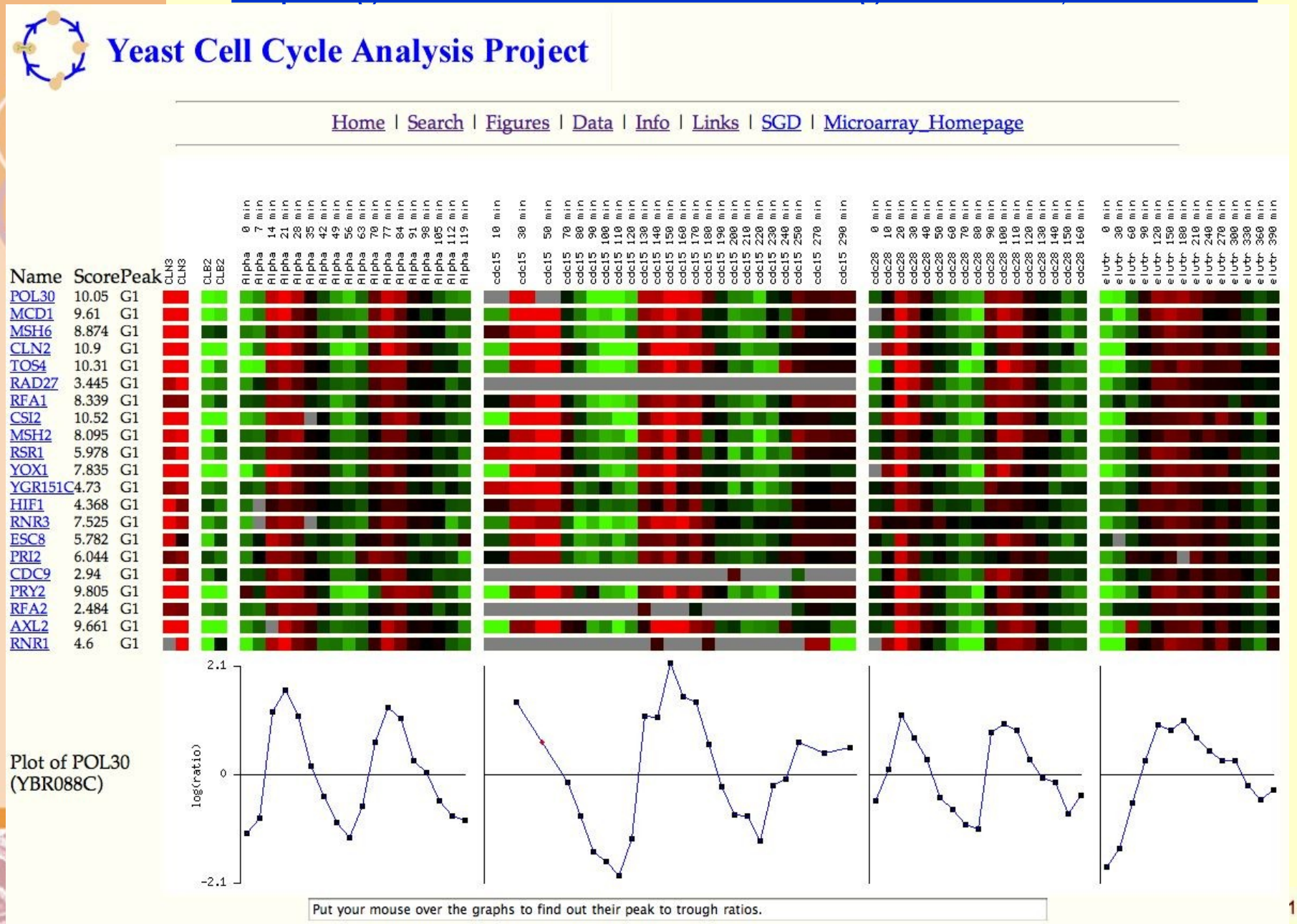
Home | Search | Figures | Data | Info | Links | SGD | Microarray_Homepage



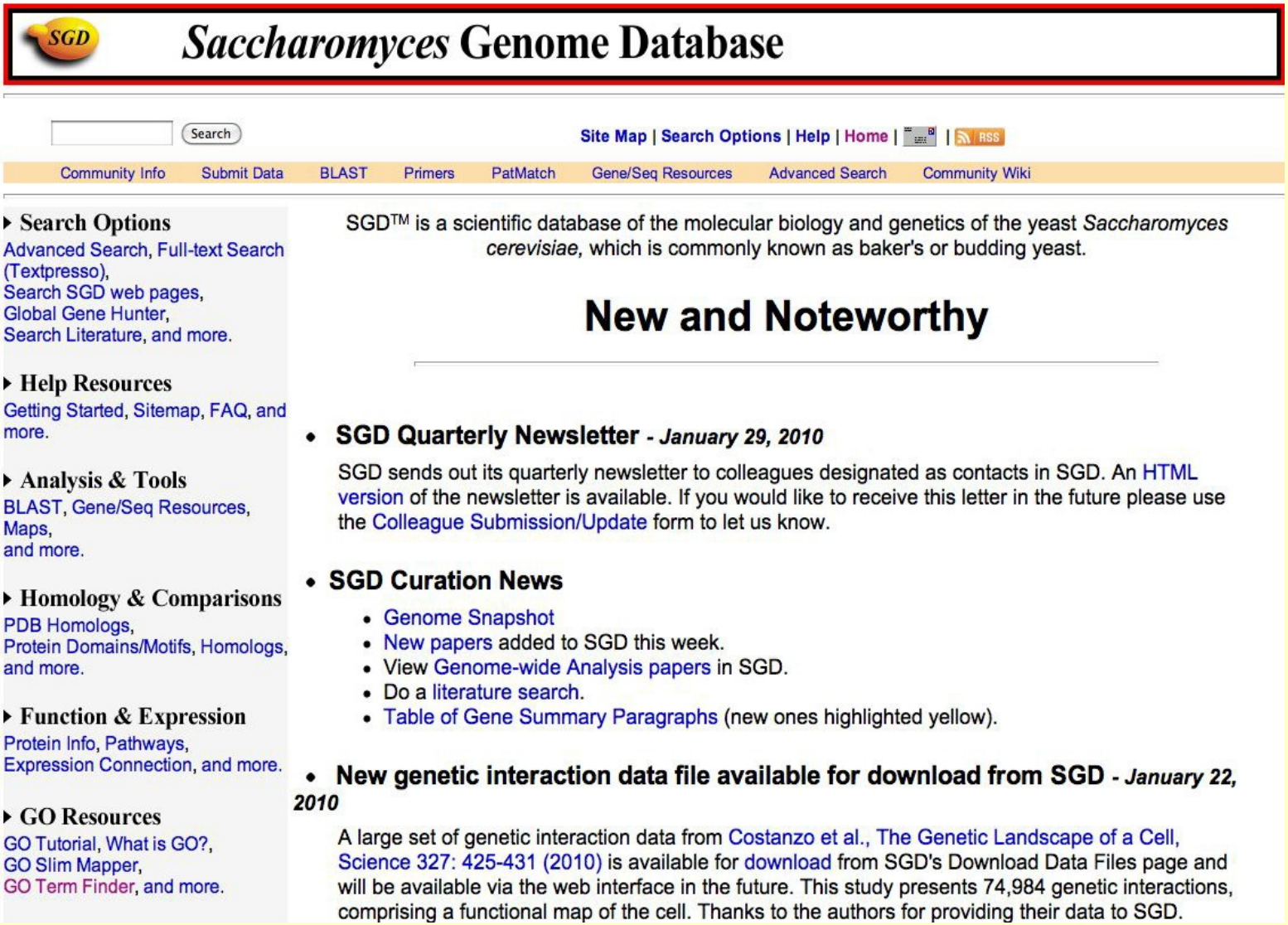
Hit submit to view 20 genes which are expressed similarly to the queried gene

Yeast Cell Cycle Database

http://genome-www.stanford.edu/cgi-bin/cellcycle/search

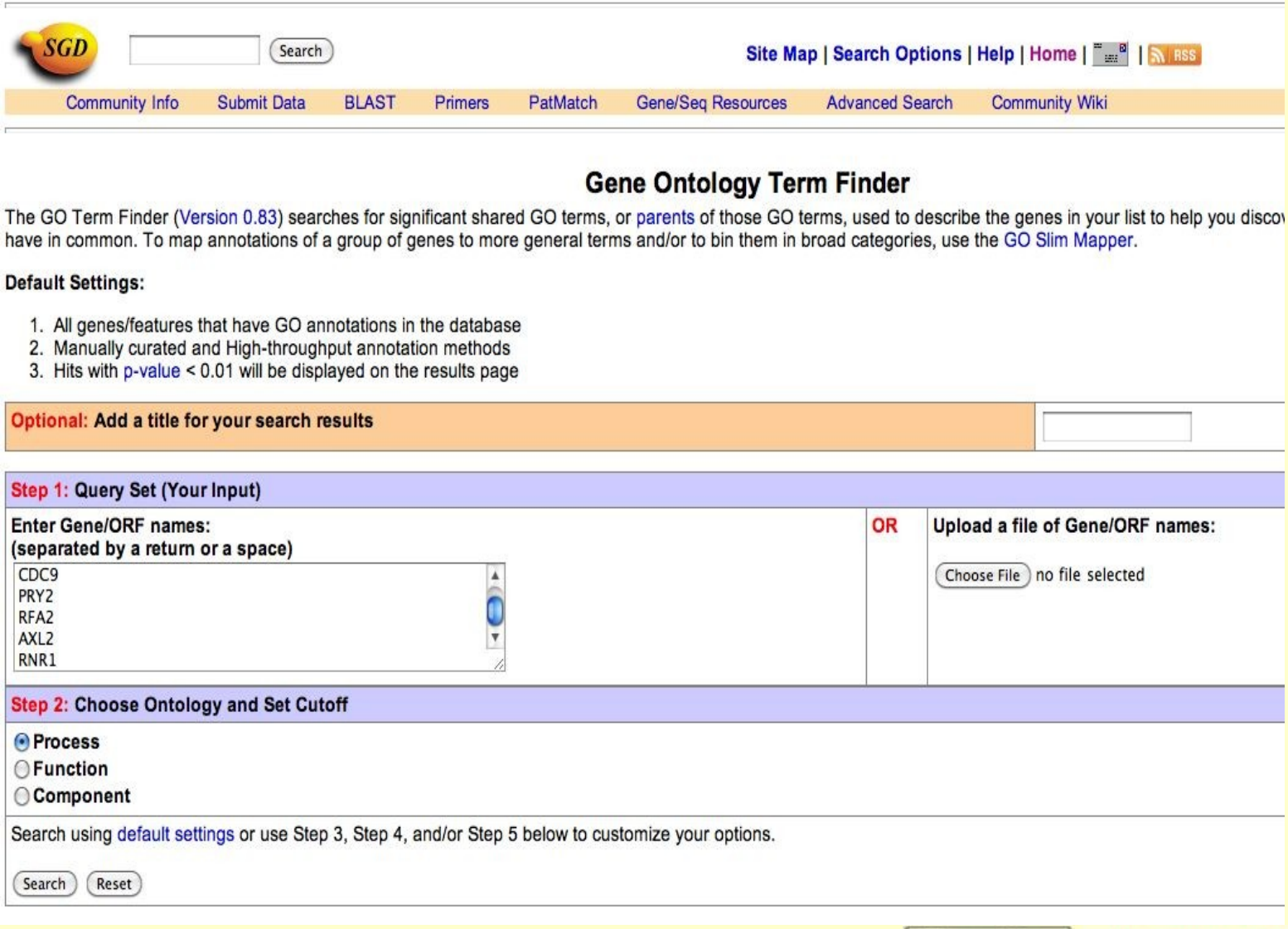


Saccharomyces Genome Database
http://www.yeastgenome.org/

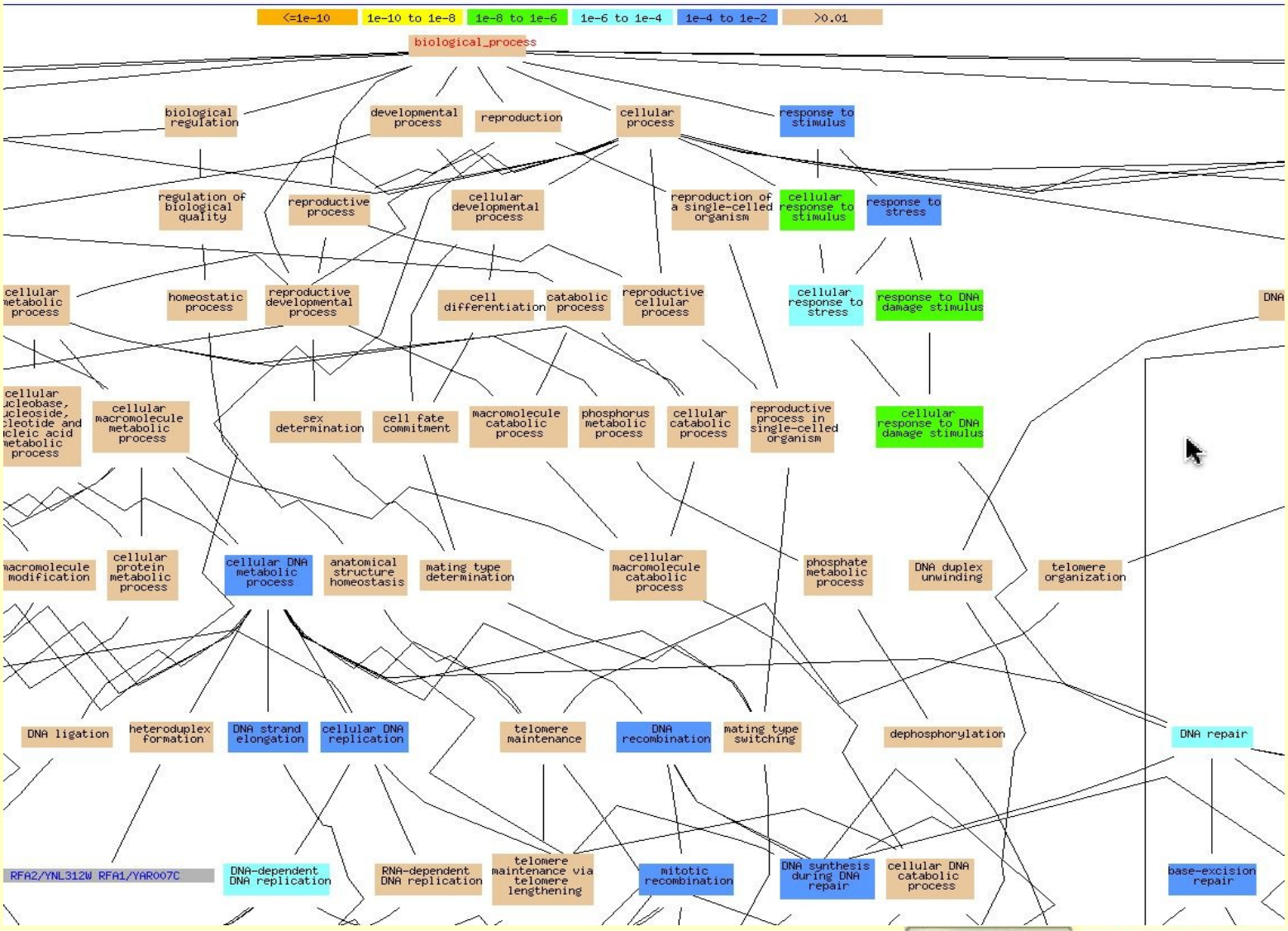


Saccharomyces Genome Database Go Term Finder

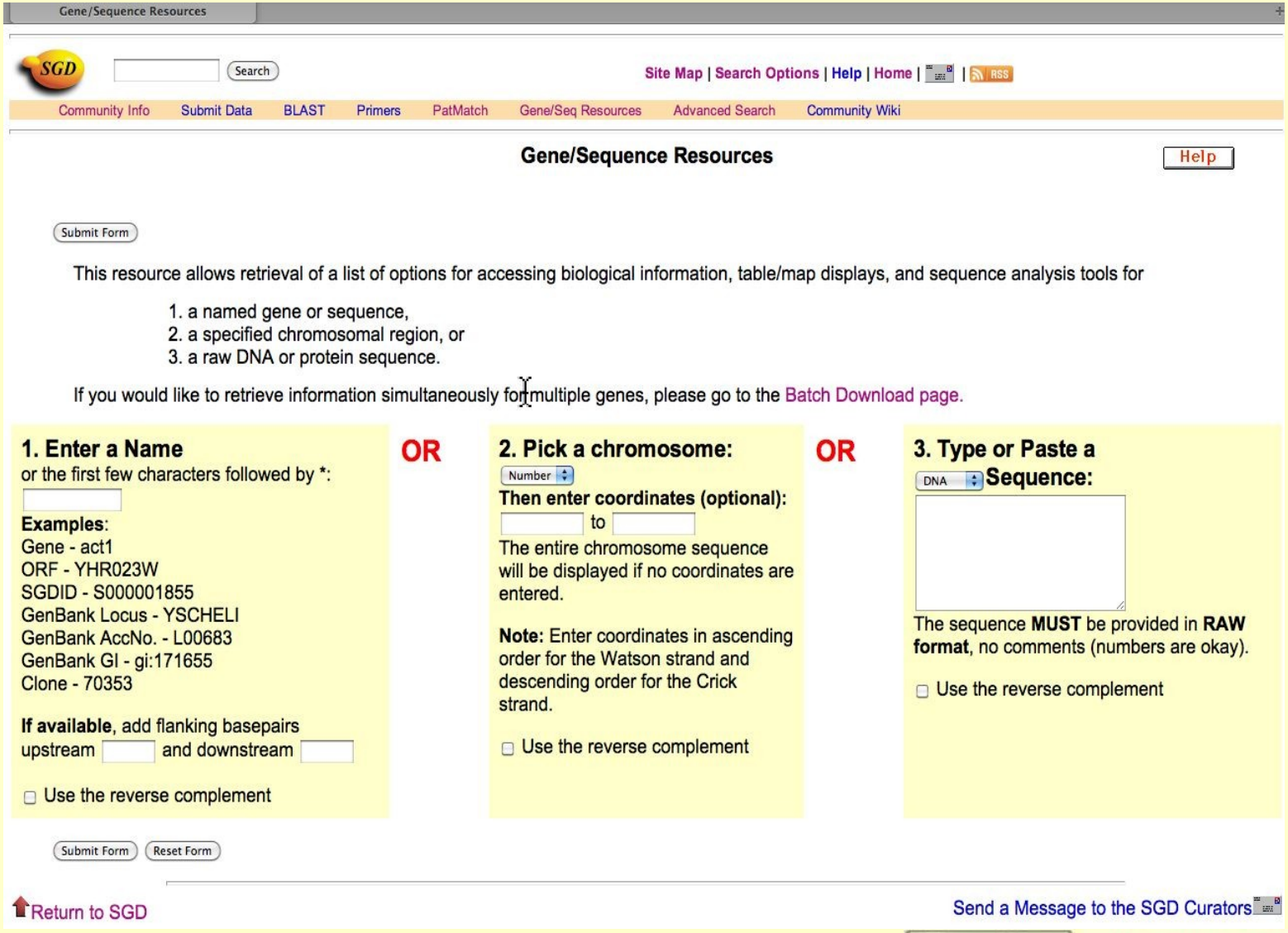
http://www.yeastgenome.org/



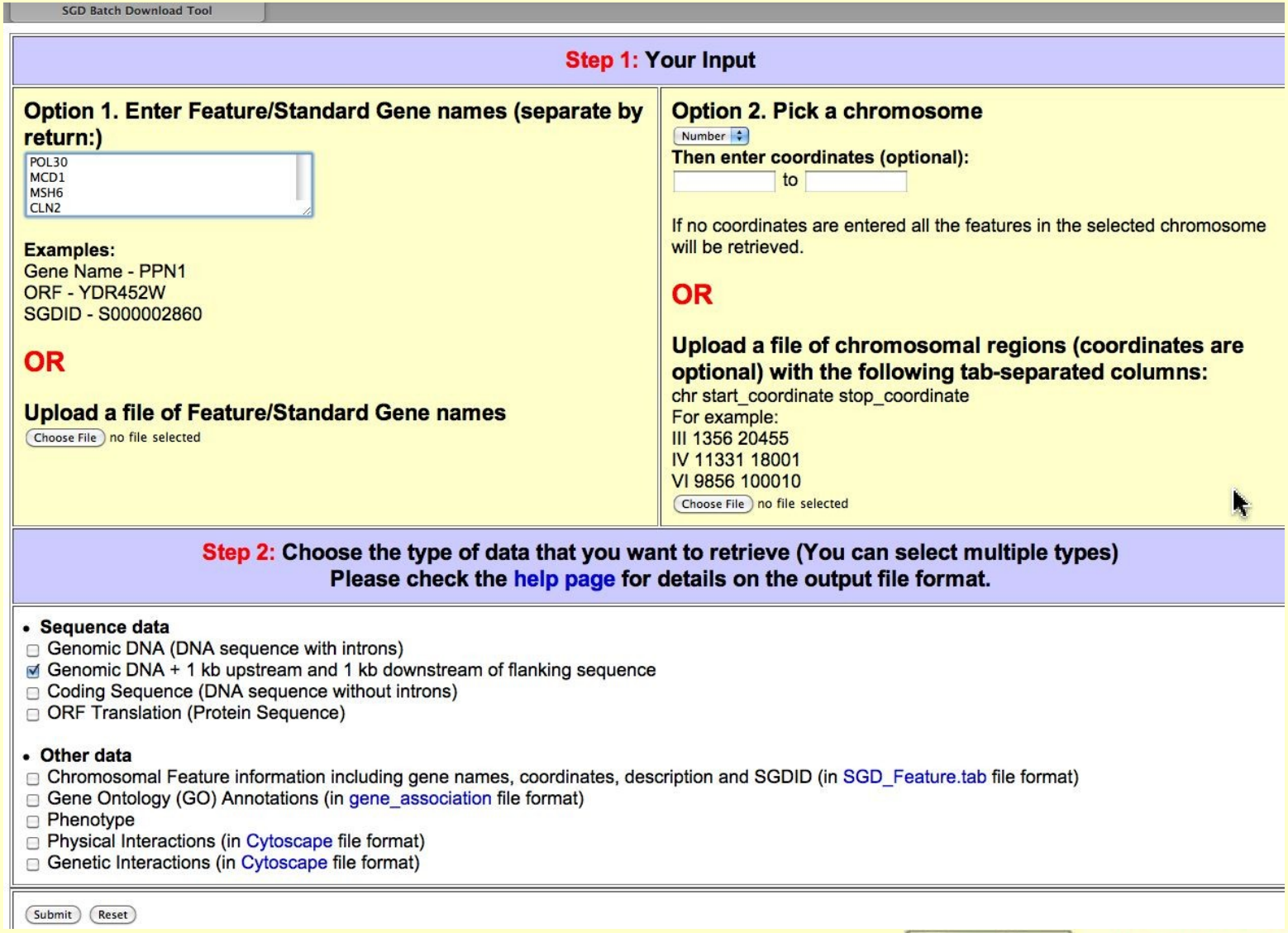
Saccharomyces Genome Database Go Term Finder Results



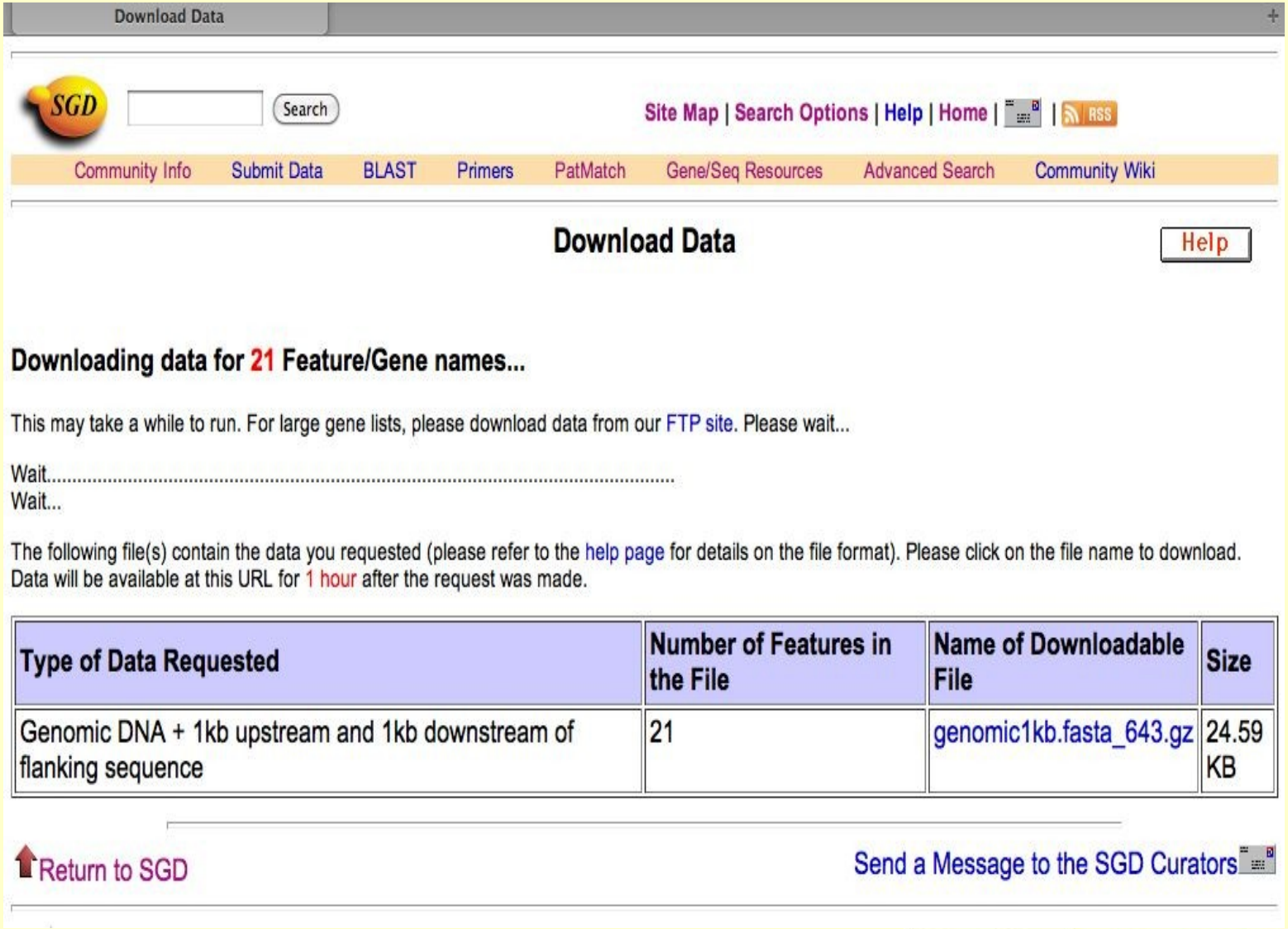
Saccharomyces Genome Database
Gene Sequence Resources
http://www.yeastgenome.org/cgi-bin/seqTools



Saccharomyces Genome Database
Batch Download Tool
http://www.yeastgenome.org/cgi-bin/batchDownload



Saccharomyces Genome Database
Batch Download Tool
http://www.yeastgenome.org/cgi-bin/batchDownload



MDScan
http://mdscan.stanford.edu/

MDscan

A Fast and Accurate Motif Finding Algorithm With Applications To Chromatin Immunoprecipitation Microarray Experiments

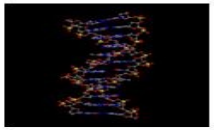
Xiaole Shirley Liu, Douglas L. Brutlag, Jun S. Liu
Stanford Medical Informatics, Stanford University

While chromatin immunoprecipitation followed by cDNA microarray (ChIP-on-chip) has become a popular procedure for studying genome-wide protein-DNA interactions and transcription regulation, it can only map the probable protein-DNA interaction loci within 1-2kb resolution. To pinpoint the interaction sites down to the base pair level, we introduce a novel computational method, Motif Discovery scan (MDscan), that examines the ChIP-array selected sequences and searches for DNA sequence motifs representing the protein-DNA interaction sites. MDscan combines the advantages of two widely adopted motif search strategies, word enumeration and position-specific weight matrix updating, and incorporates the ChIP enrichment information to accelerate the search and enhance its success rate. The intuition is to first search for similar words appearing in the sequences more likely to contain the motif (highly ChIP-enriched sequences) because these sequences have higher signal to noise ratio. Words in each similarity group can initialize a position specific motif matrix and the motif can be updated and refined with the whole input sequences (all ChIP-selected targets). The method showed both speed and accuracy advantages compared to several established motif-finding algorithms in both simulation and published yeast ChIP-on-chip experiments. MDscan can be used not only with the ChIP experiments, but also to find DNA motifs in other experiments in which a subgroup of the sequences can be inferred to contain relatively more abundant motif sites.

