

