We recently developed a new program, Motif Regressor, to better utilize mRNA expression level or ChIP-on-chip enrichment information to improve the performance of MDscan. Motif Regressor first identifies a set of non-redundant candidate motifs using MDscan, and scans the promoter region of every gene in the genome with each candidate motif to measure how good a promoter matches a motif (in terms of both the number of sites and the strength of matching). It then uses linear regression analysis to select motifs whose promoter matching scores are significantly correlated with ChIP-on-chip enrichment or downstream gene expression values. When ranking motifs by linear regression p-value, Motif Regressor automatically picks the best motif and optimal motif width. Due to its computational intensity, Motif Regressor is not currently available as a web server. However, for interested users to explore the program locally, the program is available for download at:
http://www.math.umass.edu/~conlon.mr.html

Obtaining a local copy of MDscan:
MDscan is free-of-charge to academia. Please check out:
Brutlag Bioinformatics Group Software Download and
Academic License Instructions for details.

MDScan Input
http://mdscan.stanford.edu/



Overview
MDscan finds motif from ChIP-on-chip targets

Motif Finding
Search for interesting motifs in your sequences on our server

Input Format
How to specify the input parameters

Output Explanations
How to look at the output file we send you

Motif Regressor
An improved MDscan for download.

Reference
Nat Biotechnol. 2002 20(8):835-9.

Contacts
People behind the project

SeqMotifs
See other motif finding algorithms we have developed.

Suggestions, comments, bugs to: Xiaole Shirley Liu
Last updated 4/15/2004.

MDscan Search

Please don't submit more than one job at a time!!

Submit another job only if one of the following is true:
1. You have received answer from your previous submission.
2. Your previous job was submitted more than one day ago.

User information

Please specify your email so we can send the search result to you

brutlag@stanford.edu

Run information

Please specify a run name (will be included in the result email)

Pol30 1KB upstream -TATATATATA - TTTTTTTTT

Input Sequences

You can either specify a file

Choose File no file selected

Or paste your sequences below:

```
>YKR013W PRY2 SGDID:S000001721, Chr XI from 462602-465591, Verified ORF
TCAAAAATTCACGGCGGCAAAAGGTCTGTAACTACCTAAACATTAAAAGCCTTTGAGCTA
TAAATCATTATGTACACTGAGAATGCCACATGCTAAAACAAGGAAGTTTTCAGCGAAGTA
GCACATTAACTGACCACTCAATAGAAAGAAAACAGAAAGTGGAATGGATGGCCACTTTC
A
```

Background Model

You can use the input sequence as background if you don't specify any of the following.
Or you can specify a background sequence file

Choose File no file selected

Or paste background sequences below:

Or use the precomputed genome background model

Motif Model

Motif width 8
Number of top sequences to look for candidate motifs 10
Number of candidate motifs for scanning the rest sequences 20
Report the top final 5 motifs found

Submit Search Request

MDScan Output

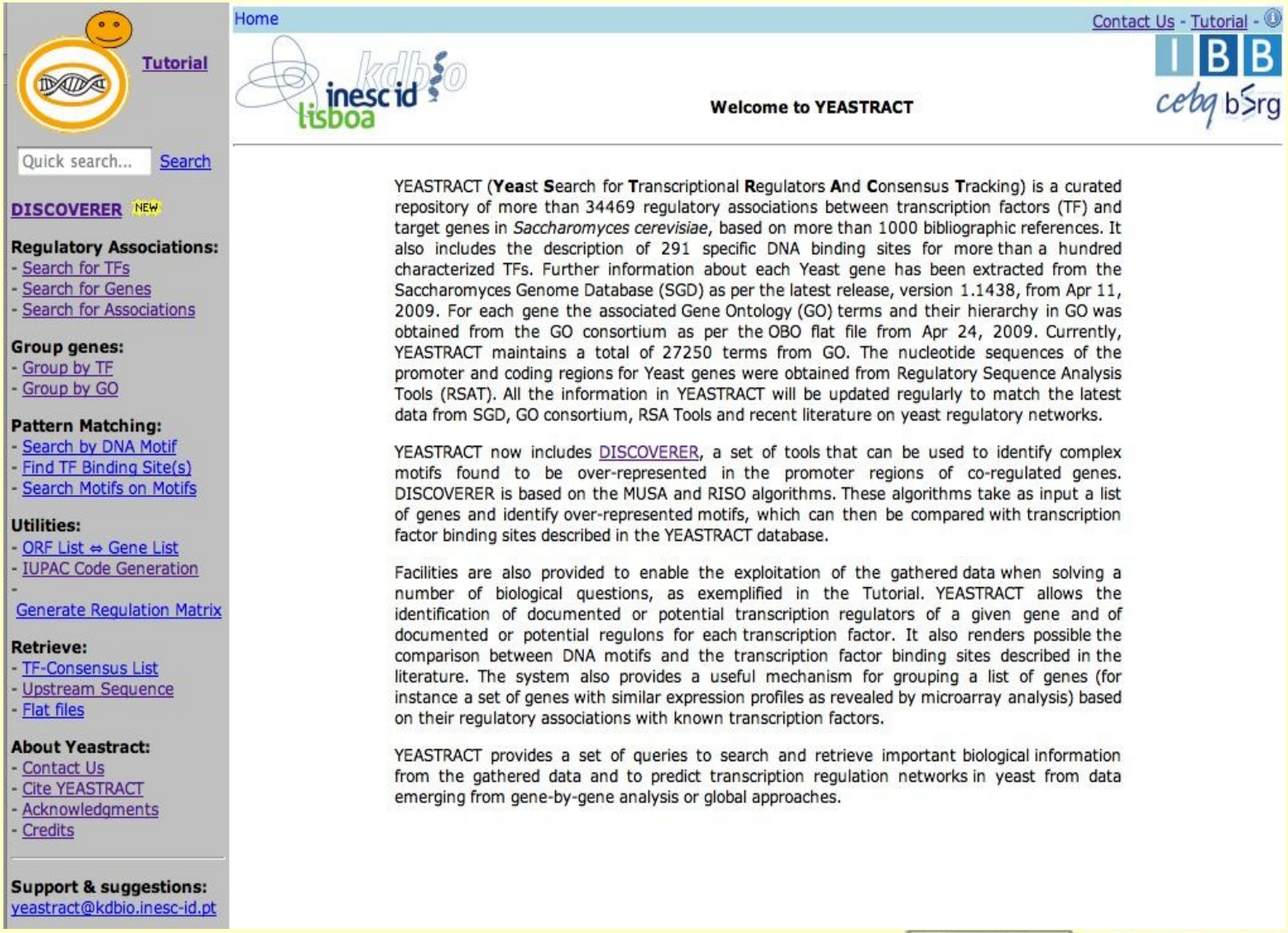
```
****************************************
*                                      *
*         MDscan Search Result         *
*                                      *
****************************************

Pm 0.2639       Minimum match (6/8)
The highest scoring 5 motifs are:

Motif 1: Wid 8; Score 4.276; Sites 68; Con AAACGCGT; RCon ACGCGTTT
********************************
       A      C      G      T       Con  rCon Deg  rDeg
1     47.78   7.85   6.44  37.93     A    T    W    W
2     47.78   7.85  27.57  16.80     A    T    R    Y
3     97.08   0.81   0.81   1.31     A    T    A    T
4      1.31  96.58   0.81   1.31     C    G    C    G
5      1.31   0.81  96.58   1.31     G    C    G    C
6      1.31  96.58   0.81   1.31     C    G    C    G
7      1.31   0.81  96.58   1.31     G    C    G    C
8     28.07   0.81   0.81  70.32     T    A    W    W
| YPL256C CLN2 SGDID:S000006177, Chr XVI from 67614-63977, reverse complement, Verified ORF       Len 1020        Site #1 r 36
TGACGCGA
| YML027W YOX1 SGDID:S000004489, Chr XIII from 220406-223563, Verified ORF        Len 1000        Site #1 f 494
AAACGCGT
| YML027W YOX1 SGDID:S000004489, Chr XIII from 220406-223563, Verified ORF        Len 1000        Site #2 r 496
TTACGCGT
| YML027W YOX1 SGDID:S000004489, Chr XIII from 220406-223563, Verified ORF        Len 1000        Site #3 f 555
AGACGCGA
| YML027W YOX1 SGDID:S000004489, Chr XIII from 220406-223563, Verified ORF        Len 1000        Site #4 f 760
AAACGCGA
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #1 r 695
TTACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #2 r 630
TGACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #3 r 559
TTACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #4 r 509
AAACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #5 f 507
TTACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #6 f 557
AAACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #7 f 628
AAACGCGT
| YER070W RNR1 SGDID:S000000872, Chr V from 297948-302614, Verified ORF   Len 1020        Site #8 f 693
```

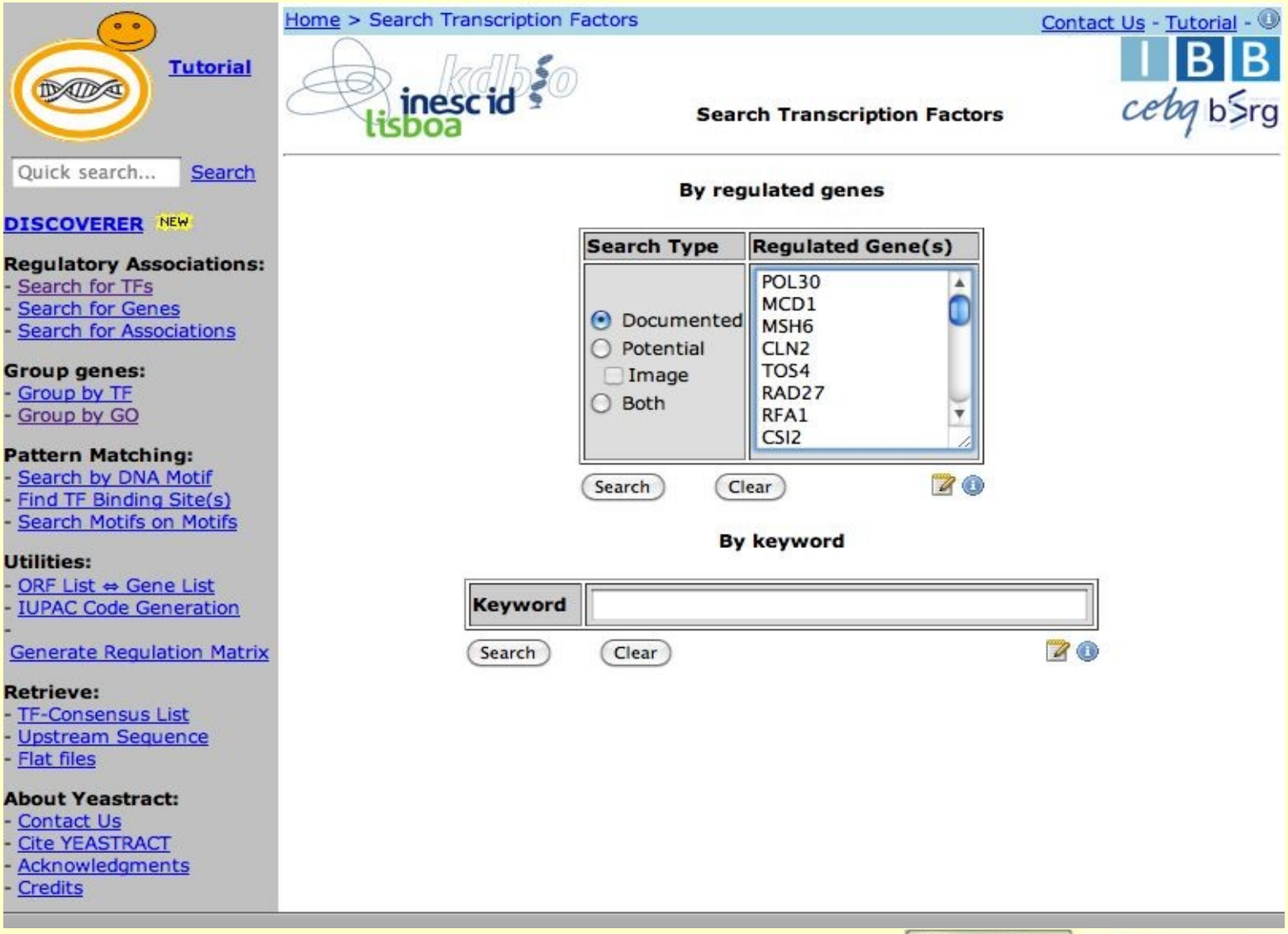
YeasTract Transcription Factor Database

http://www.yeastract.com/



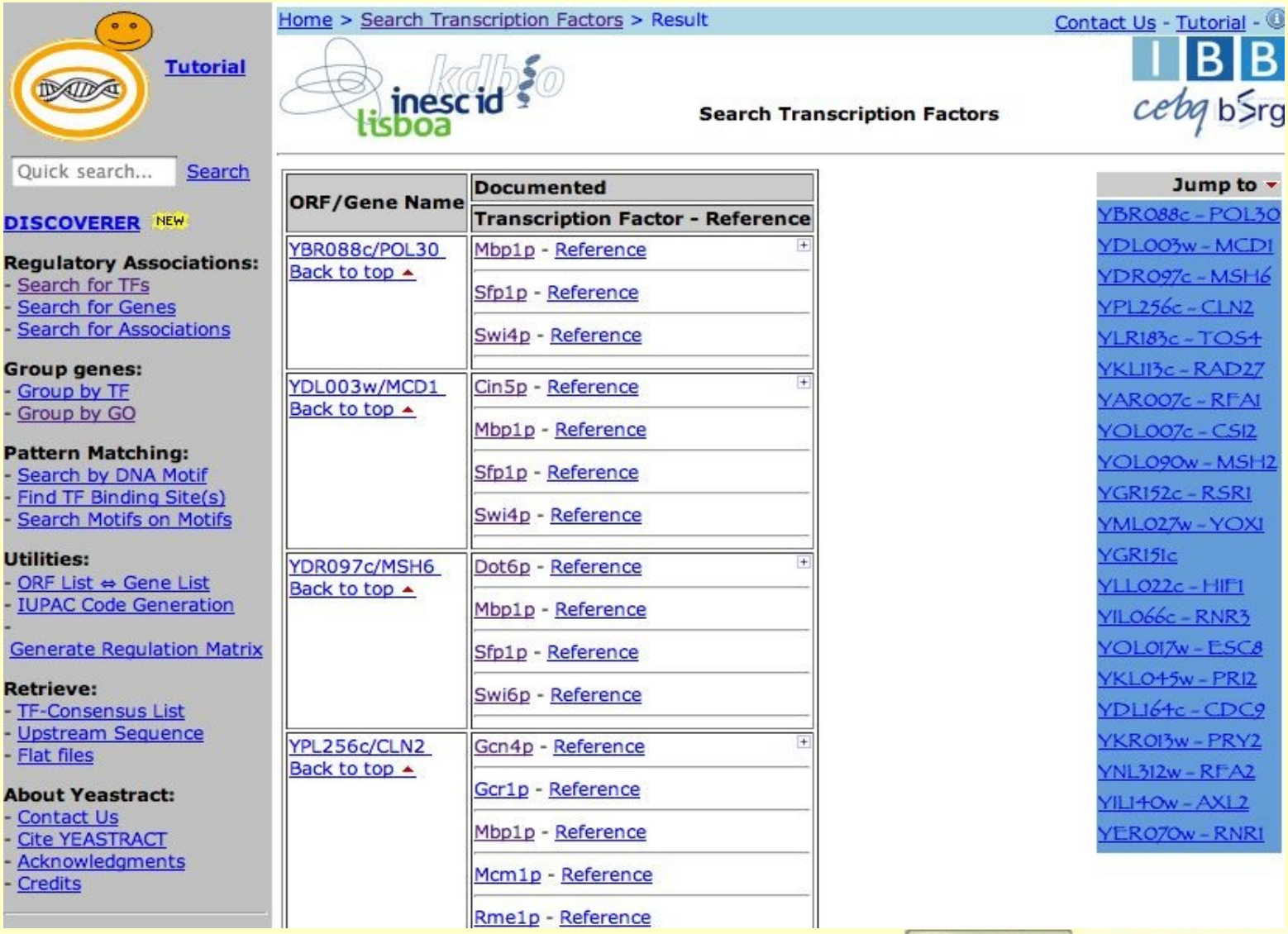
YeasTract Transcription Factor Search

http://www.yeastract.com/formfindregulators.php



YeasTract Transcription Factor Search Results

http://www.yeastract.com/formfindregulators.php



Tutorial

Quick search... Search

DISCOVERER NEW

Regulatory Associations:
- Search for TFs
- Search for Genes
- Search for Associations

Group genes:
- Group by TF
- Group by GO

Pattern Matching:
- Search by DNA Motif
- Find TF Binding Site(s)
- Search Motifs on Motifs

Utilities:
- ORF List ⇔ Gene List
- IUPAC Code Generation
- Generate Regulation Matrix

Retrieve:
- TF-Consensus List
- Upstream Sequence
- Flat files

About Yeastract:
- Contact Us
- Cite YEASTRACT
- Acknowledgments
- Credits

Home > Search Transcription Factors > Result

Contact Us - Tutorial -

Search Transcription Factors

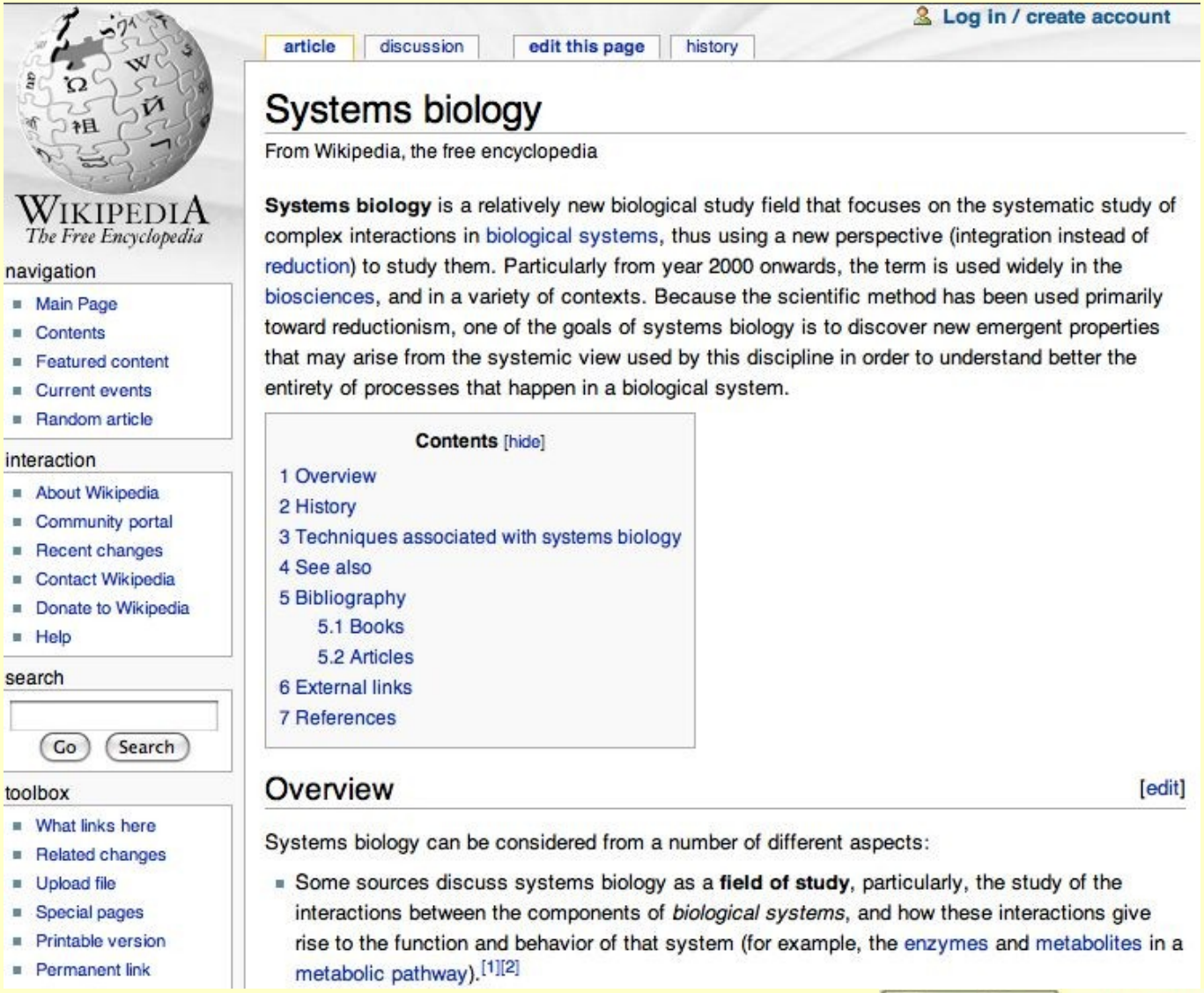
ORF/Gene Name	Documented Transcription Factor - Reference
YBR088c/POL30 Back to top	Mbp1p - Reference
	Sfp1p - Reference
	Swi4p - Reference
YDL003w/MCD1 Back to top	Cin5p - Reference
	Mbp1p - Reference
	Sfp1p - Reference
	Swi4p - Reference
YDR097c/MSH6 Back to top	Dot6p - Reference
	Mbp1p - Reference
	Sfp1p - Reference
	Swi6p - Reference
YPL256c/CLN2 Back to top	Gcn4p - Reference
	Gcr1p - Reference
	Mbp1p - Reference
	Mcm1p - Reference
	Rme1p - Reference

Jump to

YBR088c - POL30
YDL003w - MCD1
YDR097c - MSH6
YPL256c - CLN2
YLR183c - TOS4
YKL113c - RAD27
YAR007c - RFA1
YOL007c - CSI2
YOL090w - MSH2
YGR152c - RSR1
YML027w - YOX1
YGR151c
YLL022c - HIF1
YIL066c - RNR3
YOL017w - ESC8
YKL045w - PRI2
YDL164c - CDC9
YKR013w - PRY2
YNL312w - RFA2
YIL140w - AXL2
YER070w - RNR1

What is Systems Biology?

http://en.wikipedia.org/wiki/Systems_biology



WIKIPEDIA
The Free Encyclopedia

Log in / create account

article | discussion | edit this page | history

navigation
- Main Page
- Contents
- Featured content
- Current events
- Random article

interaction
- About Wikipedia
- Community portal
- Recent changes
- Contact Wikipedia
- Donate to Wikipedia
- Help

search

Go | Search

toolbox
- What links here
- Related changes
- Upload file
- Special pages
- Printable version
- Permanent link

Systems biology

From Wikipedia, the free encyclopedia

Systems biology is a relatively new biological study field that focuses on the systematic study of complex interactions in biological systems, thus using a new perspective (integration instead of reduction) to study them. Particularly from year 2000 onwards, the term is used widely in the biosciences, and in a variety of contexts. Because the scientific method has been used primarily toward reductionism, one of the goals of systems biology is to discover new emergent properties that may arise from the systemic view used by this discipline in order to understand better the entirety of processes that happen in a biological system.

Contents [hide]

1 Overview
2 History
3 Techniques associated with systems biology
4 See also
5 Bibliography
  5.1 Books
  5.2 Articles
6 External links
7 References

Overview [edit]

Systems biology can be considered from a number of different aspects:

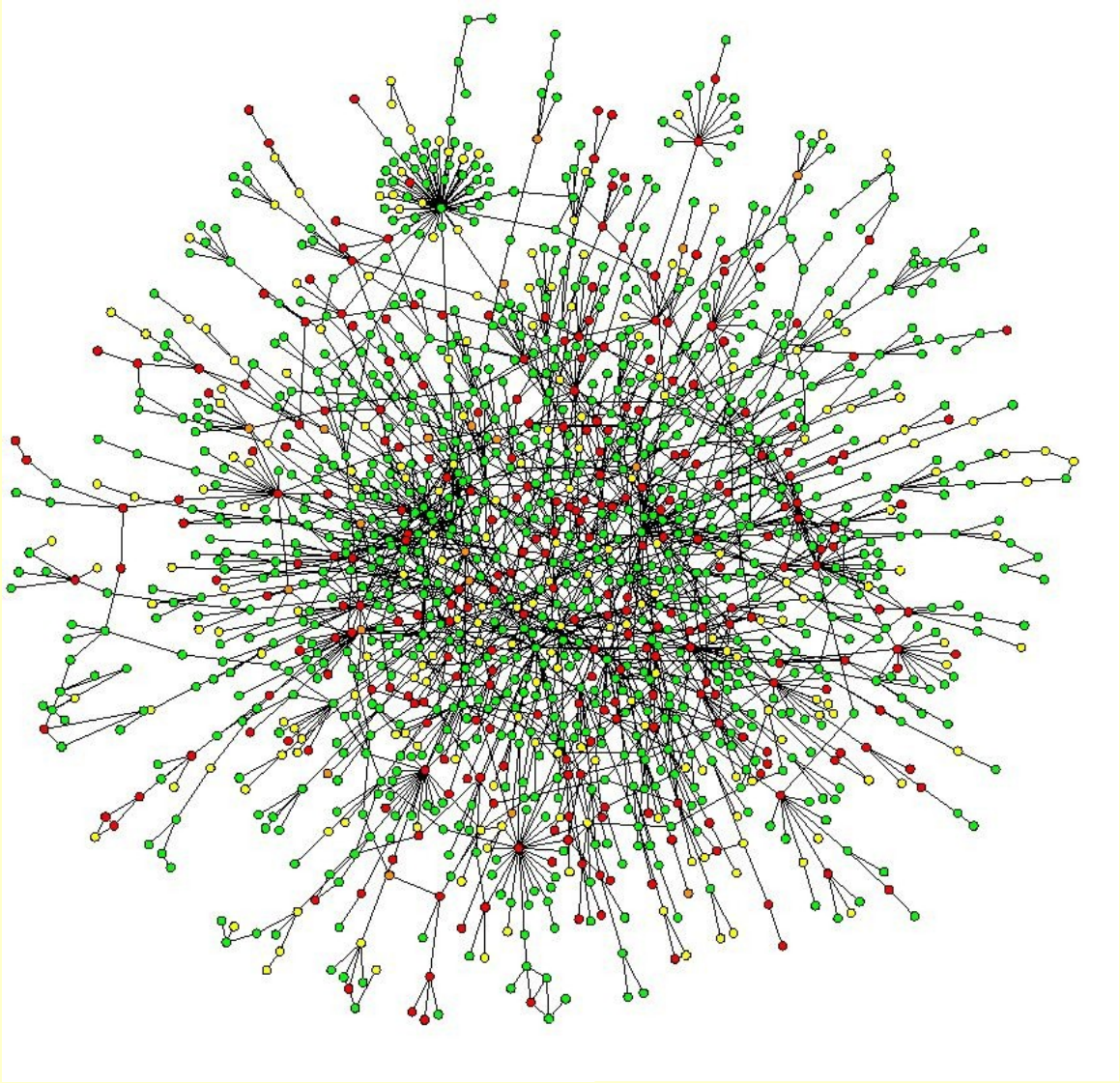
- Some sources discuss systems biology as a **field of study**, particularly, the study of the interactions between the components of *biological systems*, and how these interactions give rise to the function and behavior of that system (for example, the enzymes and metabolites in a metabolic pathway).[1][2]

What is Systems Biology?

- Systematic approach to biology
  - Protein networks involving all protein-protein interactions
  - Metabolic networks involving all enzymes & pathways
  - Gene regulation involving all regulatory interactions
  - Cell behavior involving all environmental responses
  - Population biology involving all selectable traits.
- Challenges of System Biology
  - Identifying all selectable traits, all regulatory interactions, all protein complexes, all metabolic pathways, all regulatory pathways, all selectable traits
  - Computationally representing and simulating cell structures, metabolism, regulation, behavior and selection
  - Usually approached by simplification, studying one protein complex, one metabolic pathway, one regulatory network or one trait at a time. Then combining individual examples to represent the entire cell or organism.

Yeast Protein Interaction Network

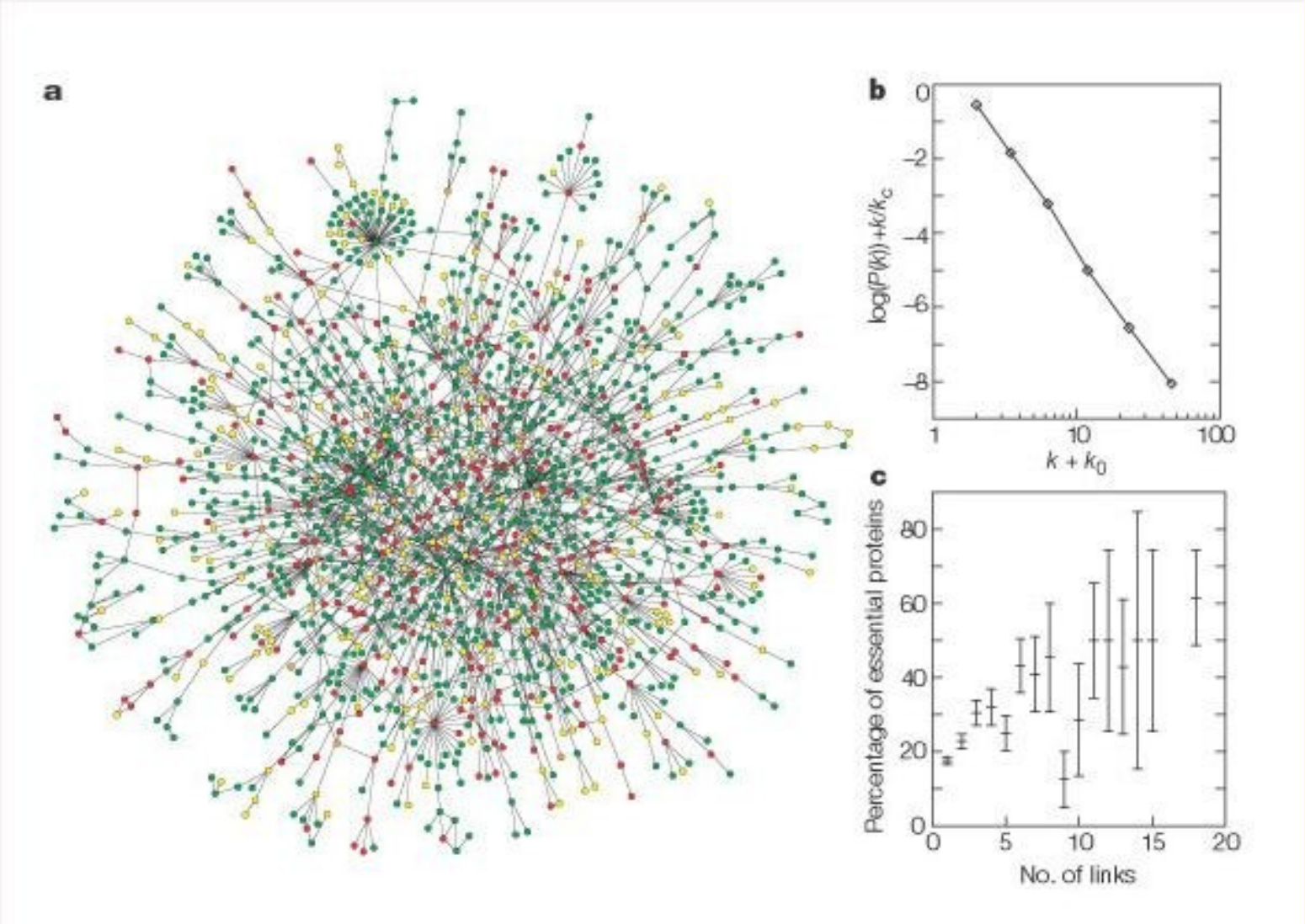
H. Jeong, et al., Nature, Vol 411, p41 (2001)



H. Jeong, et al., Nature, Vol 411, p41 (2001)

Yeast Protein Interaction Network

H. Jeong, et al., Nature, Vol 411, p41 (2001)



H. Jeong, et al., Nature, Vol 411, p41 (2001)

Evolutionary Rate in the Protein Interaction Network

Hunter B. Fraser,[1*†] Aaron E. Hirsh,[2*] Lars M. Steinmetz,[3] Curt Scharfe,[3] Marcus W. Feldman[2]

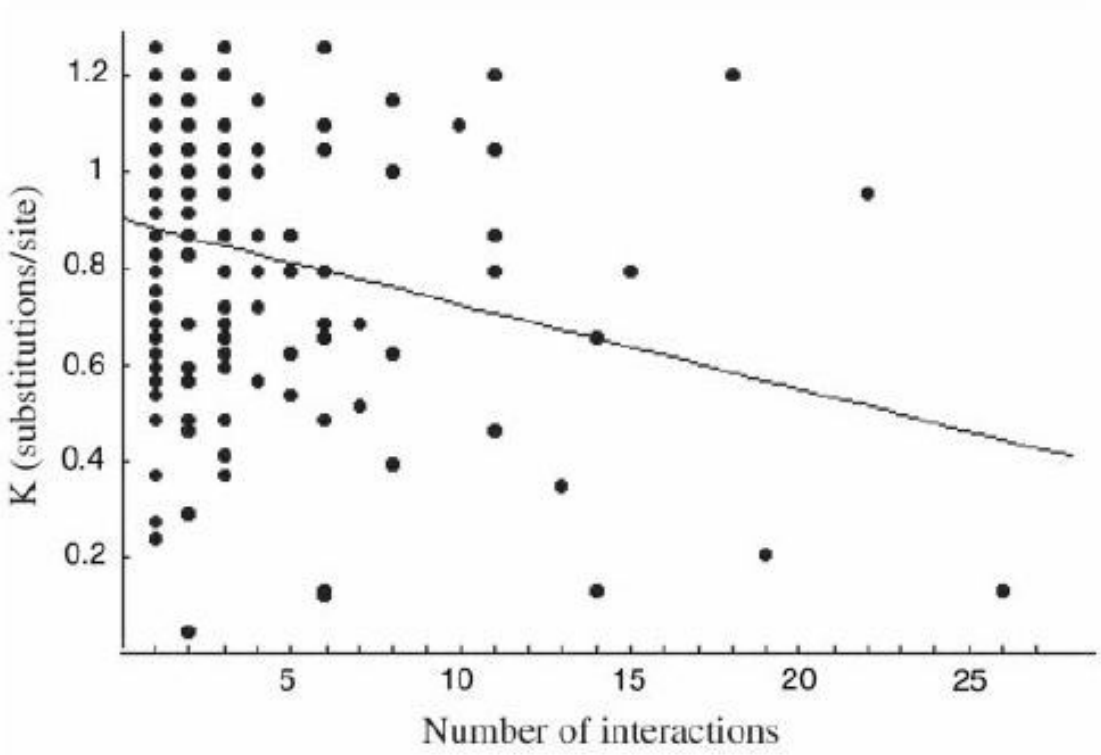
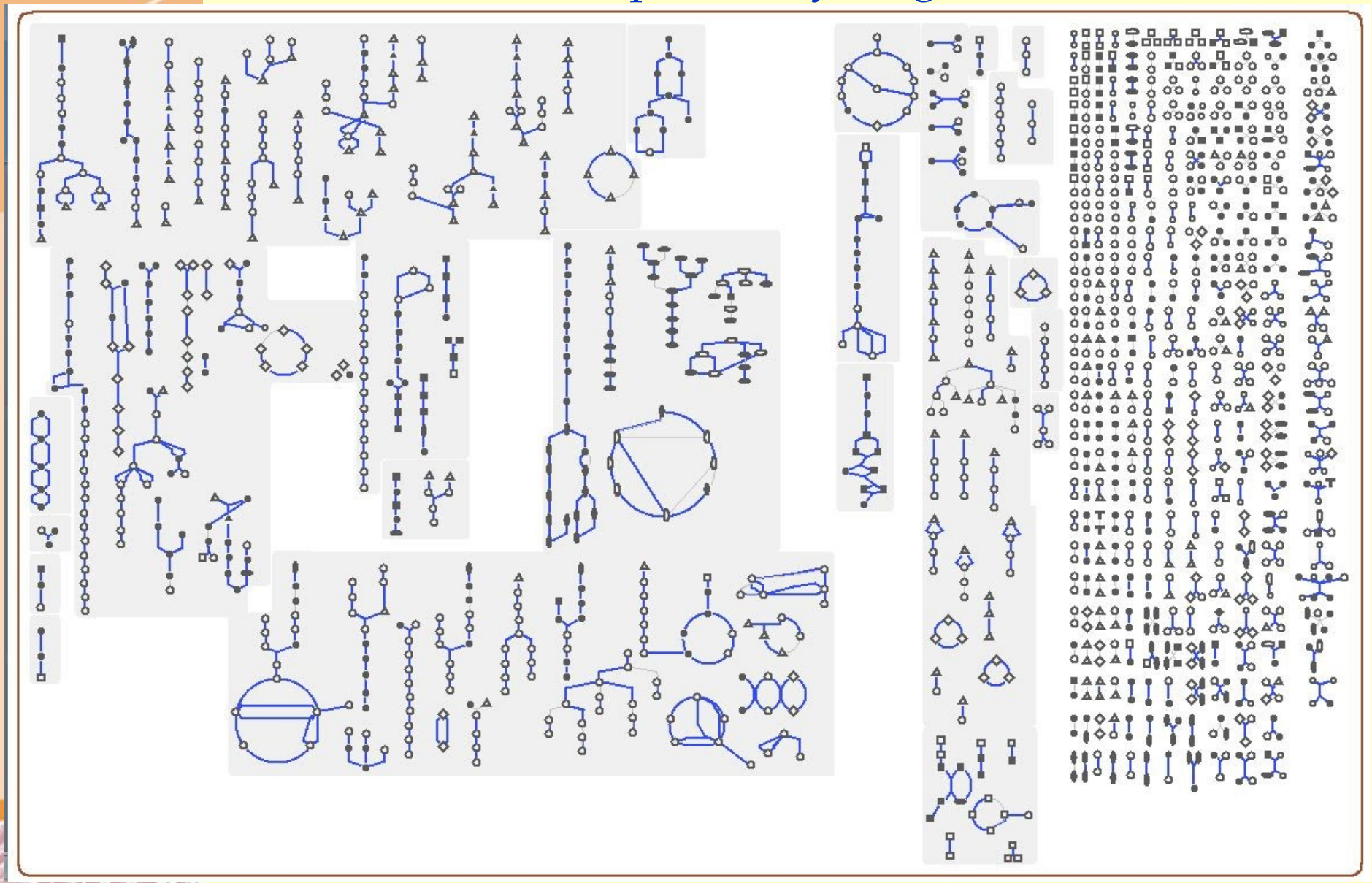


Fig. 1. The relation between the number of protein-protein interactions (*I*) in which a yeast protein participates and that protein's evolutionary rate, as estimated by the evolutionary distance (*K*) to the protein's well-conserved ortholog in the nematode *C. elegans*.

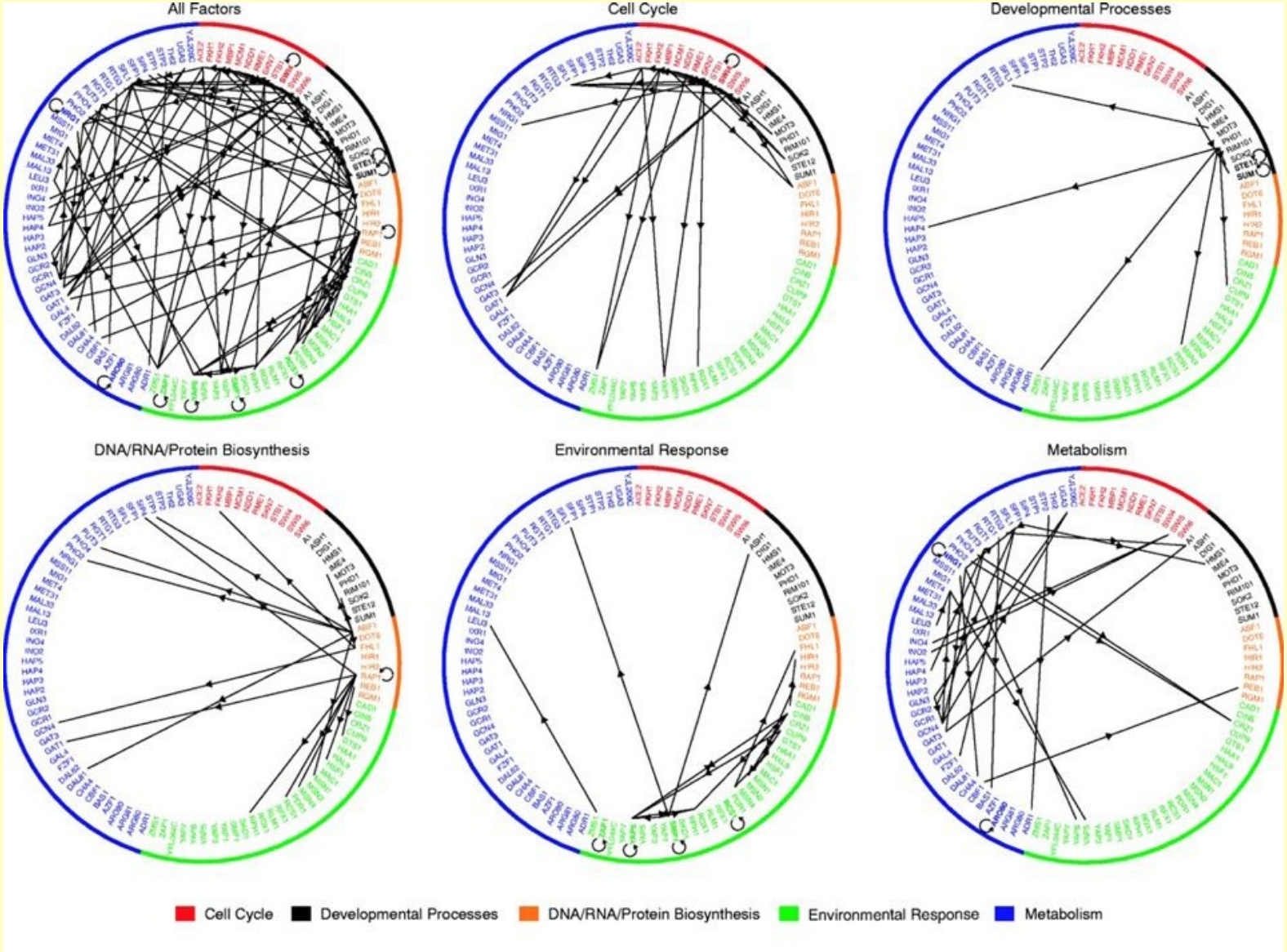
Fraser, et al. Science 296, p750 (2002)

Yeast Metabolic Pathways

http://biocyc.org/



Yeast Gene Regulatory Networks

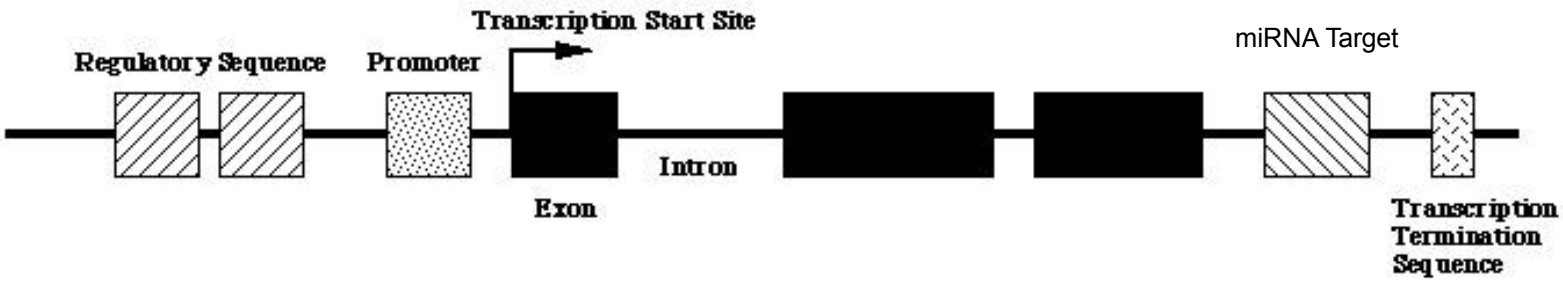


Lee et al. Science 298, 799-804 (2002)

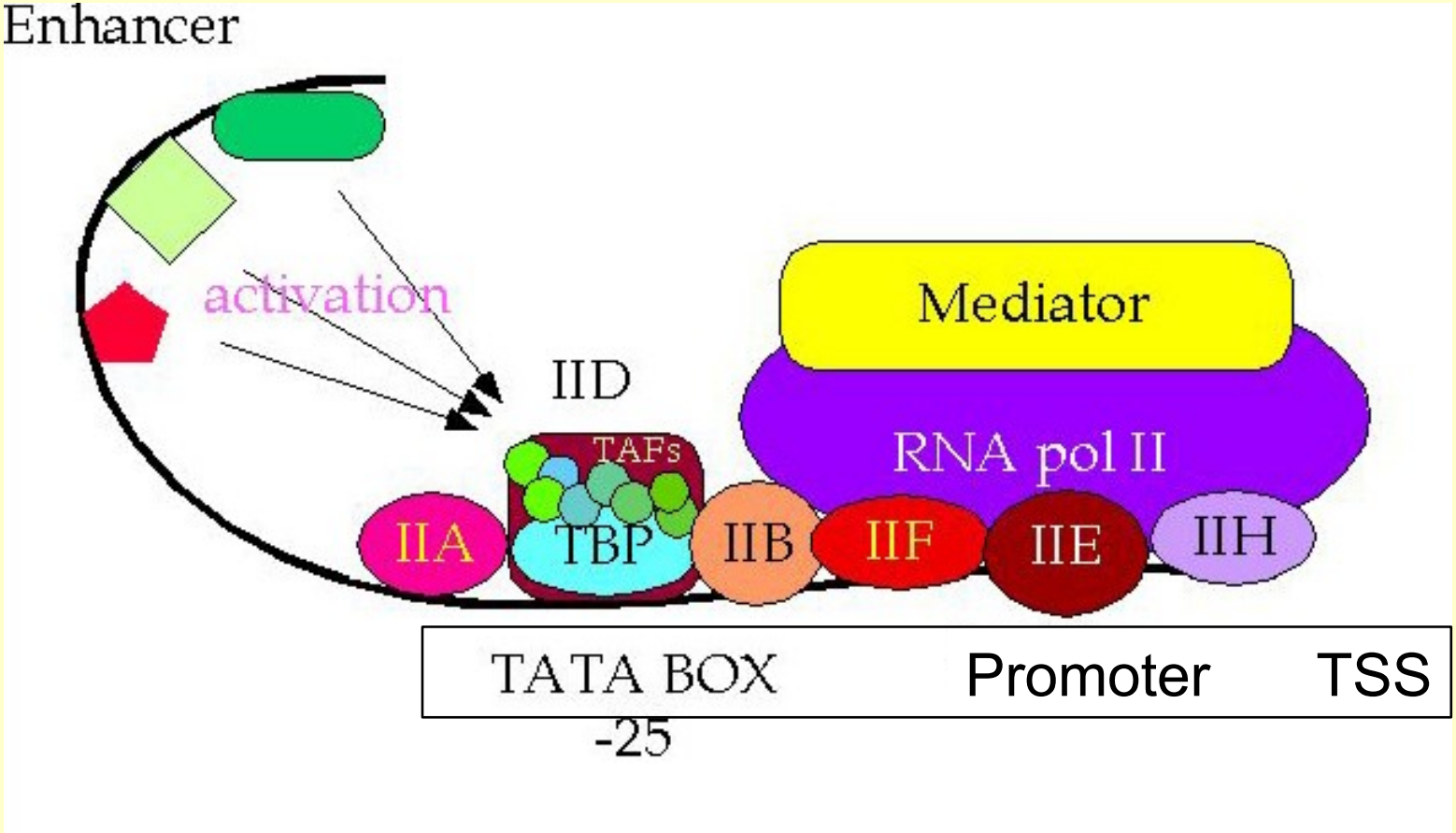
Genomics Provides

- Location and number of genes
  - By mapping messenger RNAs (cDNAs or ESTs)
  - By computational gene finding programs
    - Glimmer
    - GenScan
    - Grail & Grail EXP
    - FGENESH
    - GeneMark & GenMarkHMM
  - By homology and conservation
- Gene expression and regulation information
  - From tissue specific mRNAs (cDNAs or ESTs)
  - From DNA chips or microarrays
    - Gene probes and tiling arrays
  - ChIP-Chip experiments
  - Stanford Microarray Database
  - NCBI GEO
  - EBI ArrayExpress
- Gene regulatory sequences
  - Promoter signals near beginning (5′ end) of transcripts
  - Transcription factor binding sites
  - Transcription start sites (TSS)
  - Transcription termination signals (TTS)
  - RNA processing signals
    - Splice donor and acceptor sites
    - 5′ Cap site
    - 3′ polyAdemylation

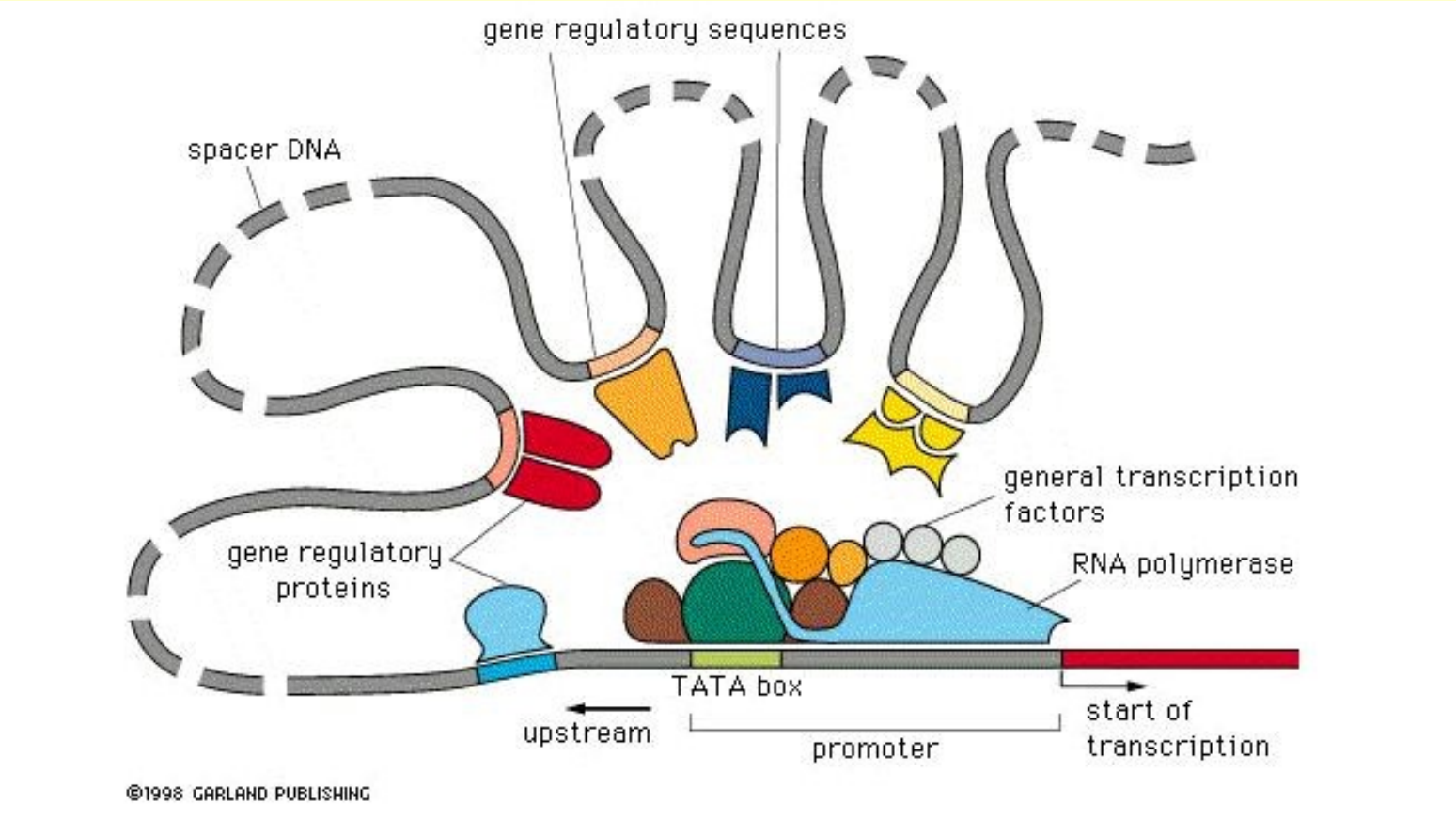
Eukaryotic Gene Structure



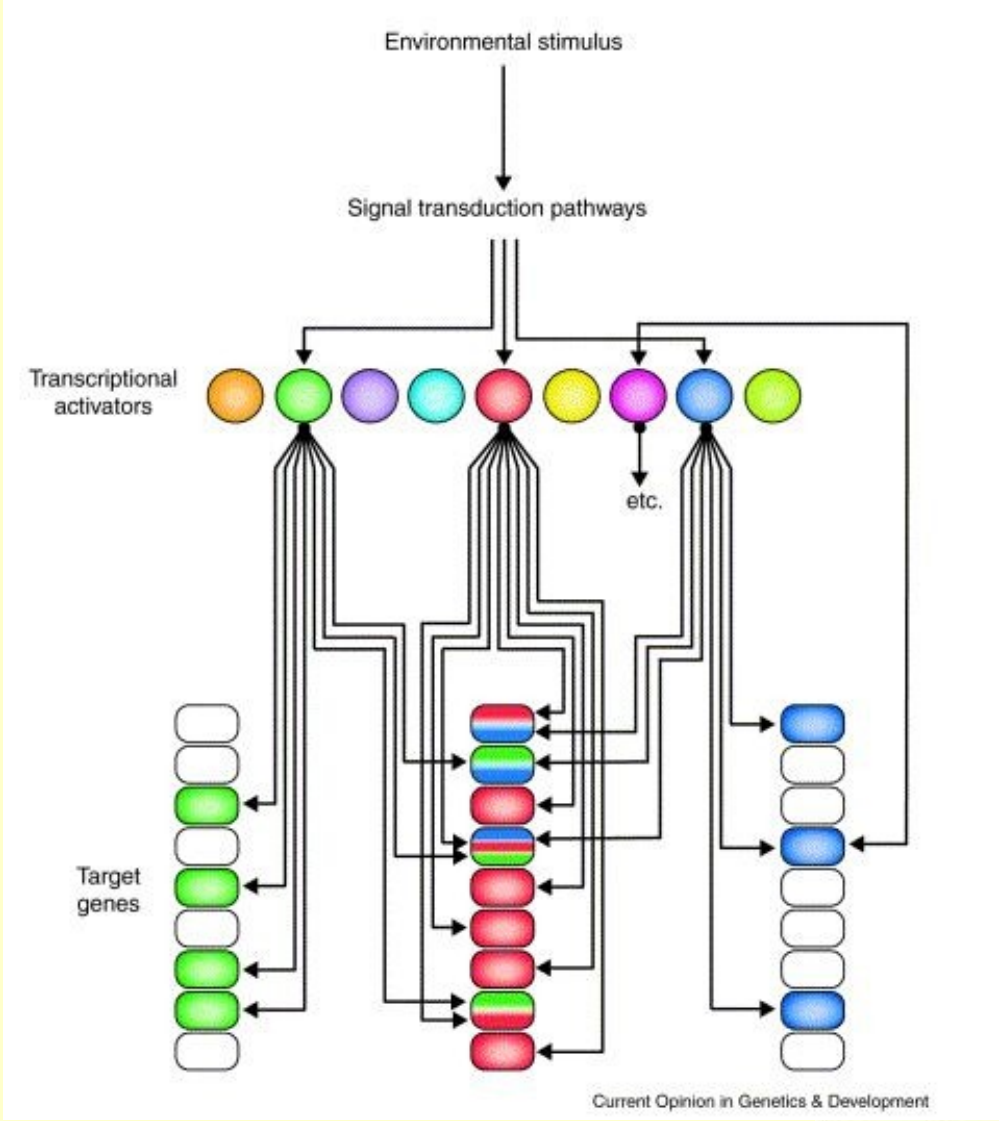
RNA Polymerase Promoter Binding



Multiple Enhancer Sequences



Gene Expression Regulatory Network



Enhanceosome

http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb122_1.html

February 2010 **Molecule of the Month** by David Goodsell
Previous Features
doi: **10.2210/rcsb_pdb/mom_2010_2**

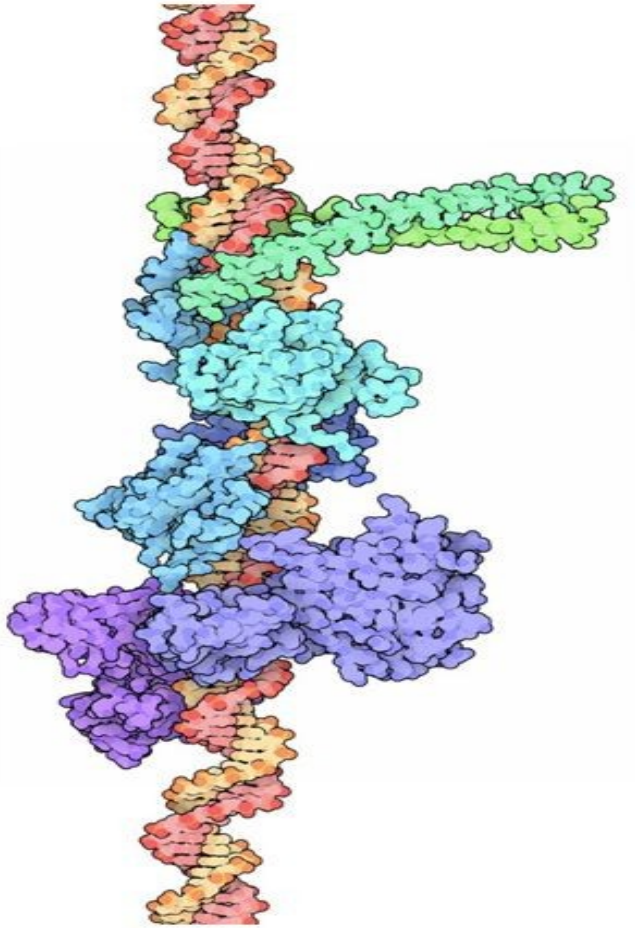
Enhanceosome

keywords: transcription factor, gene expression, CBP, CREB-binding protein, transcriptional enhancers, enhanceosome

Take a moment to ponder the form of your body: the shape of your face, the color of your eyes, the length of your fingers, the perfect articulation of your bones and muscles, the way your hair grows curly or straight. Now let your imagination travel inward, and think of the complex shapes and functions of your different cells, and the teeming molecular world inside each one. Remarkably, this amazing structure and form and function is specified by information in the genome, which encodes a mere 20,000-25,000 protein-coding genes. One of the great puzzles being pieced together by scientists is the mechanism by which these genes, and the methods used to control their expression, specify all of these different aspects of life.

Combinatorial Control

In order to specify which gene will be expressed in a given situation, your cells use a diverse collection of DNA-binding proteins to control access to the DNA. Surprisingly, there are relatively few of these proteins: by some estimates, the human genome encodes about 2,600 of them. But then, the capabilities of this limited set are greatly expanded by using them in combination, by requiring two or more to bind simultaneously to activate a gene. In this way, each protein may be used in many ways and the spectrum of responses is far more varied.



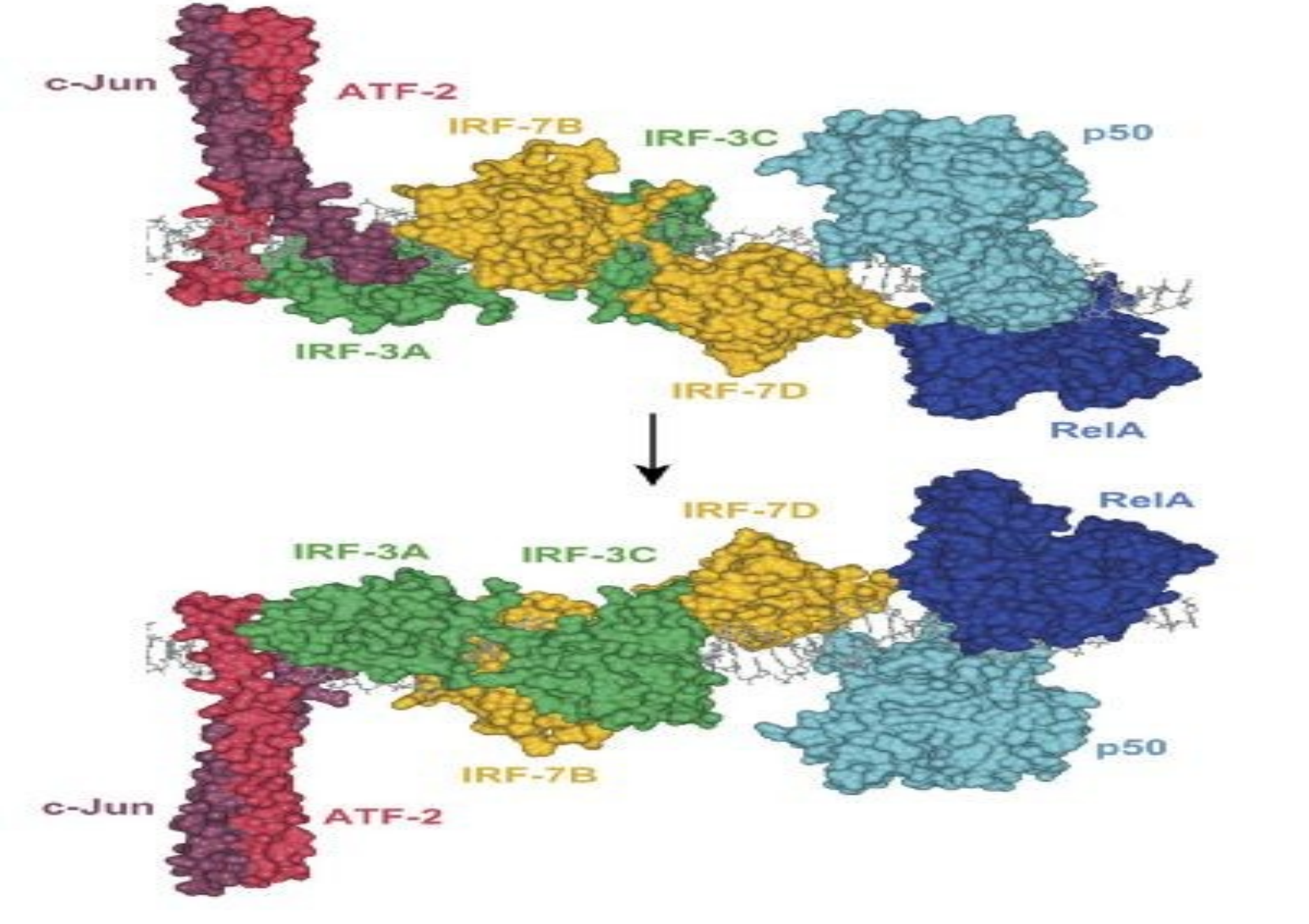
Enhancing Transcription

The assembly of DNA and proteins pictured here is a transcriptional enhanceosome (PDB entries **1t2k**, **2pi0**, **2o6g** and **2o61**) that controls expression of interferon-beta, an important protein for fighting viral infection. When the cell is infected by viruses, several different DNA-binding proteins are produced, including ATF-2/c-Jun (in green at the top), interferon response factors (IRF, shown in turquoise at the center), and nuclear factor kB (NF-kB, shown in blue and magenta at the bottom). Individually, each one is not sufficient to activate the gene, and each one also plays other roles in the activation of other genes (for instance, Nf-kB is also important in immune responses, inflammation, apoptosis, and many other processes). But when they all bind together, they activate the gene and interferon is made.

Next: **Integrating the Signal**

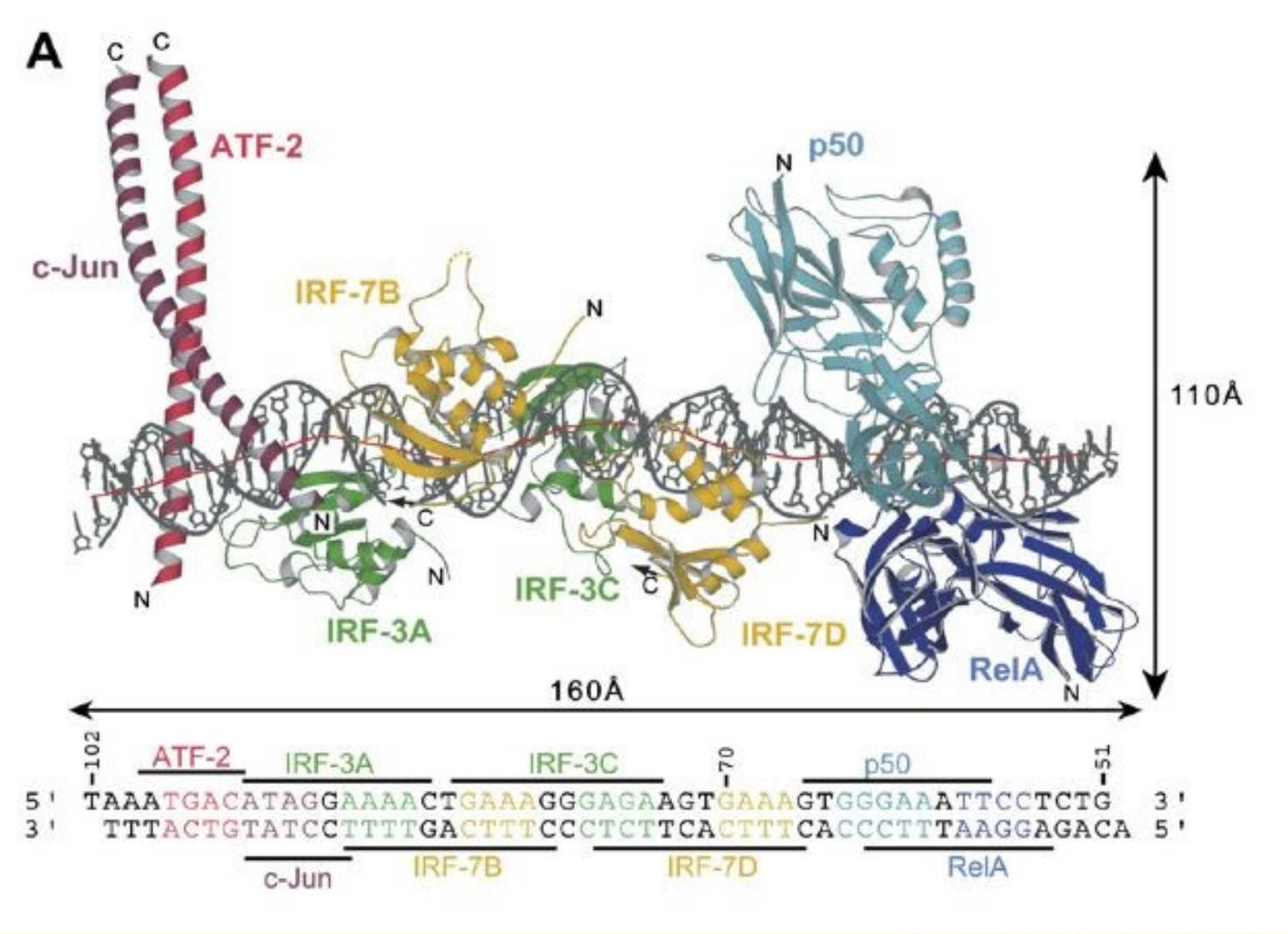
Enhanceosome

http://www.cell.com/retrieve/pii/S0092867407006563



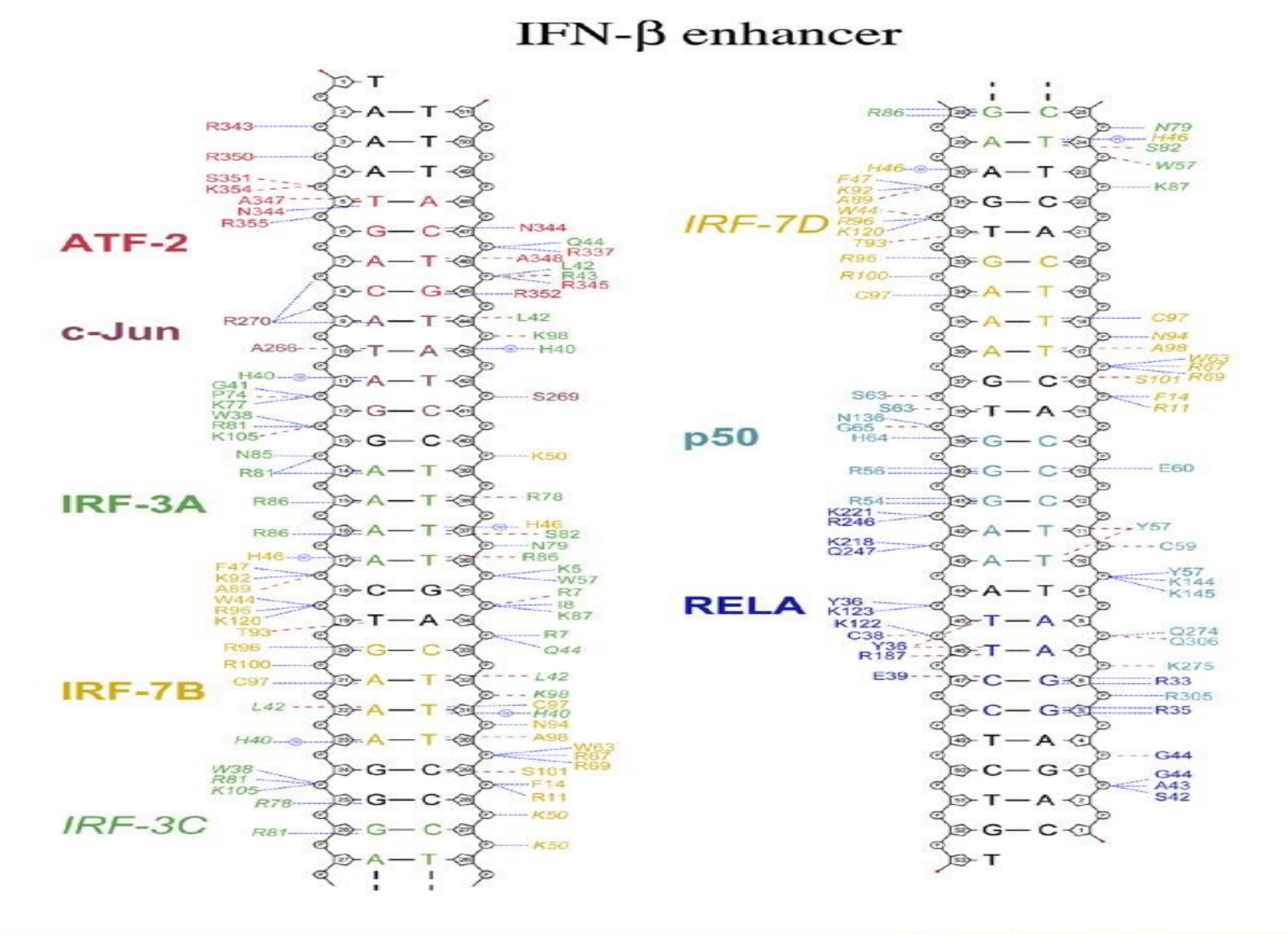
Enhanceosome

http://www.cell.com/retrieve/pii/S0092867407006563



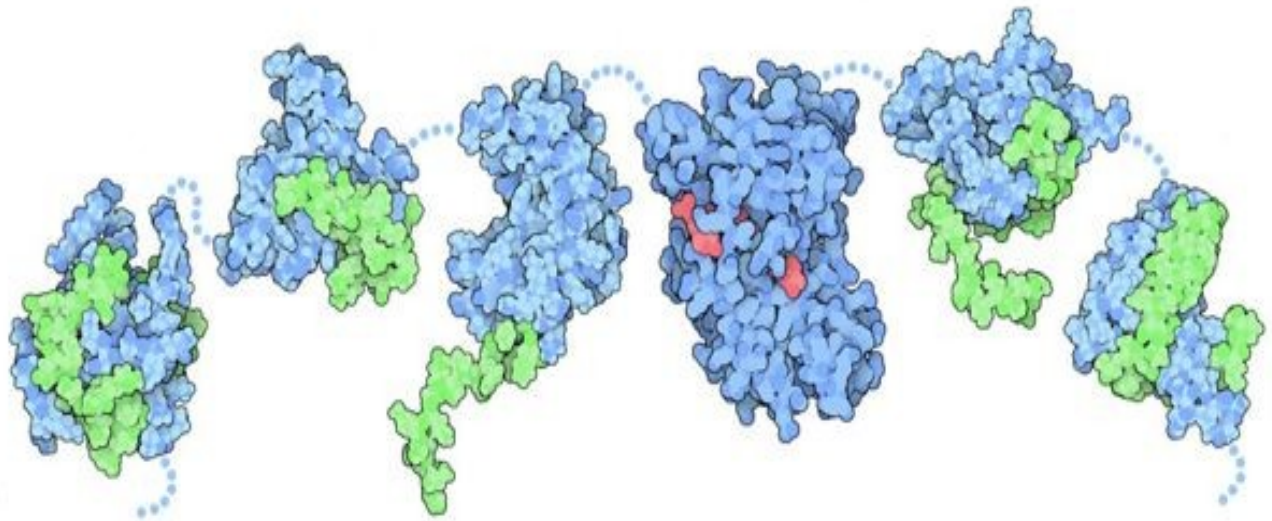
IFN-beta Enhancer

http://www.cell.com/retrieve/pii/S0092867407006563



Enhanceosome: Integrating the Signal

http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb122_2.html



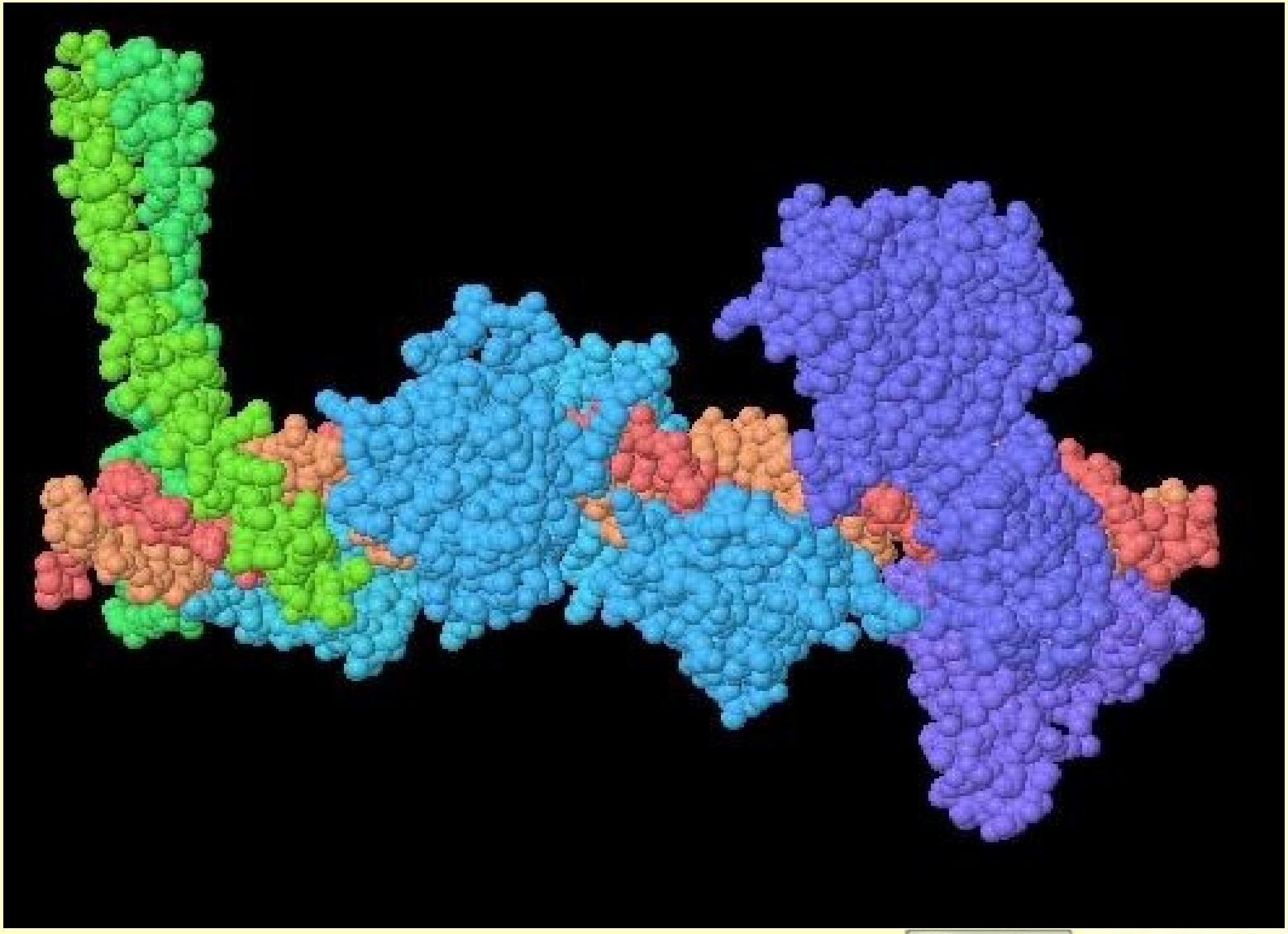
Integrating the Signal

Once the transcription factors bind to the different sites in the enhancer DNA sequence, the signal must somehow be sensed and used to activate the gene. In many cases, this is performed using CREB-binding protein or the similar protein p300. This protein is composed of many connected domains, (PDB entries **1l8c**, **1kdx**, **1jsp**, **3biy**, **2ka6** and **1kbh**), which bind to different proteins in the assembled enhanceosome. Then, a large domain in the center acts as a histone acetyltransferase, modifying histones in nucleosomes and causing them to disassemble and reveal the gene. In the intereferon-b gene, a nucleosome normally covers the start site of transcription, blocking transcription. Assembly of the enhanceosome leads to removal of this nucleosome, allowing the gene to be expressed.

Previous: **Enhanceosome** Home: **Enhanceosome** Next: **Exploring the Structure**

Enhanceosome: Exploring the Structure

http://www.rcsb.org/pdb/static.do?p=education_discussion/molecule_of_the_month/pdb122_2.html



BIOCARTA

FEATURES | PATHWAYS | CUSTOM SERVICES | GENES | PRODUCTS | LOGIN

Interactive graphic models of molecular and cellular pathways

ht asp

PATHWAYS ▸ The information-processing pathway at the IFN-beta enhancer

Submitted by: Guru:

COMMENT ON THIS PATHWAY | DESCRIPTION | CONTRIBUTORS | SAVE THIS LINK | SUBMIT | LEGEND

PRODUCT INDEX

PRODUCT SEARCH

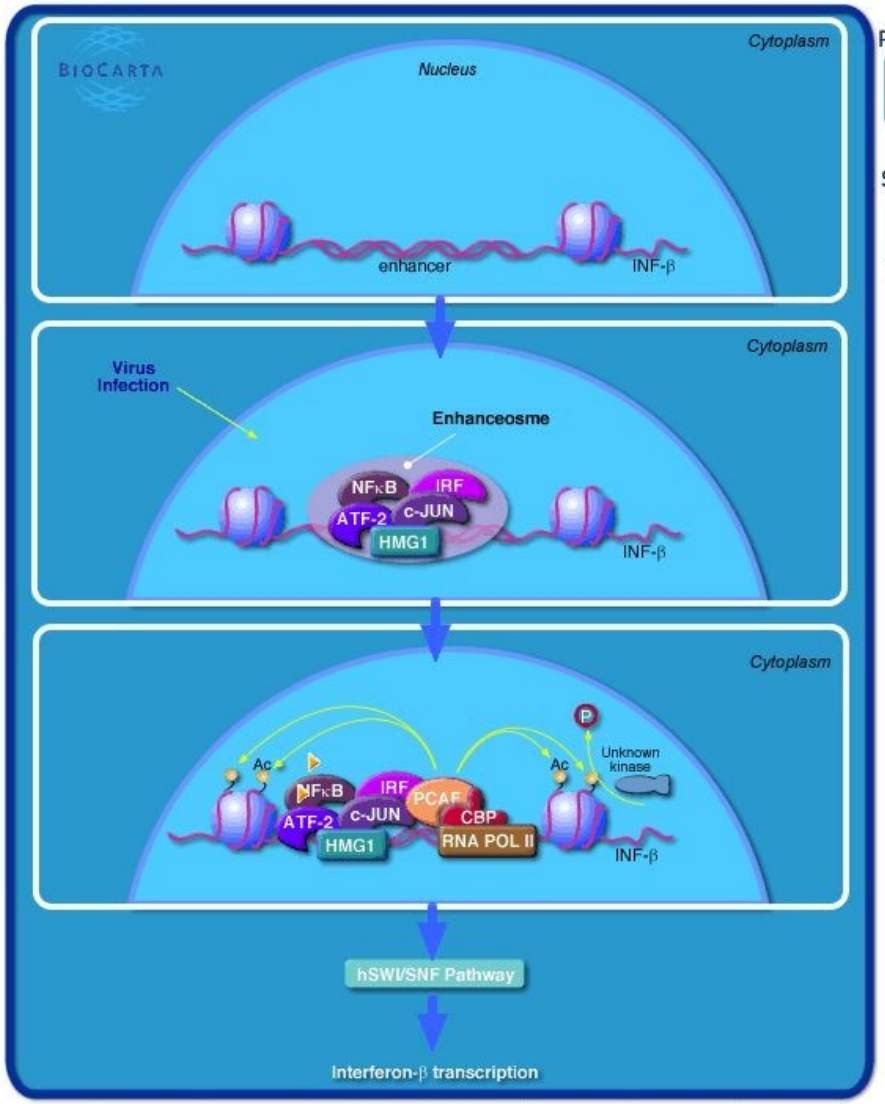
SEARCH

Contains
Exact
Advanced Search

PRODUCT HIGHLIGHT
ON OFF

PROTEIN LIST

REQUEST A CATALOG



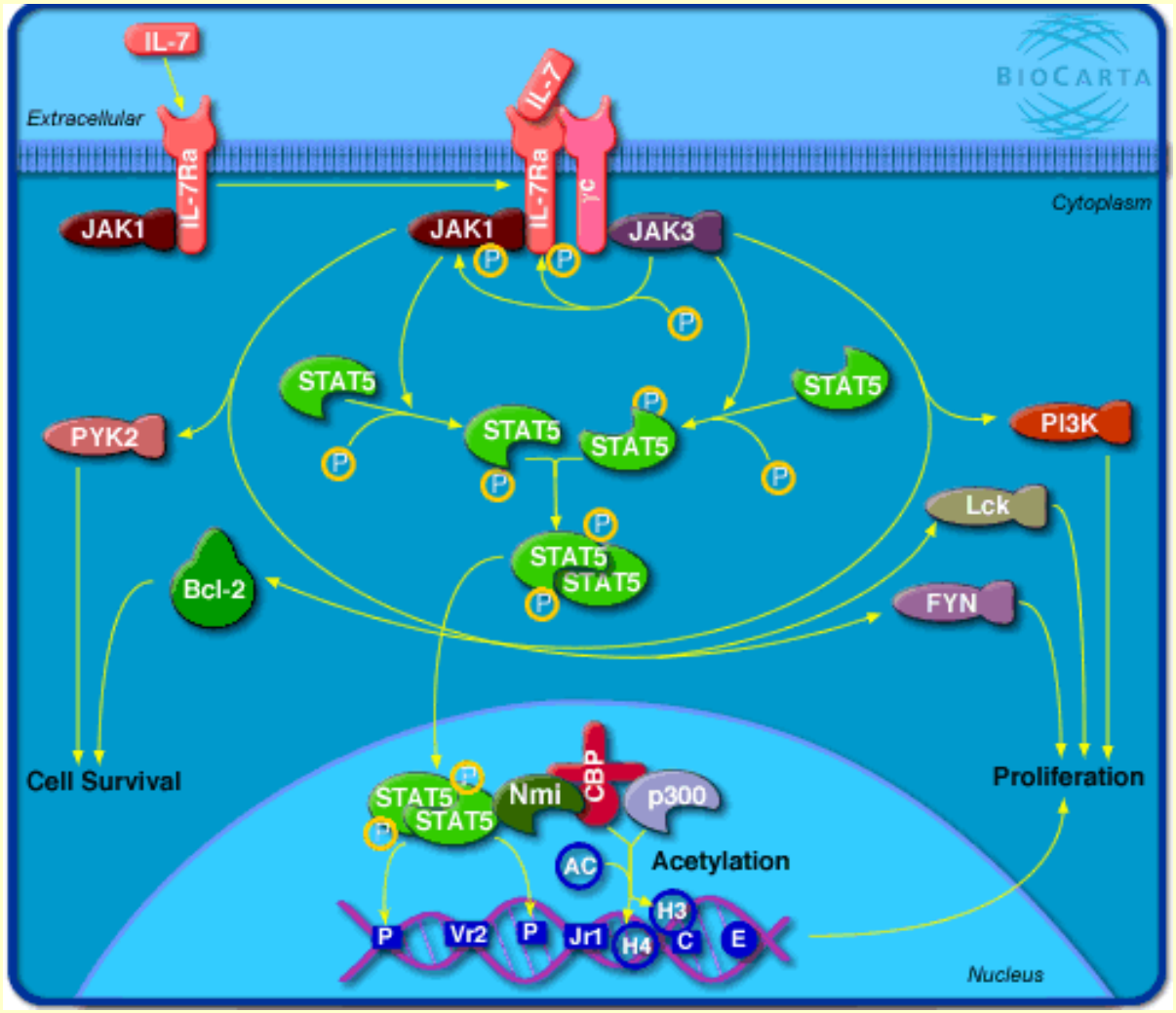
This Pathway:

Other Species:

arta

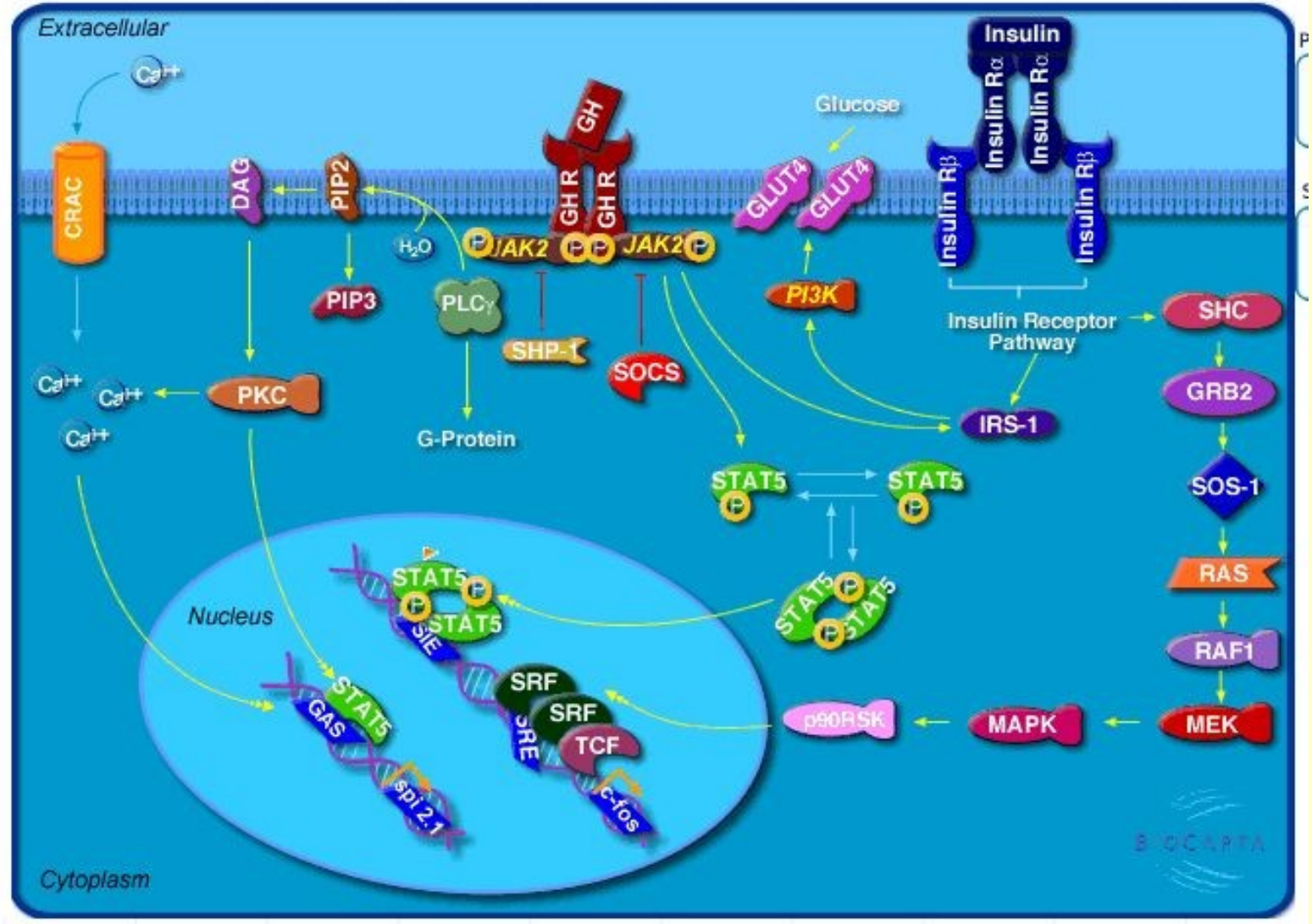
IL7 Regulatory Pathway

http://www.biocarta.com/pathfiles/h_il7Pathway.asp/



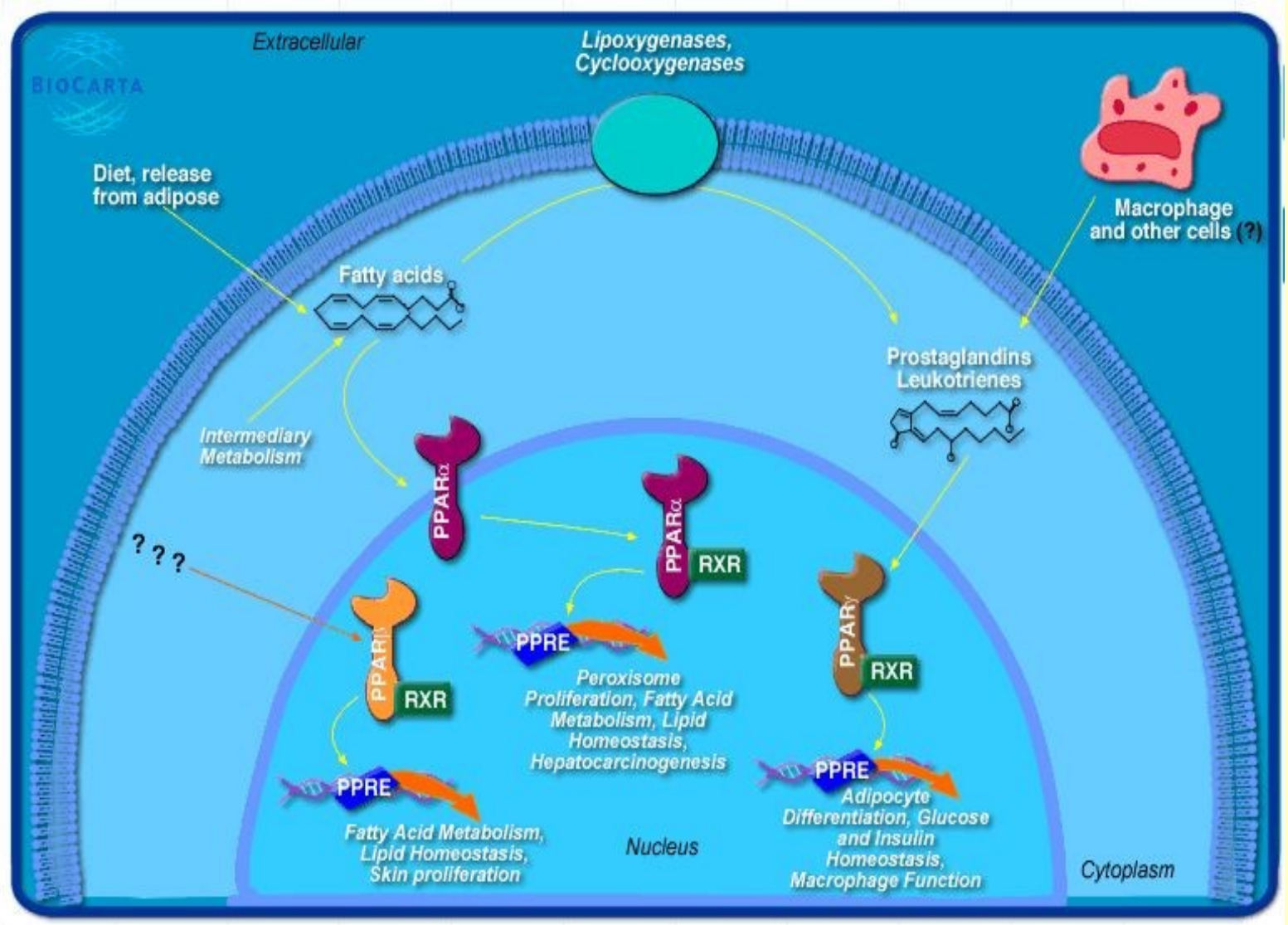
Growth Hormone Receptor Pathway

http://www.biocarta.com/pathfiles/h_ghPathway.asp/



PPAR-Mechanism of Action

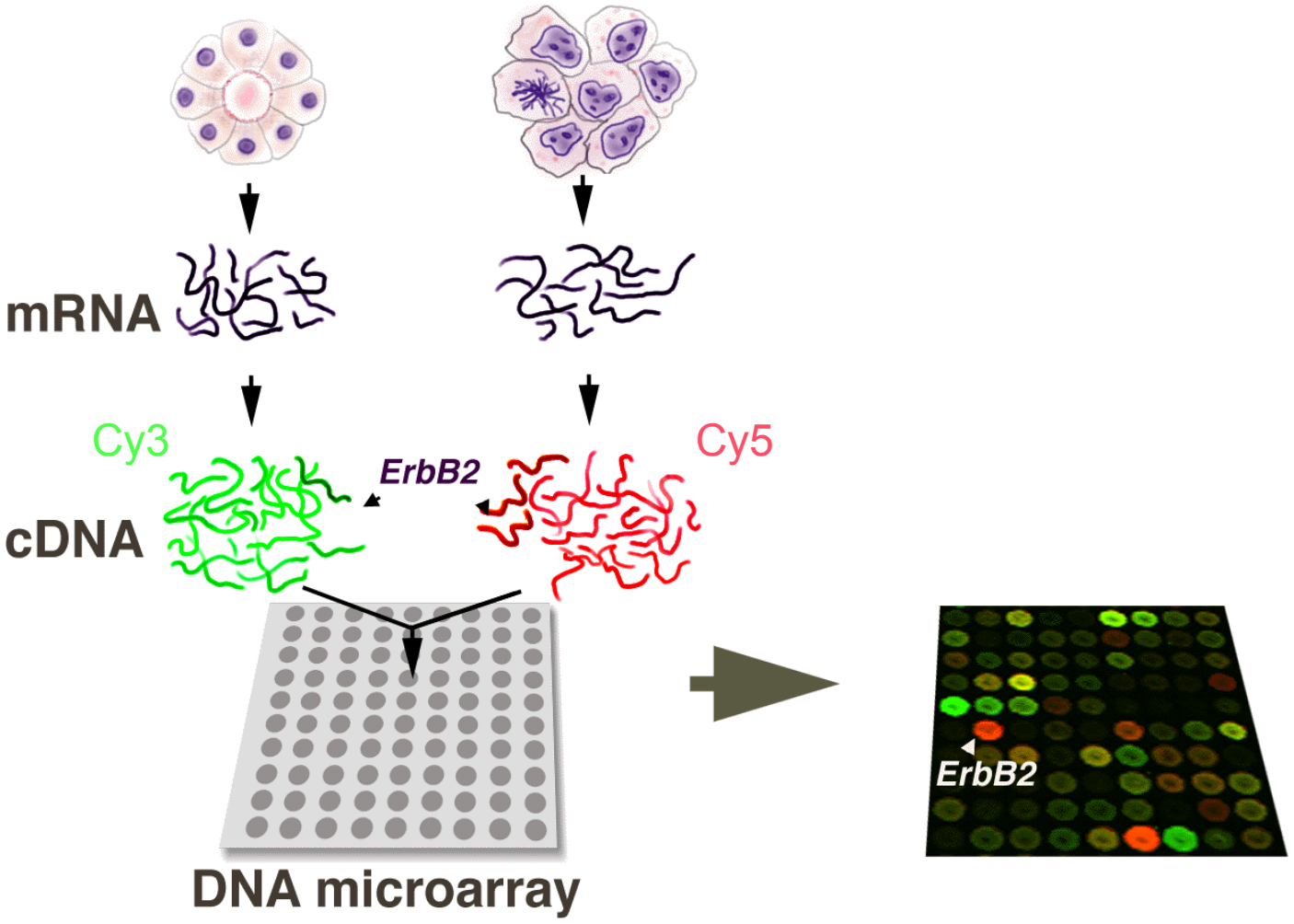
http://www.biocarta.com/pathfiles/h_pparPathway.asp



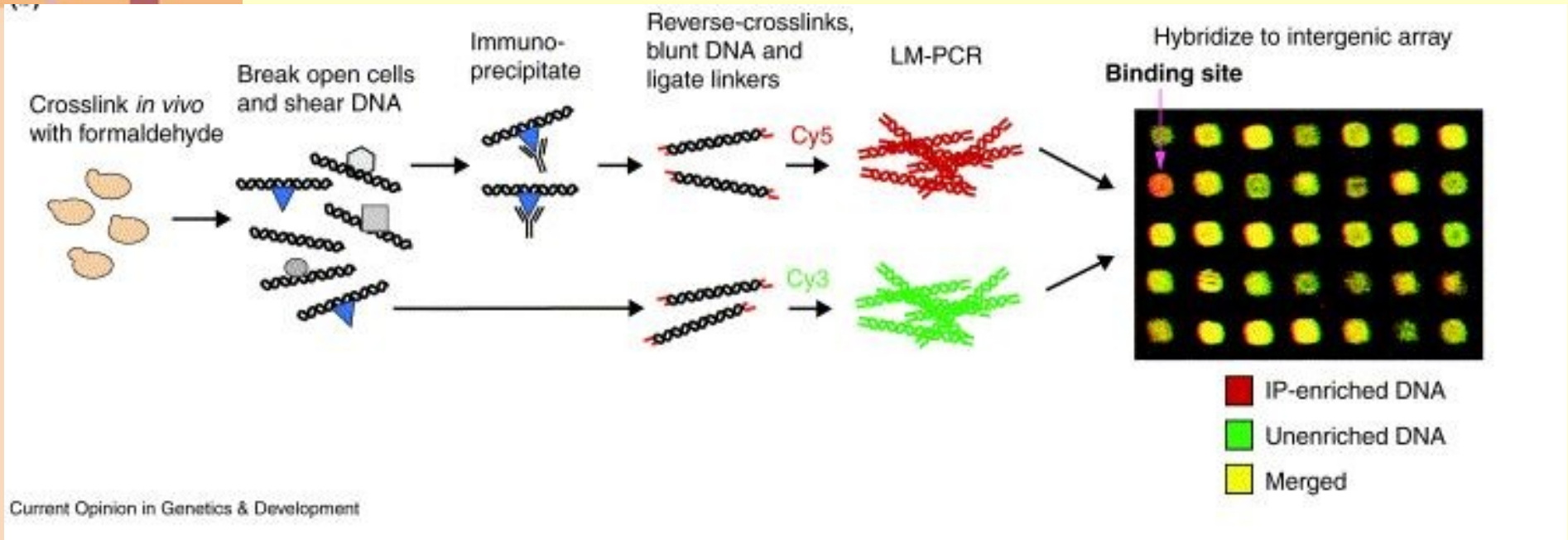
Gene Regulatory Mechanisms

- Transcriptional Mechanisms
  - Type of promoters & RNA polymerase
  - Control of Transcription
    - Constitutive
    - Inducible
    - Repressible
  - Transcription Factors and TFBS
- RNA processing
  - Capping
  - Splicing and Alternative Splicing
  - Poly-Adenylation
  - RNA export to cytoplasm
  - RNA degradation rates
- Translational Mechanisms
  - Micro RNAs (miRNAs)
  - Silencer RNAs (siRNAs or RNAi) degrading mRNA
- Epigenetic Mechanisms
  - Chromatin remodeling
  - Histone acetylation
  - DNA methylation

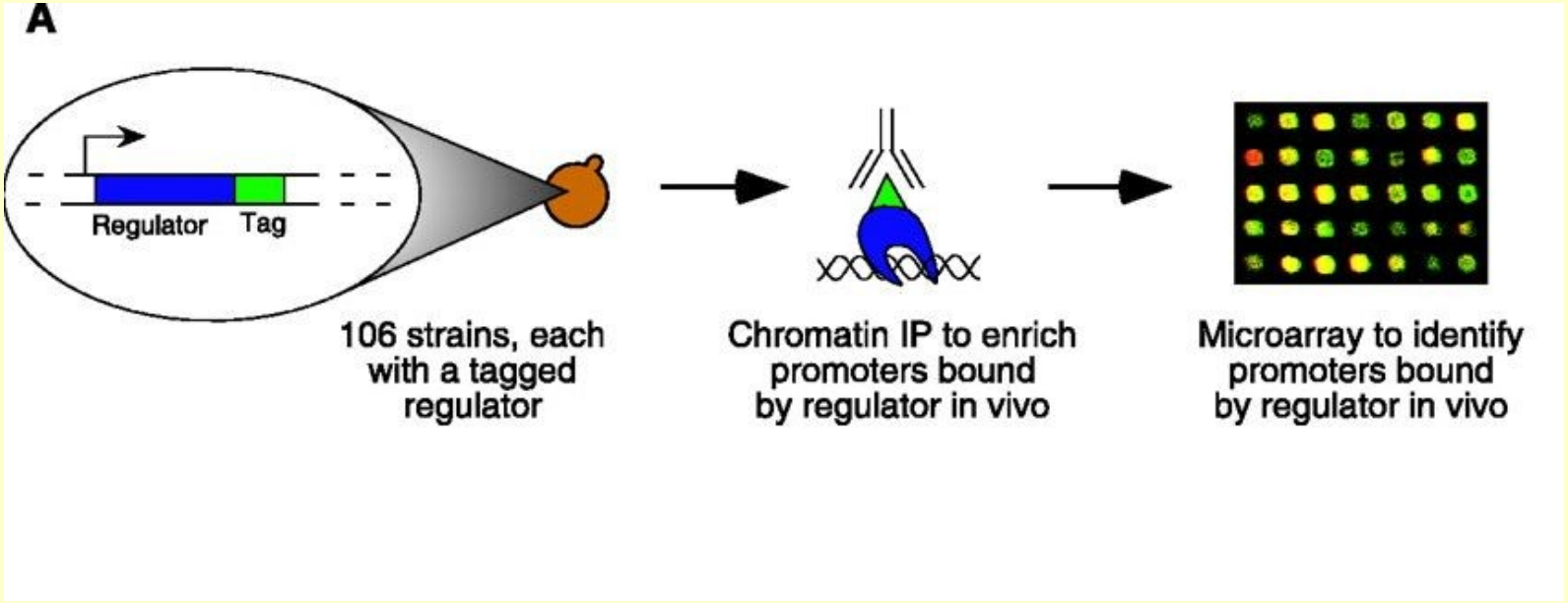
Measuring Gene Expression with DNA Microarrays and cDNA Labeling

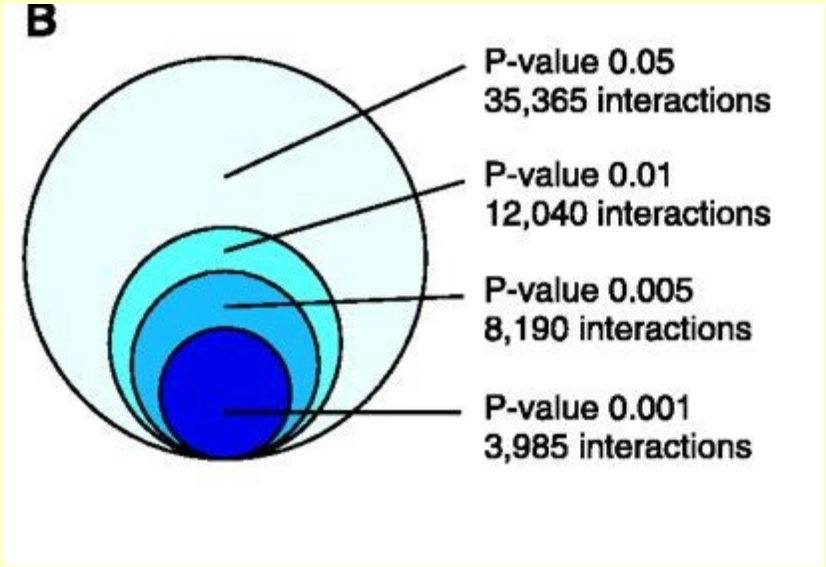


ChIP-Chip Methodology



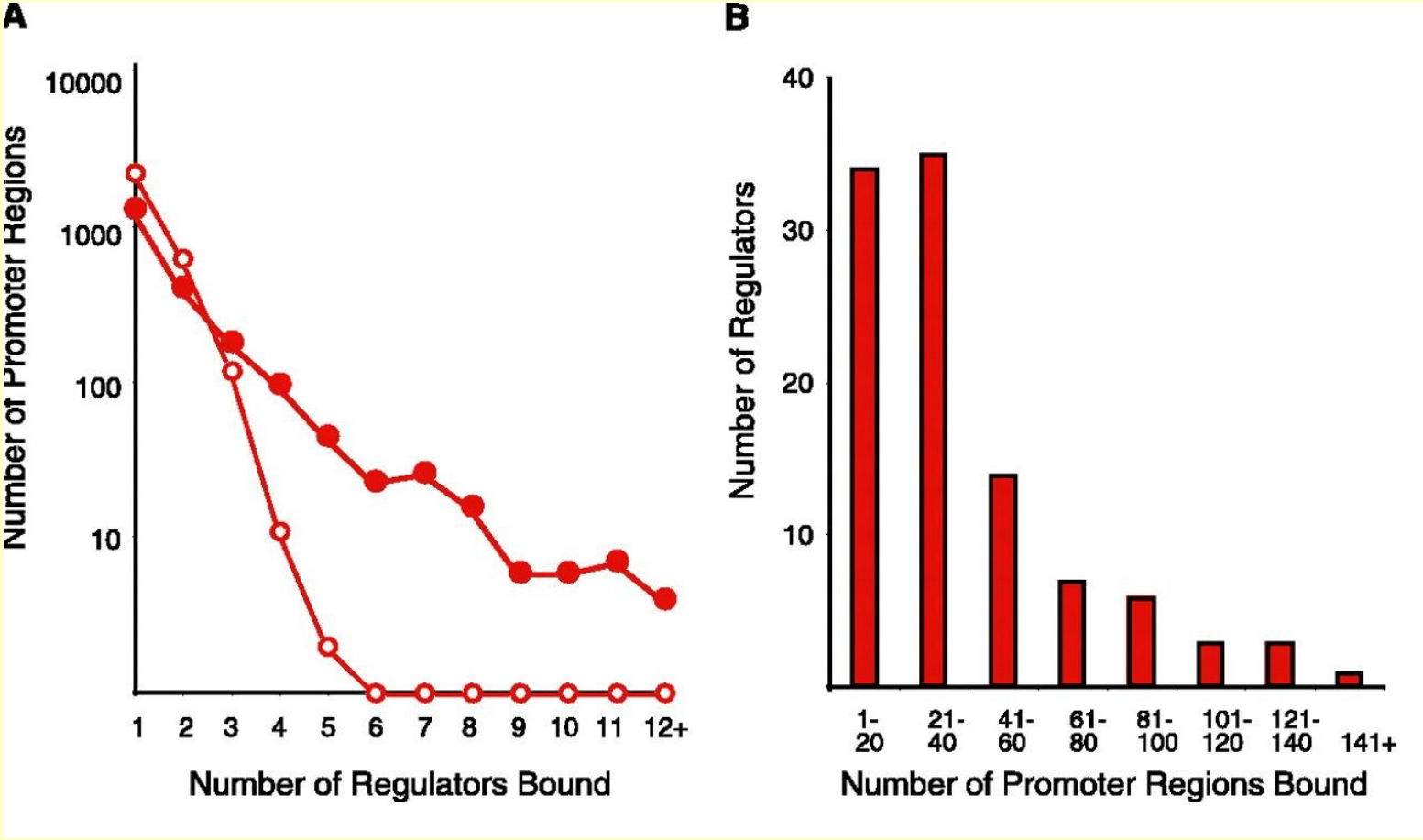
ChIP-Chip Identification of Transcription factor Binding Sites





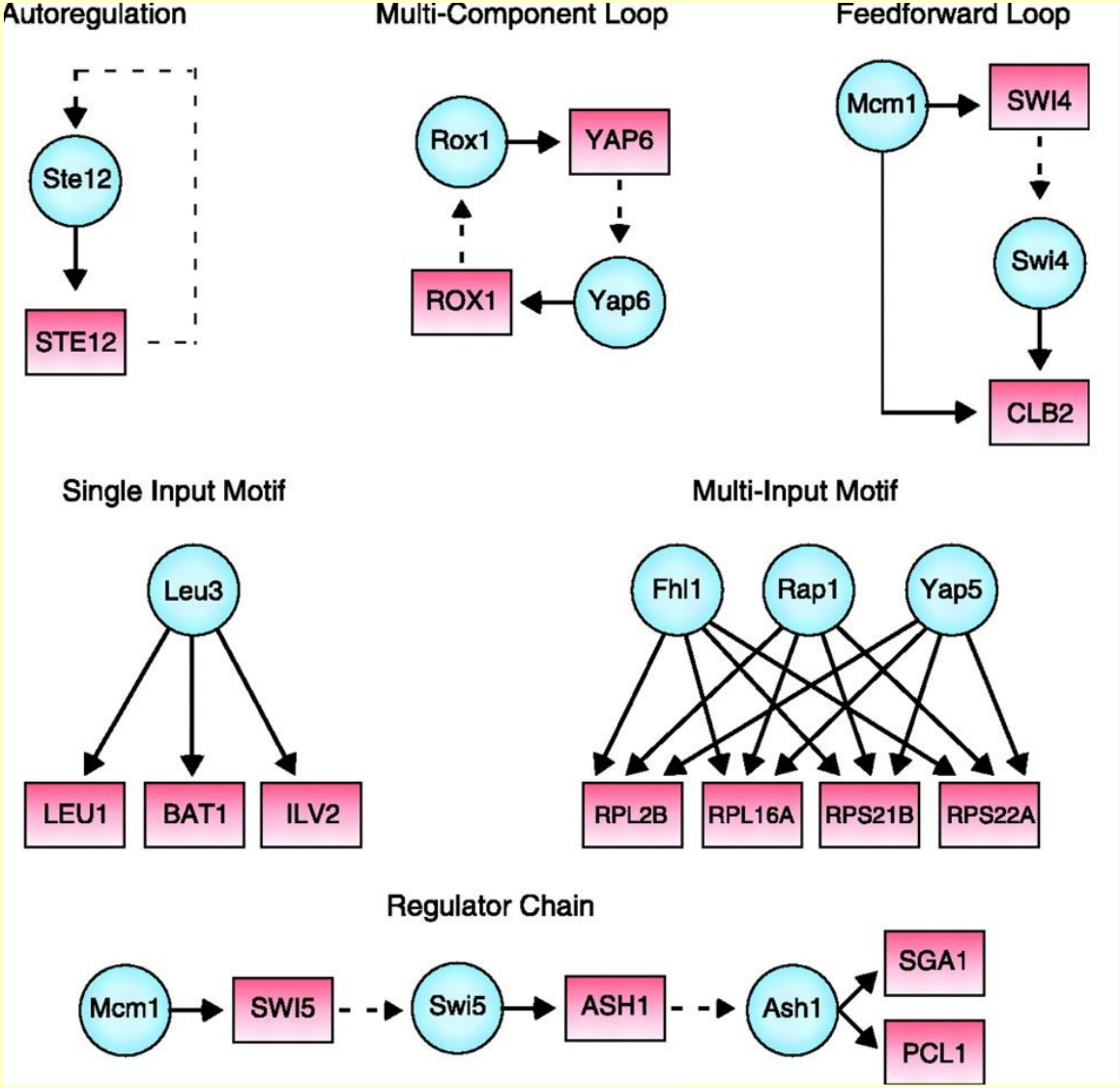
Lee et al. Science 298, 799-804 (2002)

Distribution of TFBS Among 6300 Promoters



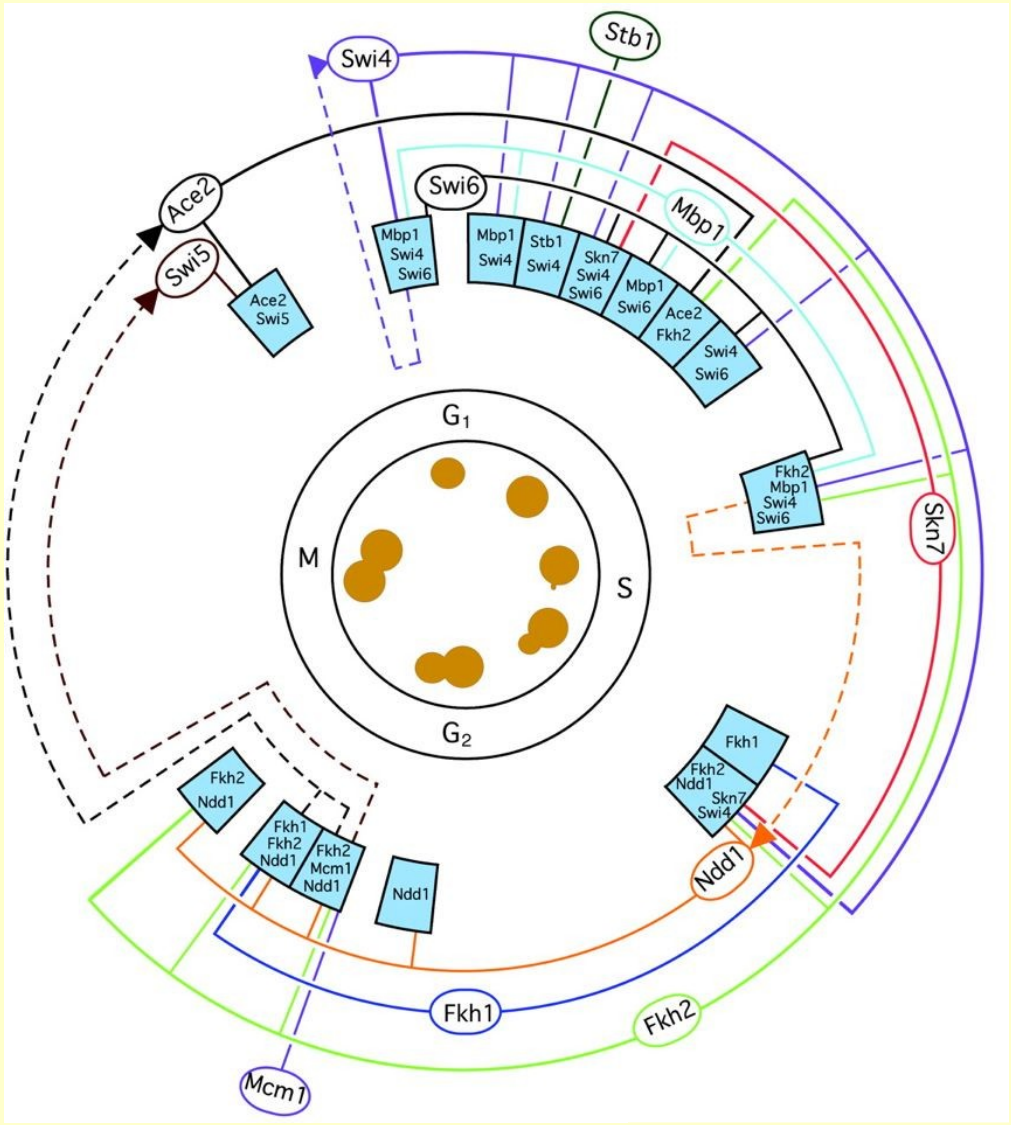
Lee et al. Science 298, 799-804 (2002)

Network Motifs in Yeast Regulatory Networks



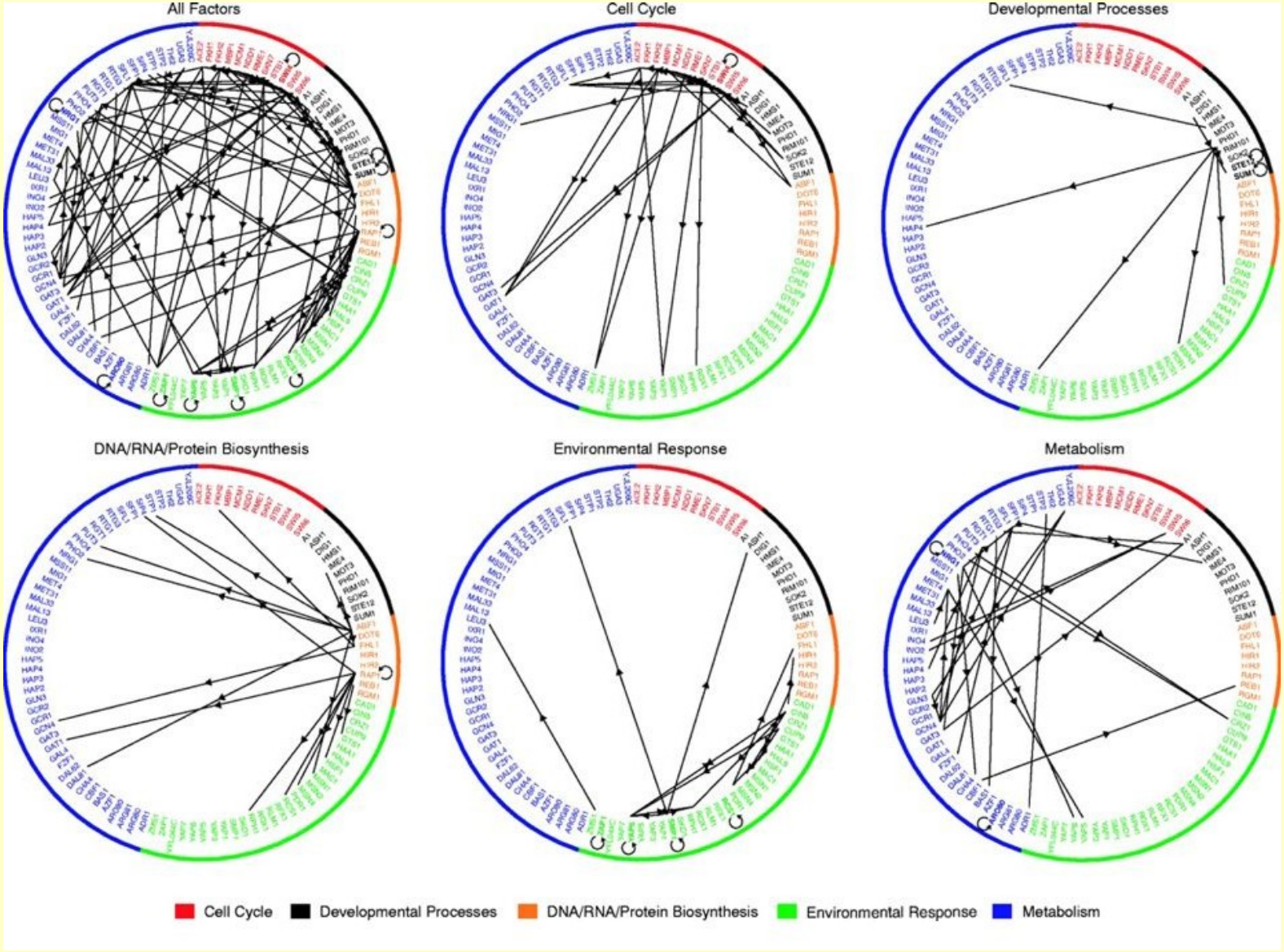
Lee et al. Science 298, 799-804 (2002)

Yeast Cell Cycle Regulatory Motifs



Lee et al. Science 298, 799-804 (2002)

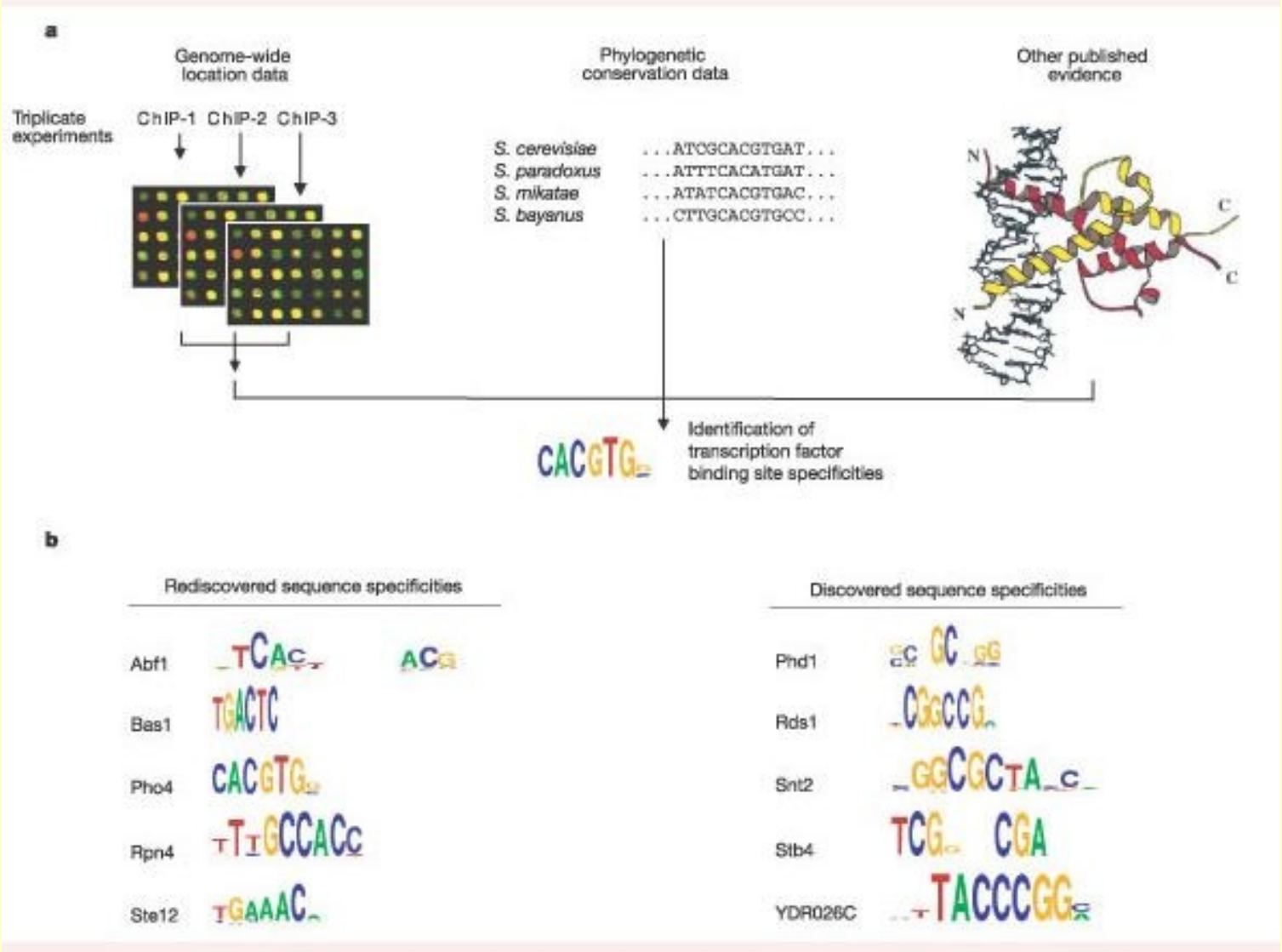
Yeast Regulatory Motifs



Lee et al. Science 298, 799-804 (2002)

Transcription Factor Binding Site Specificities

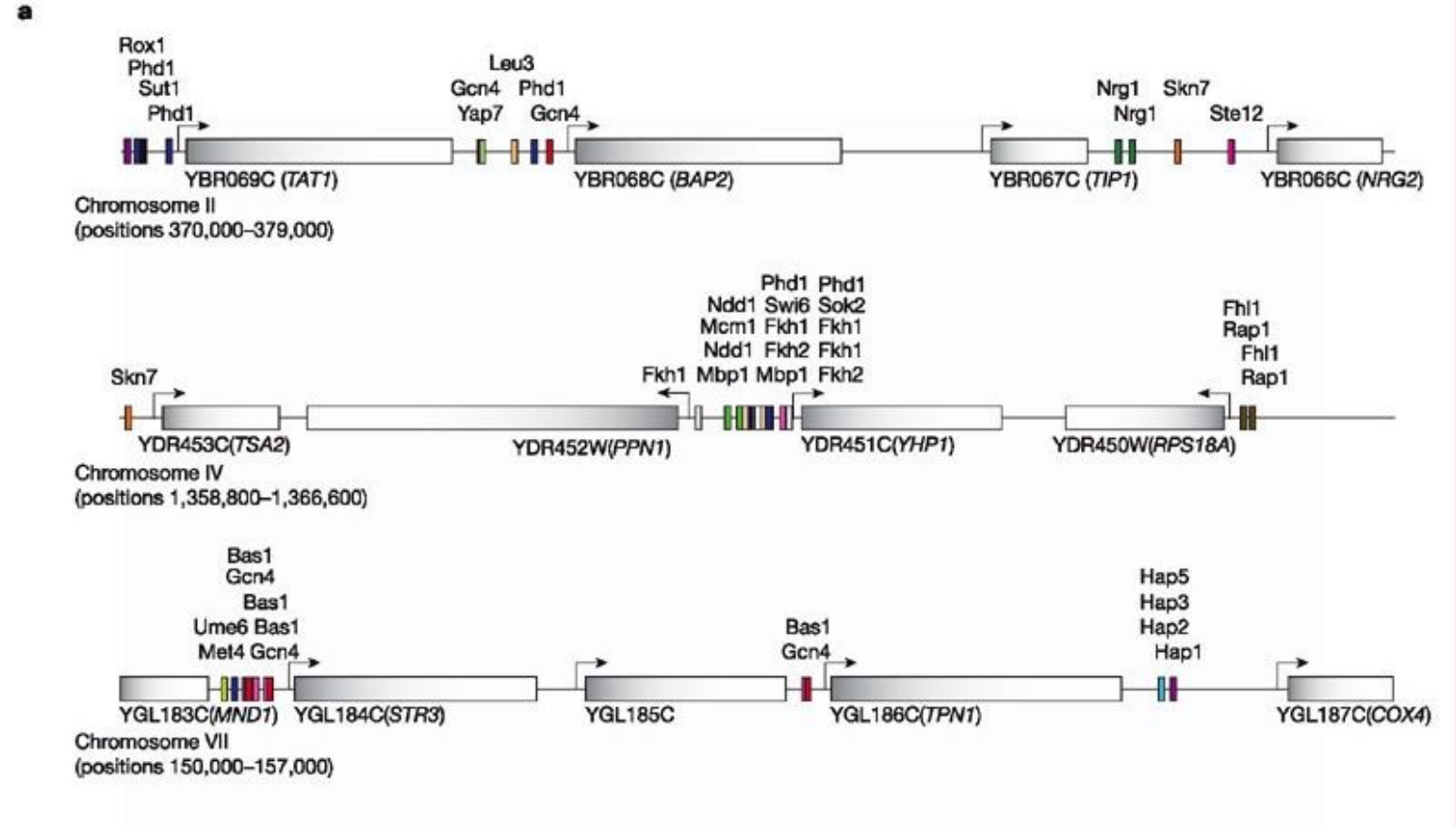
http://web.wi.mit.edu/young/regulatory_code/



Harbison et al, *Nature 431, 99-104 (2004)*

Yeast Regulatory Map

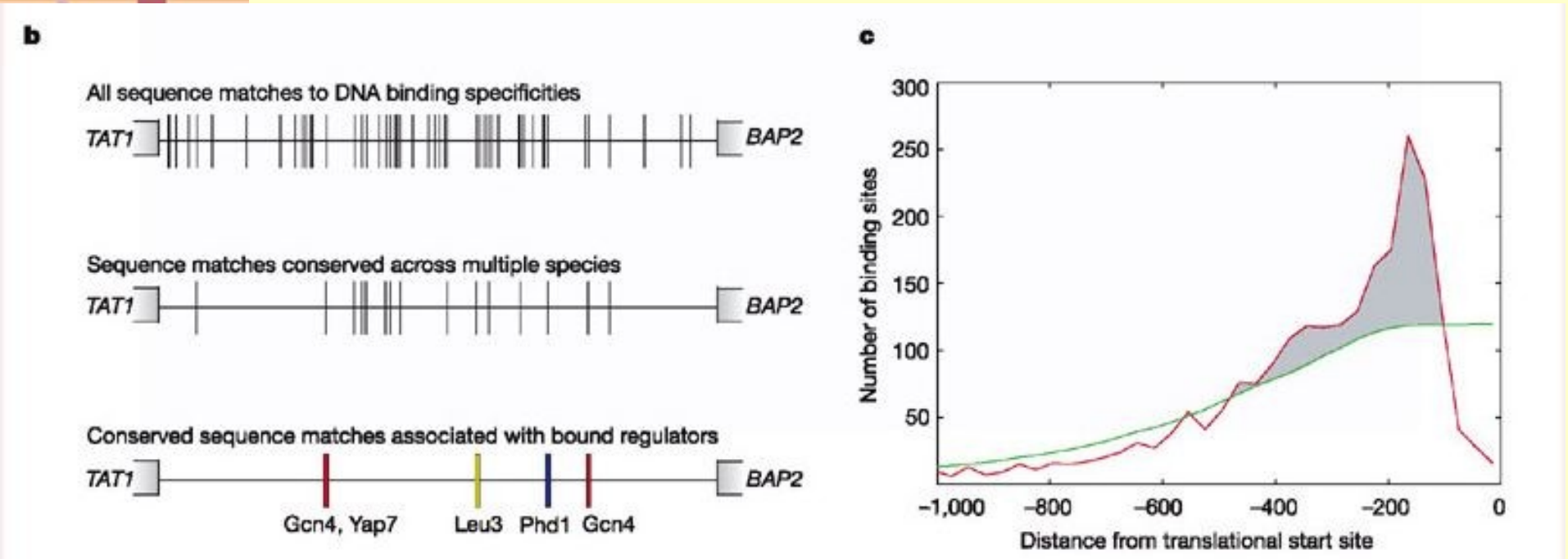
http://web.wi.mit.edu/young/regulatory_code/



Harbison et al, *Nature 431, 99-104 (2004)*

Yeast Regulatory Map

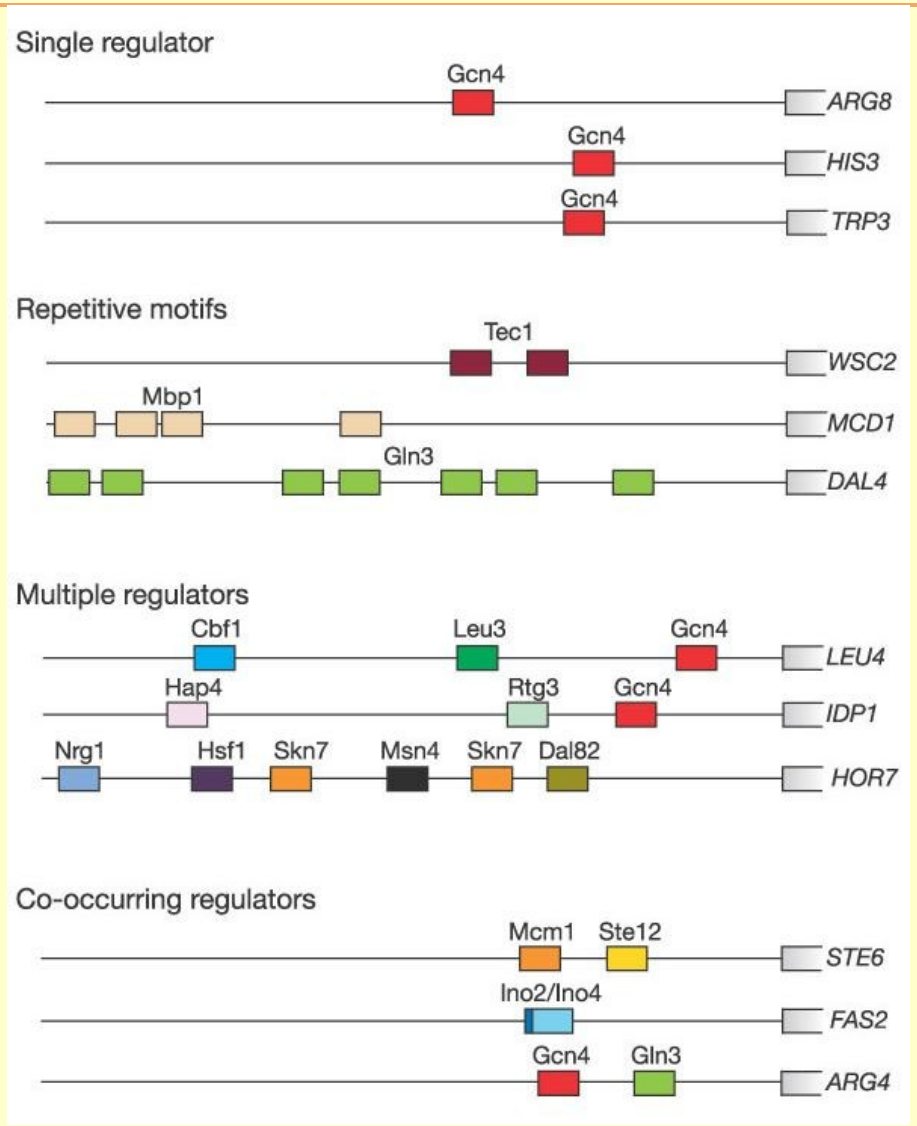
http://web.wi.mit.edu/young/regulatory_code/



Harbison et al, *Nature 431, 99-104 (2004)*

Yeast Promoter Architecture

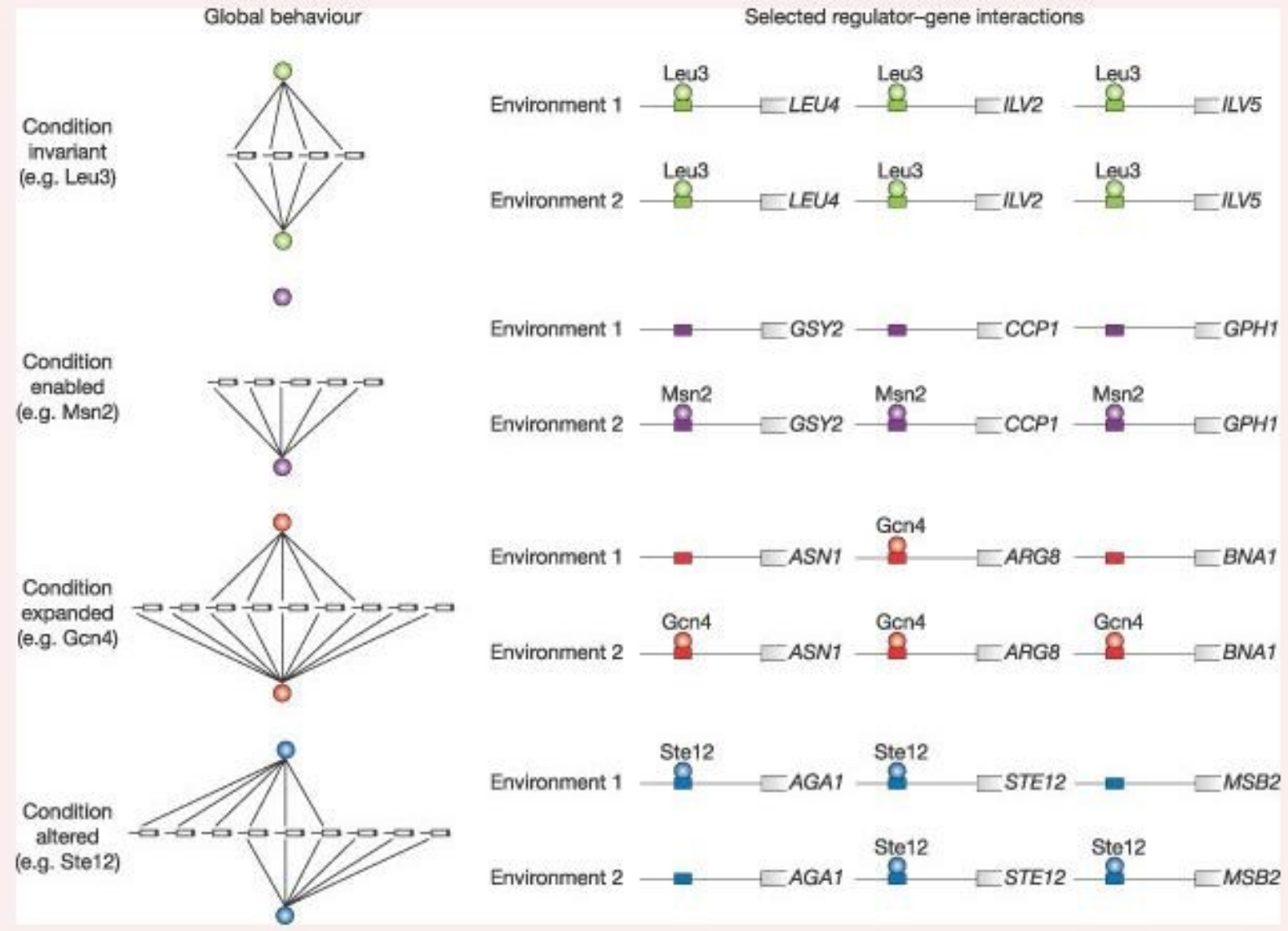
http://web.wi.mit.edu/young/regulatory_code/



Harbison et al, *Nature 431, 99-104 (2004)*

Environment Specific Regulators

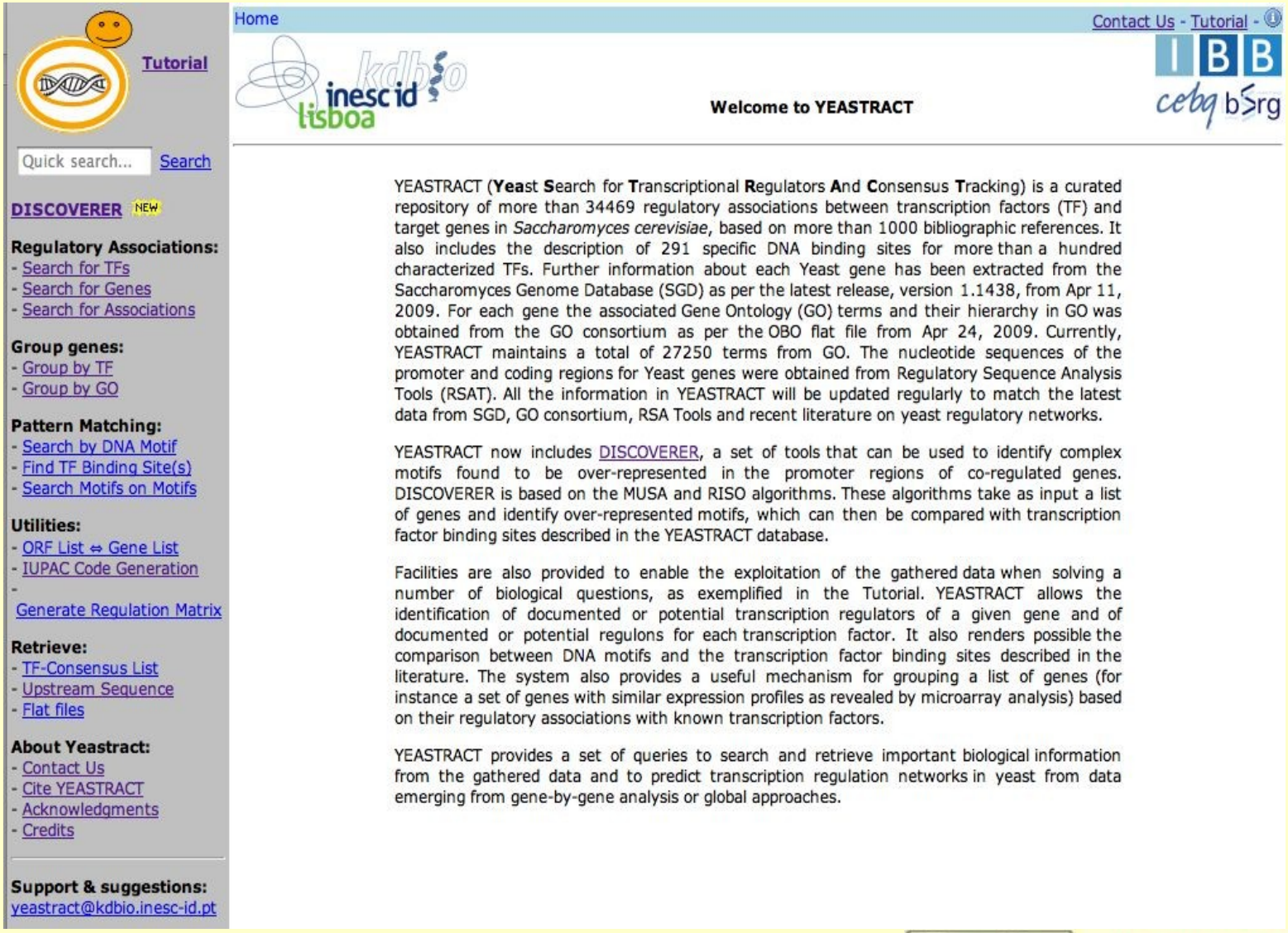
http://web.wi.mit.edu/young/regulatory_code/



Harbison et al, *Nature 431, 99-104 (2004)*

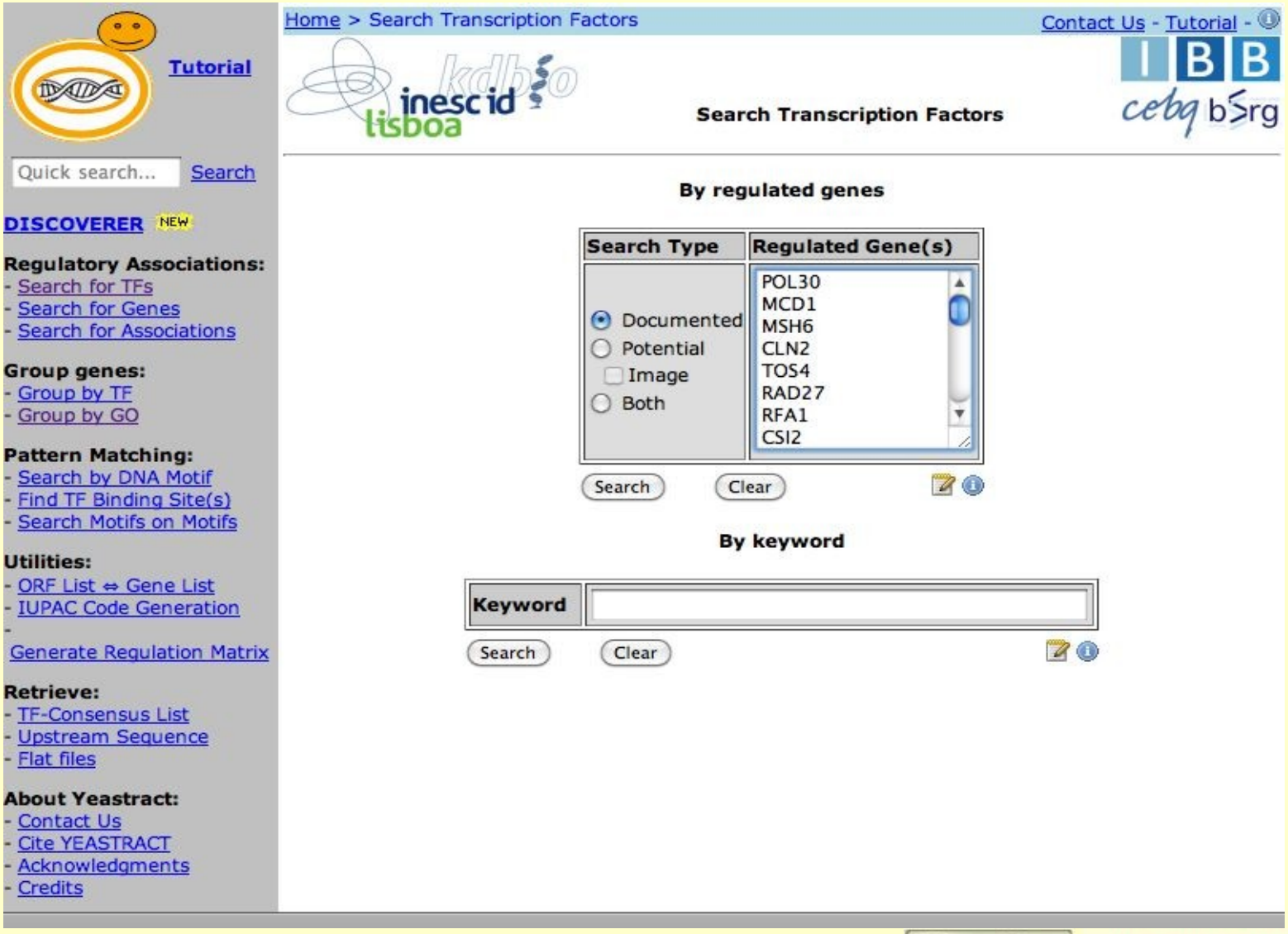
YeasTract Transcription Factor Database

http://www.yeastract.com/



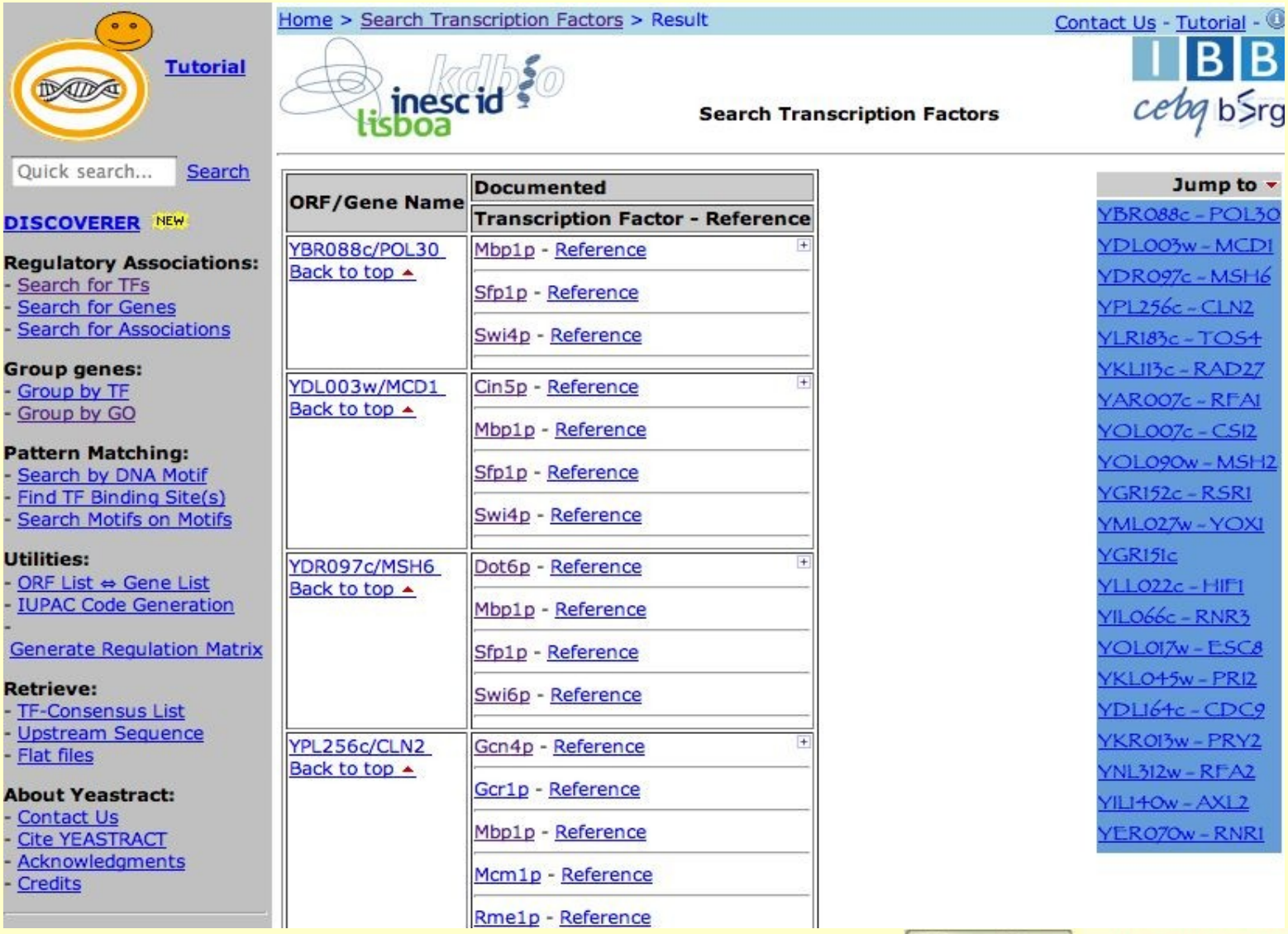
YeasTract Transcription Factor Search

http://www.yeastract.com/formfindregulators.php



YeasTract Transcription Factor Search Results

http://www.yeastract.com/formfindregulators.php



Tutorial

Quick search... Search

DISCOVERER NEW

Regulatory Associations:
- Search for TFs
- Search for Genes
- Search for Associations

Group genes:
- Group by TF
- Group by GO

Pattern Matching:
- Search by DNA Motif
- Find TF Binding Site(s)
- Search Motifs on Motifs

Utilities:
- ORF List ⇔ Gene List
- IUPAC Code Generation
- Generate Regulation Matrix

Retrieve:
- TF-Consensus List
- Upstream Sequence
- Flat files

About Yeastract:
- Contact Us
- Cite YEASTRACT
- Acknowledgments
- Credits

Home > Search Transcription Factors > Result

Contact Us - Tutorial -

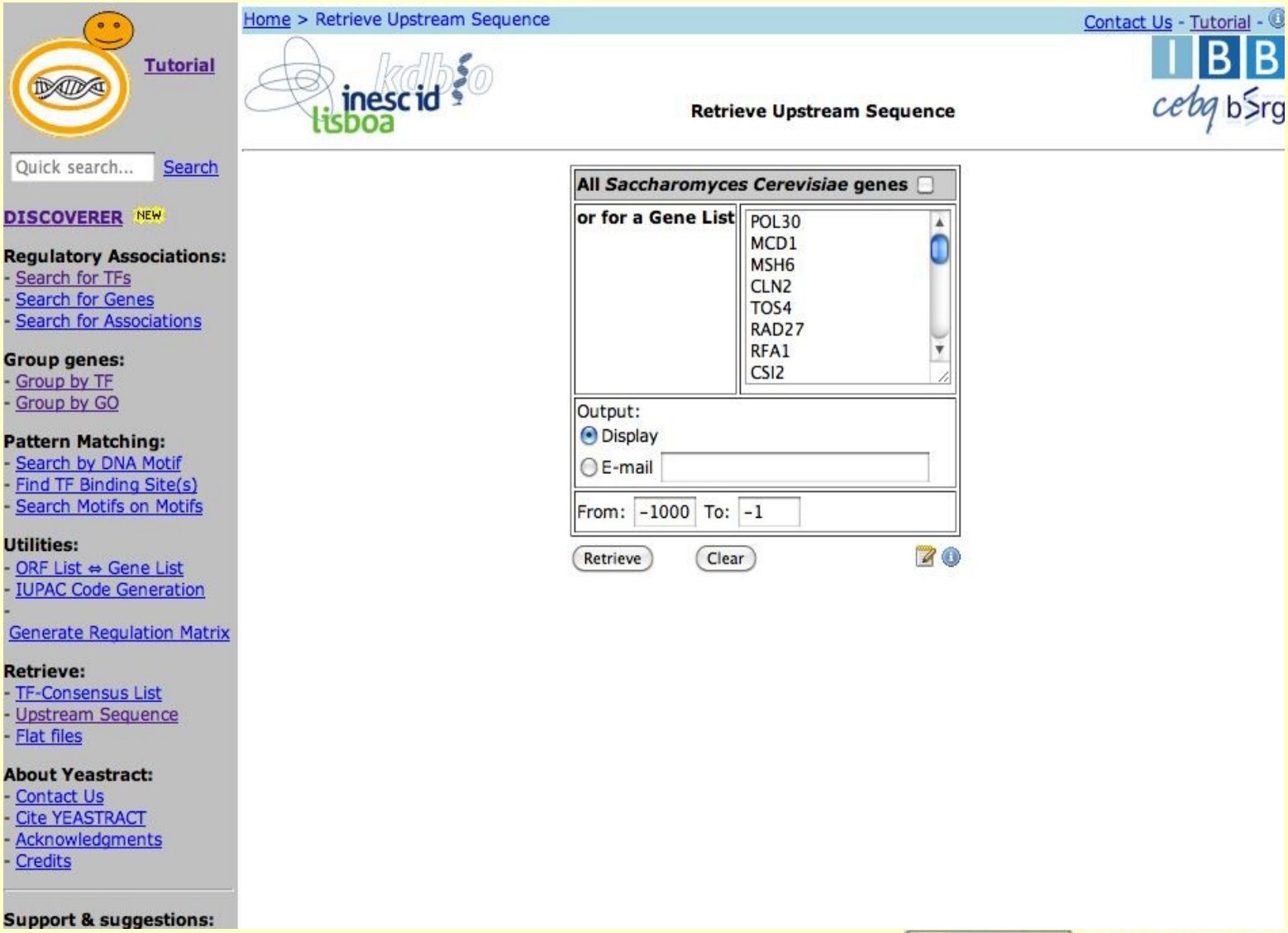
Search Transcription Factors

ORF/Gene Name	Documented Transcription Factor - Reference
YBR088c/POL30 Back to top	Mbp1p - Reference
	Sfp1p - Reference
	Swi4p - Reference
YDL003w/MCD1 Back to top	Cin5p - Reference
	Mbp1p - Reference
	Sfp1p - Reference
	Swi4p - Reference
YDR097c/MSH6 Back to top	Dot6p - Reference
	Mbp1p - Reference
	Sfp1p - Reference
	Swi6p - Reference
YPL256c/CLN2 Back to top	Gcn4p - Reference
	Gcr1p - Reference
	Mbp1p - Reference
	Mcm1p - Reference
	Rme1p - Reference

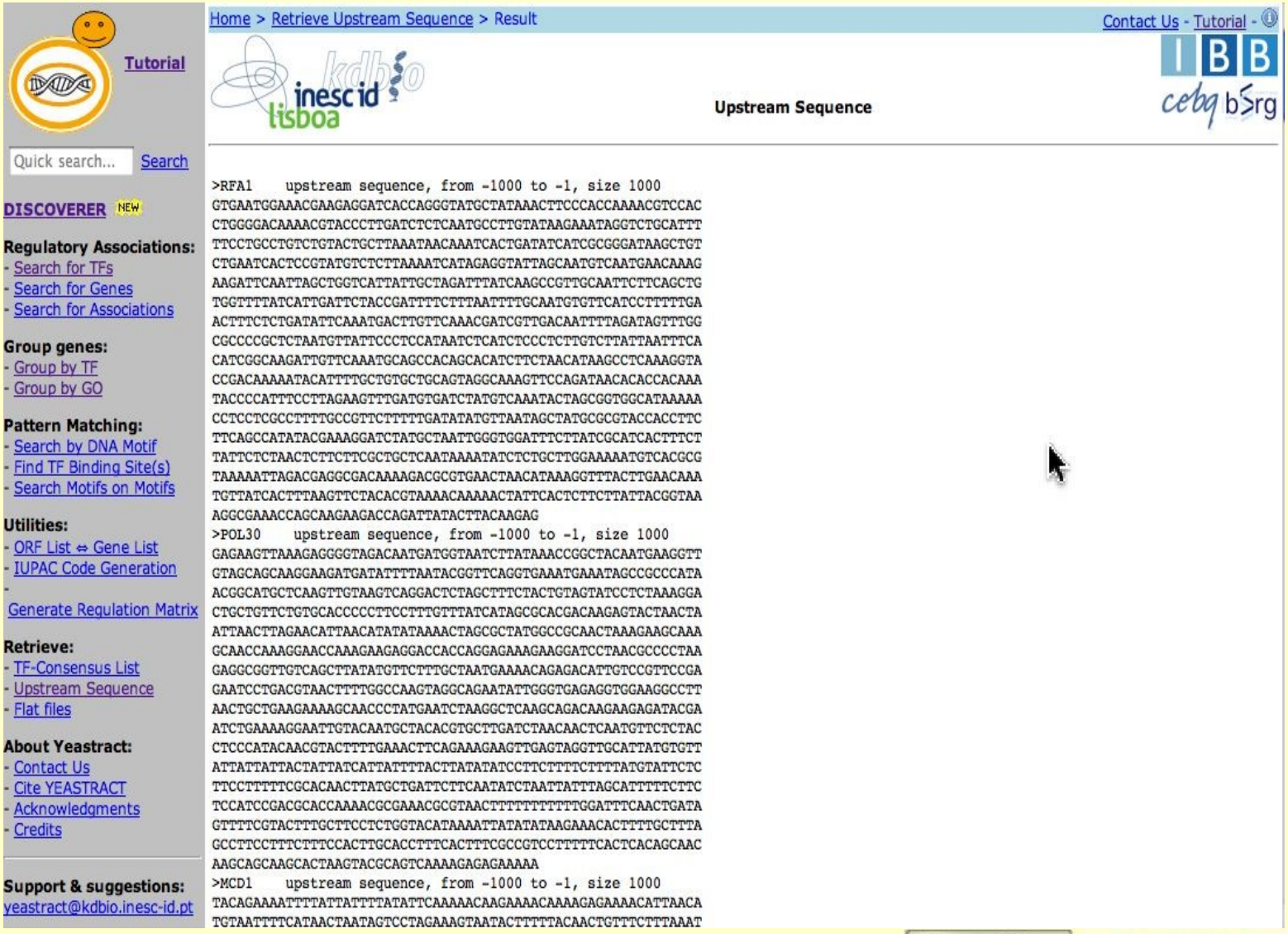
Jump to

YBR088c - POL30
YDL003w - MCD1
YDR097c - MSH6
YPL256c - CLN2
YLR183c - TOS4
YKL113c - RAD27
YAR007c - RFA1
YOL007c - CSI2
YOL090w - MSH2
YGR152c - RSR1
YML027w - YOX1
YGR151c
YLL022c - HIF1
YIL066c - RNR3
YOL017w - ESC8
YKL045w - PRI2
YDL164c - CDC9
YKR013w - PRY2
YNL312w - RFA2
YIL140w - AXL2
YER070w - RNR1

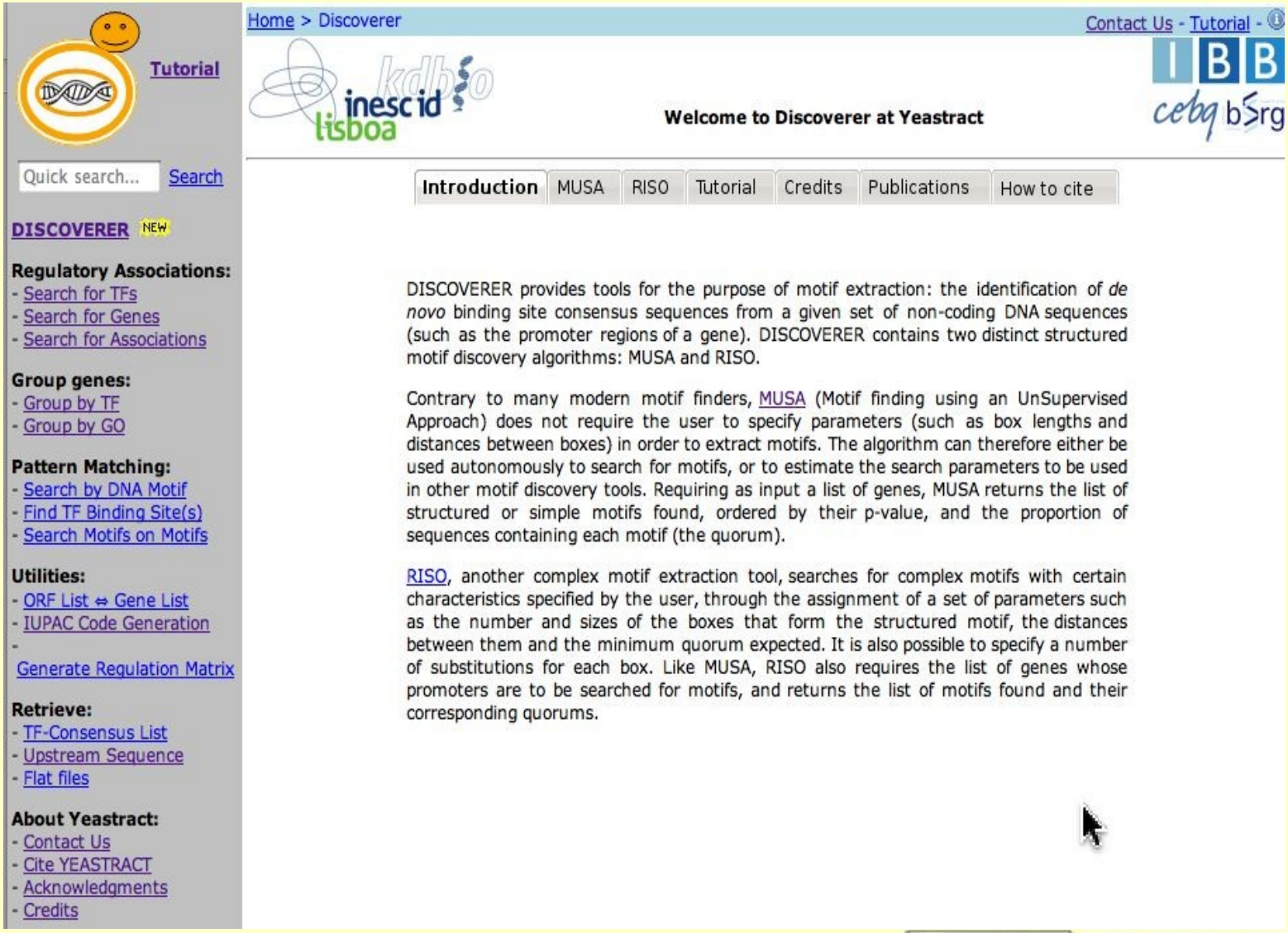
YeasTract UpStream Sequences
http://www.yeastract.com/formseqretrieval.php



YeasTract UpStream Sequences Retrieval
http://www.yeastract.com/formseqretrieval.php

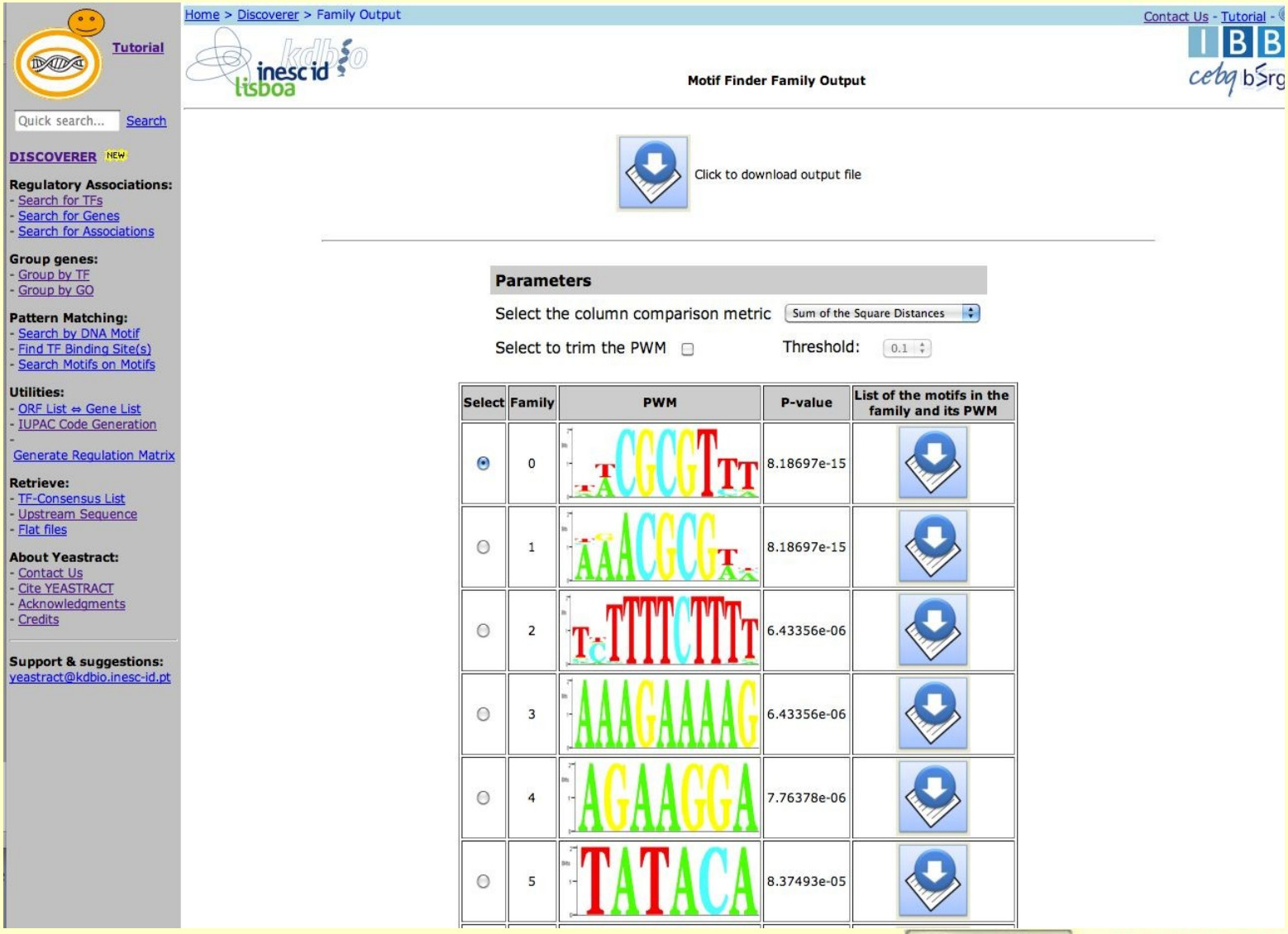


YeasTract Motif Discoverer
http://www.yeastract.com/discoverer/



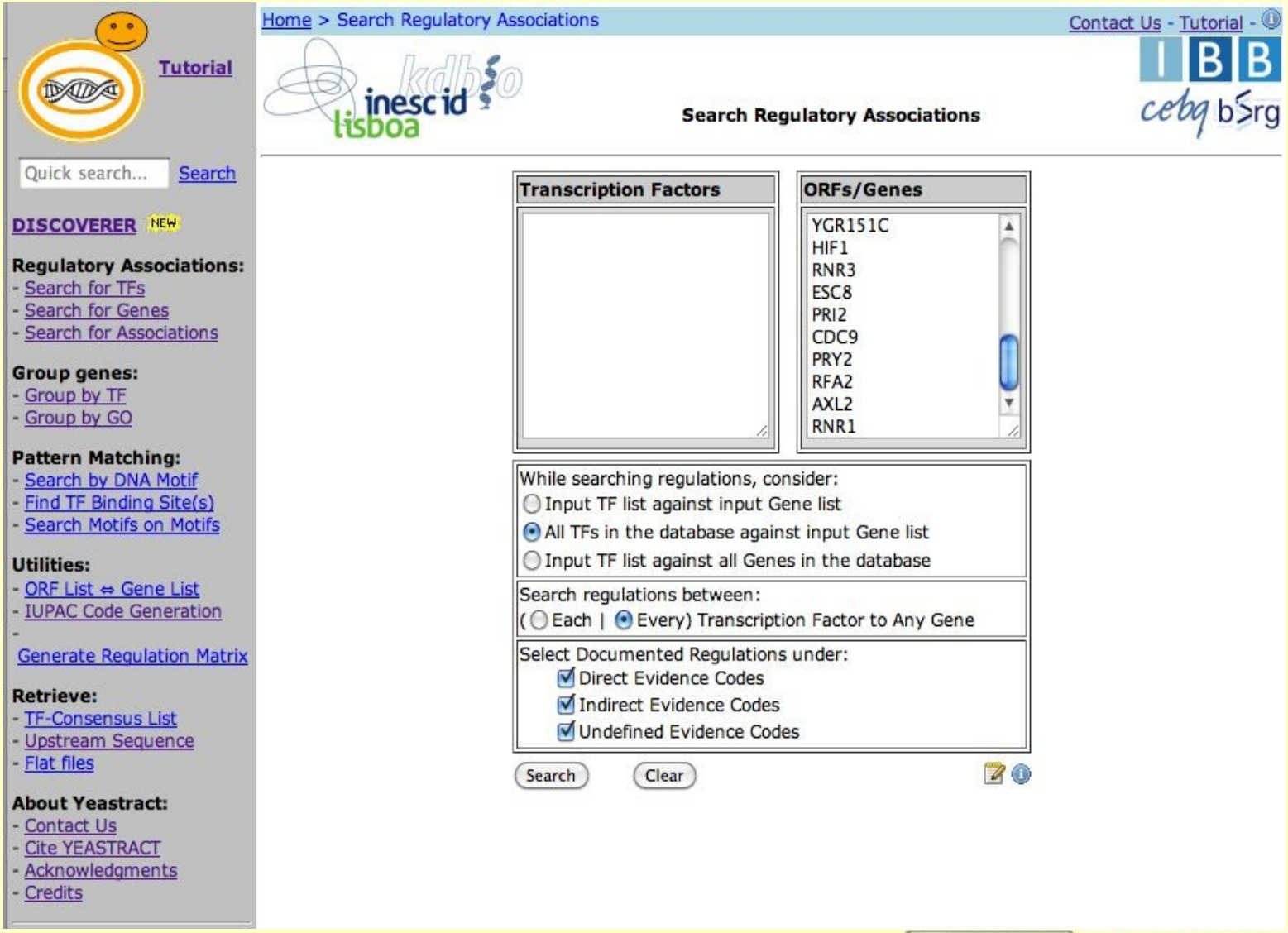
YeasTract Motif Discoverer Results

http://www.yeastract.com/discoverer/



YeasTract Transcription Factor Associations

http://www.yeastract.com/formregassociations.php



YeasTract Transcription Factor Associations

http://www.yeastract.com/formregassociations.php

Regulatory Associations

Quick search... Search

DISCOVERER NEW

Regulatory Associations:
- Search for TFs
- Search for Genes
- Search for Associations

Group genes:
- Group by TF
- Group by GO

Pattern Matching:
- Search by DNA Motif
- Find TF Binding Site(s)
- Search Motifs on Motifs

Utilities:
- ORF List ⇔ Gene List
- IUPAC Code Generation
- Generate Regulation Matrix

Retrieve:
- TF-Consensus List
- Upstream Sequence
- Flat files

About Yeastract:
- Contact Us
- Cite YEASTRACT
- Acknowledgments
- Credits

Support & suggestions:
yeastract@kdbio.inesc-id.pt

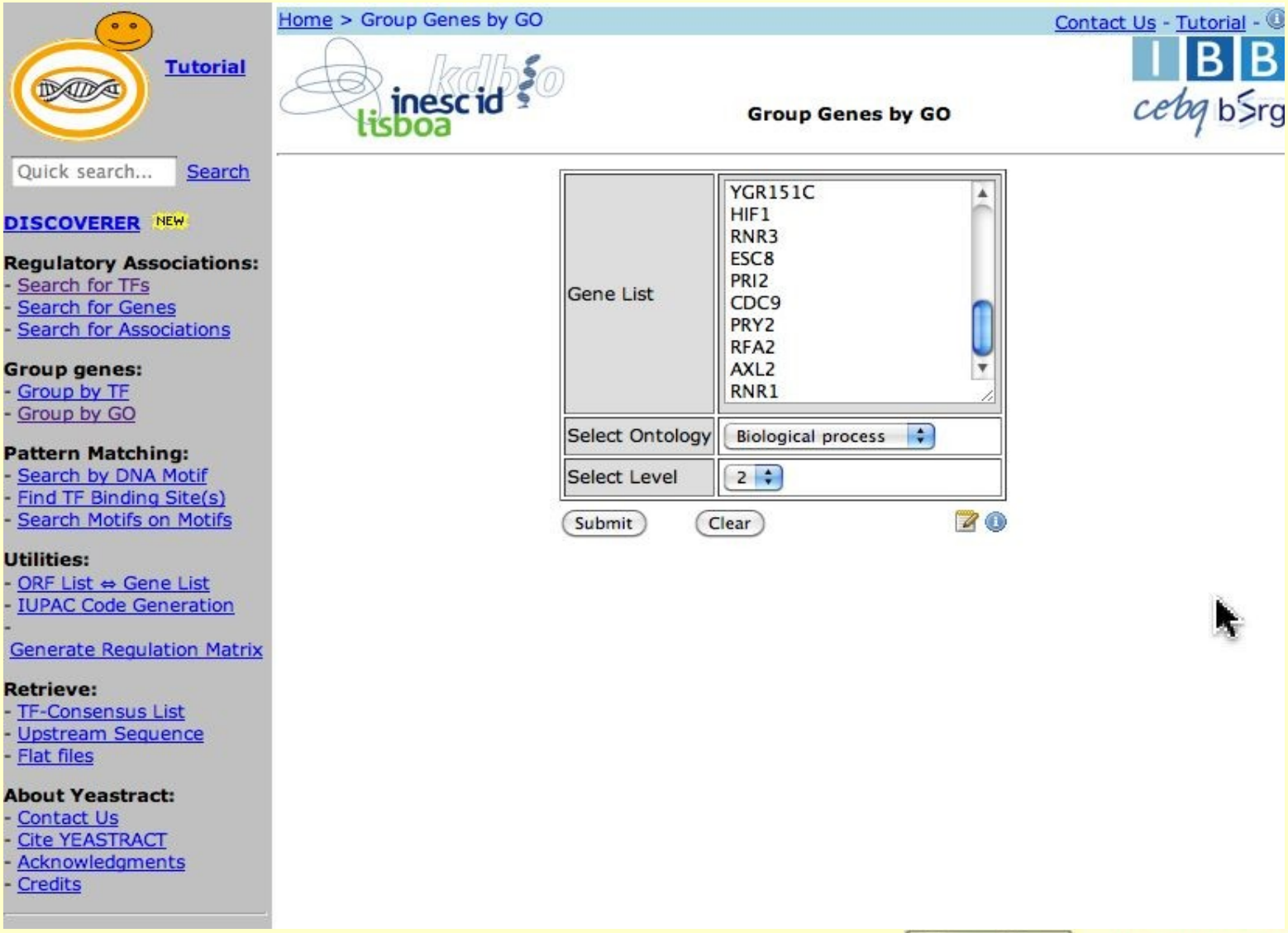
Transcription Factors	Documented Regulated Genes - Ref	Potential Regulated Genes - Promoter
Abf1p DNA binding protein with possible chromatin-reorganizing activity involved in transcriptional activation, gene silencing, and DNA replication and repair	RFA2 - Reference	HIF1 - Promoter POL30 - Promoter PRI2 - Promoter PRY2 - Promoter RAD27 - Promoter RFA1 - Promoter RFA2 - Promoter RNR1 - Promoter RNR3 - Promoter YOX1 - Promoter
Ace2p Transcription factor that activates expression of early G1-specific genes, localizes to daughter cell nuclei after cytokinesis and delays G1 progression in daughters, localization is regulated by phosphorylation; potential Cdc28p substrate	**No rows were found!**	AXL2 - Promoter CDC9 - Promoter RAD27 - Promoter RFA1 - Promoter RFA2 - Promoter RNR1 - Promoter YGR151c - Promoter
Adr1p Carbon source-responsive zinc-finger transcription factor, required for transcription of the glucose-repressed gene ADH2, of peroxisomal protein genes, and of genes required for ethanol, glycerol, and fatty acid utilization	RNR3 - Reference	CDC9 - Promoter CLN2 - Promoter CSI2 - Promoter ESC8 - Promoter MCD1 - Promoter MSH2 - Promoter MSH6 - Promoter PRI2 - Promoter RAD27 - Promoter RNR1 - Promoter TOS4 - Promoter YOX1 - Promoter
Aft1p Transcription factor involved in iron utilization and homeostasis; binds the consensus site PyPuCACCCPu and activates the expression of target genes in response to changes in iron availability	**No rows were found!**	POL30 - Promoter PRI2 - Promoter RAD27 - Promoter RNR1 - Promoter
Aft2p Iron-regulated transcriptional activator; activates genes involved in intracellular iron use and required for iron homeostasis and resistance to oxidative stress; similar to Aft1p	**No rows were found!**	POL30 - Promoter PRI2 - Promoter PRY2 - Promoter RAD27 - Promoter RNR1 - Promoter

Jump to

Abf1p
Aca1p
Ace2p
Adr1p
Aft1p
Aft2p
Arg80p
Arg81p
Aro80p
Arr1p
Ash1p
Azf1p
Bas1p
Bye1p
Cad1p
Cat8p
Cbf1p
Cdc14p
Cha4p
Cin5p
Crz1p
Cst6p
Cup2p
Cup9p
Dal80p
Dal81p
Dal82p
Dig1p
Dig2p

YeasTract Group Genes by Go Terms

http://www.yeastract.com/formregassociations.php



YeasTract Group Genes by Go Terms

GO Associations for Genes Using Biological Process at level 2

Go terms	Percentage	N. Genes	Genes/ORF
cellular process	90.0 %	18	Rsr1 Axl2 Yox1 Cdc9 Pri2 Rfa2 Tos4 Rad27 Cln2 Esc8 Hif1 Msh6 Pry2 Mcd1 Pol30 Msh2 Rnr1 Rnr3
metabolic process	65.0 %	13	Axl2 Yox1 Cdc9 Pri2 Rfa2 Rad27 Hif1 Pry2 Pol30 Msh6 Rnr1 Rnr3 Tos4
biological regulation	30.0 %	6	Cln2 Esc8 Yox1 Pol30 Rnr1 Rsr1
developmental process	20.0 %	4	Axl2 Mcd1 Msh6 Rad27
response to stimulus	20.0 %	4	Rnr1 Tos4 Cln2 Rnr3
reproductive process	10.0 %	2	Msh6 Rad27
localization	5.0 %	1	Rfa2
establishment of localization	5.0 %	1	Rfa2
multi-organism process	5.0 %	1	Cln2
reproduction	5.0 %	1	Axl2
anatomical structure formation	5.0 %	1	Hif1

Go Back

Quick search... Search

DISCOVERER NEW

Regulatory Associations:
- Search for TFs
- Search for Genes
- Search for Associations

Group genes:
- Group by TF
- Group by GO

Pattern Matching:
- Search by DNA Motif
- Find TF Binding Site(s)
- Search Motifs on Motifs

Utilities:
- ORF List ⇔ Gene List
- IUPAC Code Generation
- Generate Regulation Matrix

Retrieve:
- TF-Consensus List
- Upstream Sequence
- Flat files

About Yeastract:
- Contact Us
- Cite YEASTRACT
- Acknowledgments
- Credits

Swiss Regulon

SIB Swiss Institute of Bioinformatics

SIB Swiss Institute of Bioinformatics

BIOZENTRUM Universität Basel

SwissRegulon Portal

Swissregulon Quick links: Annotations Downloads FANTOM4 results

Services:

- Swissregulon
- Phylogibbs
- MARA
- TCS

Publications

Links

Contact

Swissregulon is a database of genome-wide annotations of regulatory sites. The predictions are based on Bayesian probabilistic analysis of a combination of input information including:

- Experimentally determined binding sites reported in the literature.
- Known sequence-specificities of transcription factors.
- ChIP-chip and ChIP-seq data.
- Alignments of orthologous non-coding regions.

Predictions were made using the PhyloGibbs, MotEvo, IRUS and MARA algorithms developed in our group, depending on the data available for each organism. Annotations can be viewed in a Gbrowse genome browser and can also be downloaded in flat file format.

Please cite this site as following:

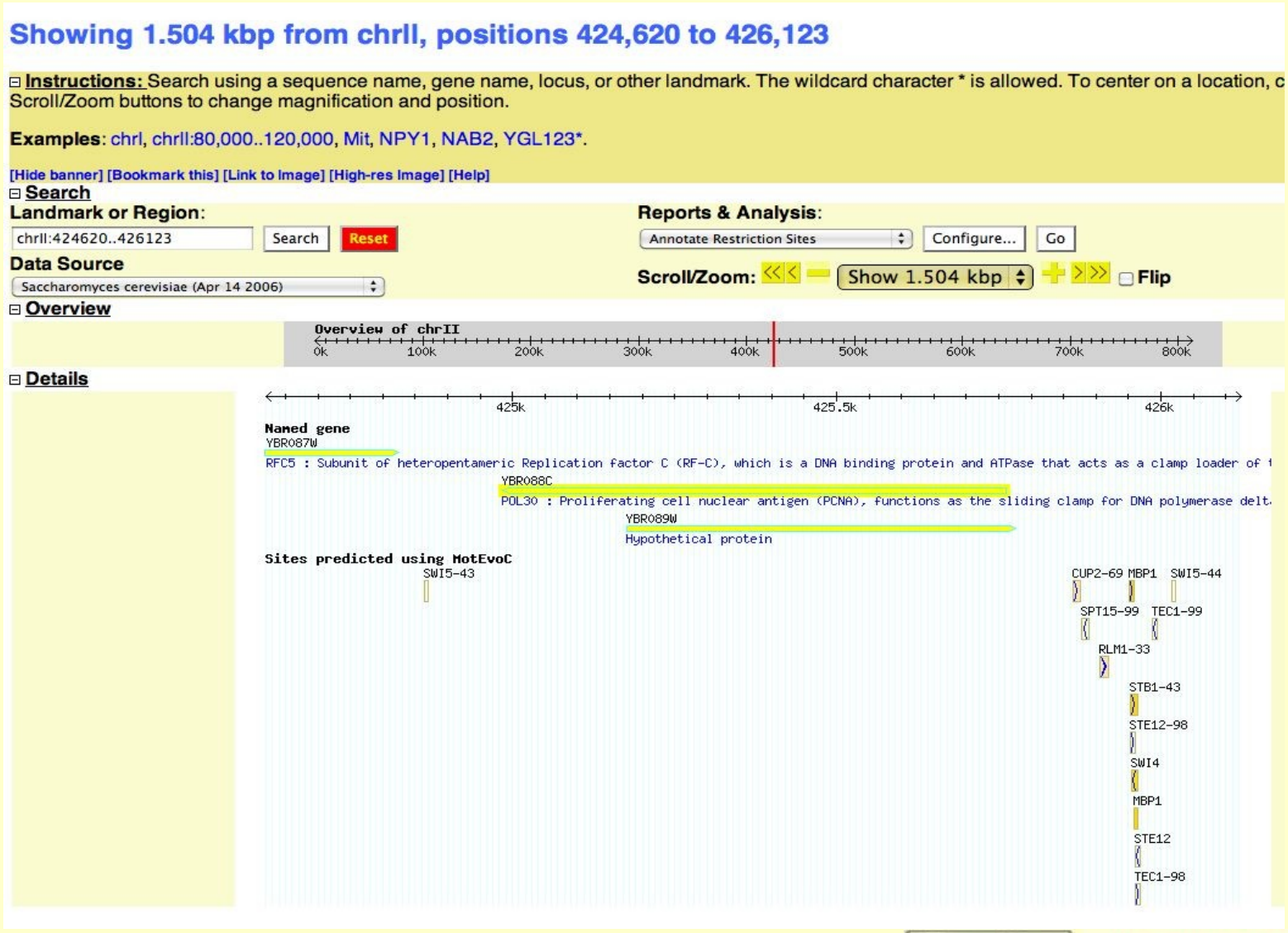
SwissRegulon: a database of genome-wide annotations of regulatory sites
Mikhail Pachkov, Ionas Erb, Nacho Molina and Erik van Nimwegen
Nucleic Acids Research, 2007, Vol. 35, Database issue D127-D131

Swiss Regulon Organisms

http://www.swissregulon.unibas.ch/cgi-bin/regulon

Swiss Regulon Yeast TF Binding Sites Near Pol30

http://www.swissregulon.unibas.ch/cgi-bin/gbrowse/yeast/



Swiss Regulon Yeast Transcription Factor Ste12

http://www.swissregulon.unibas.ch/cgi-bin/gbrowse/yeast/

Saccharomyces cerevisiae (Apr 14 2006)

back to main page

Name:	**STE12_DIG1-68**
Class:	TF_binding_site
Type:	TF_binding_site
Source:	MotEvoC
Position:	chrII:425967..425961 (- strand)
Length:	7
Score:	0.071272
Alias:	STE12_DIG1
Note:	CGTTTTG
factor_name:	STE12_DIG1
weight_matrix:	0.00 32.62 0.96 250.22 42.97 0.00 241.37 1.98 271.22 0.00 12.78 0.00 252.90 1.67 27.03 0.00 278.87 0.00 0.00 0.00 0.98 277.79 0.00 0.97 153.01 30.02 93.87 22.75

```
>STE12_DIG1-68 class=TF_binding_site position=chrII:425967..425961 (- strand)
CCAAAAC
```

Corresponding Weight Matrix


```
//
NA STE12_DIG1
PO       A         C         G         T
01       0.00      32.62     0.96      250.22
02       42.97     0.00      241.37    1.98
03       271.22    0.00      12.78     0.00
04       252.90    1.67      27.03     0.00
05       278.87    0.00      0.00      0.00
06       0.98      277.79    0.00      0.97
07       153.01    30.02     93.87     22.75
```

Swiss Regulon Yeast Transcription Factor Ste12 Sites

http://www.swissregulon.unibas.ch/cgi-bin/gbrowse/yeast/

Swiss Institute of Bioinformatics

Other locations of STE12_DIG1 sites:

Upstream gene	Upstream region coordinates	Site coordinates	Probability	Stra
---	---	chrVII:878320..878314	0.984212	-1
---	---	chrII:408661..408655	0.972057	-1
---	---	chrIII:229021..229027	0.972057	1
---	---	chrXIV:110765..110759	0.970263	-1
---	---	chrX:585873..585867	0.969271	-1
---	---	chrIII:13116..13110	0.966688	-1
---	---	chrII:192199..192193	0.966341	-1
---	---	chrXII:561858..561864	0.966132	1
---	---	chrXI:114864..114870	0.961639	1
---	---	chrVI:82256..82262	0.958973	1
---	---	chrXII:1041610..1041604	0.953367	-1
---	---	chrX:445331..445337	0.948736	1
---	---	chrVII:437032..437038	0.94671	1
---	---	chrXII:790167..790161	0.946707	-1
---	---	chrXIII:411475..411481	0.944509	1
---	---	chrV:188204..188210	0.944509	1
---	---	chrIII:13188..13194	0.944509	1
---	---	chrVIII:273742..273748	0.942583	1
---	---	chrIV:543992..543986	0.941509	-1
---	---	chrXIV:276112..276118	0.939875	1
---	---	chrXIV:703355..703349	0.935385	-1
---	---	chrXVI:462194..462188	0.933203	-1
---	---	chrVIII:273703..273709	0.932334	1
---	---	chrX:126581..126587	0.931886	1
---	---	chrVII:516467..516473	0.930058	1
---	---	chrV:167341..167335	0.930058	-1
---	---	chrIV:976108..976102	0.92753	-1
---	---	chrVII:437228..437234	0.922055	1
---	---	chrXIII:664436..664442	0.916896	1
---	---	chrXVI:781892..781898	0.916798	1
---	---	chrXIII:181232..181226	0.914531	-1
---	---	chrIV:50735..50729	0.914278	-1
---	---	chrXIII:659599..659605	0.908352	1

Datafiles for download:

- PhyloGibbs predictions made using orthologs only
- PhyloGibbs predictions based on CHIP-on-chip data
- MotEvo predictions based on CHIP-on-chip data
- Weight matrices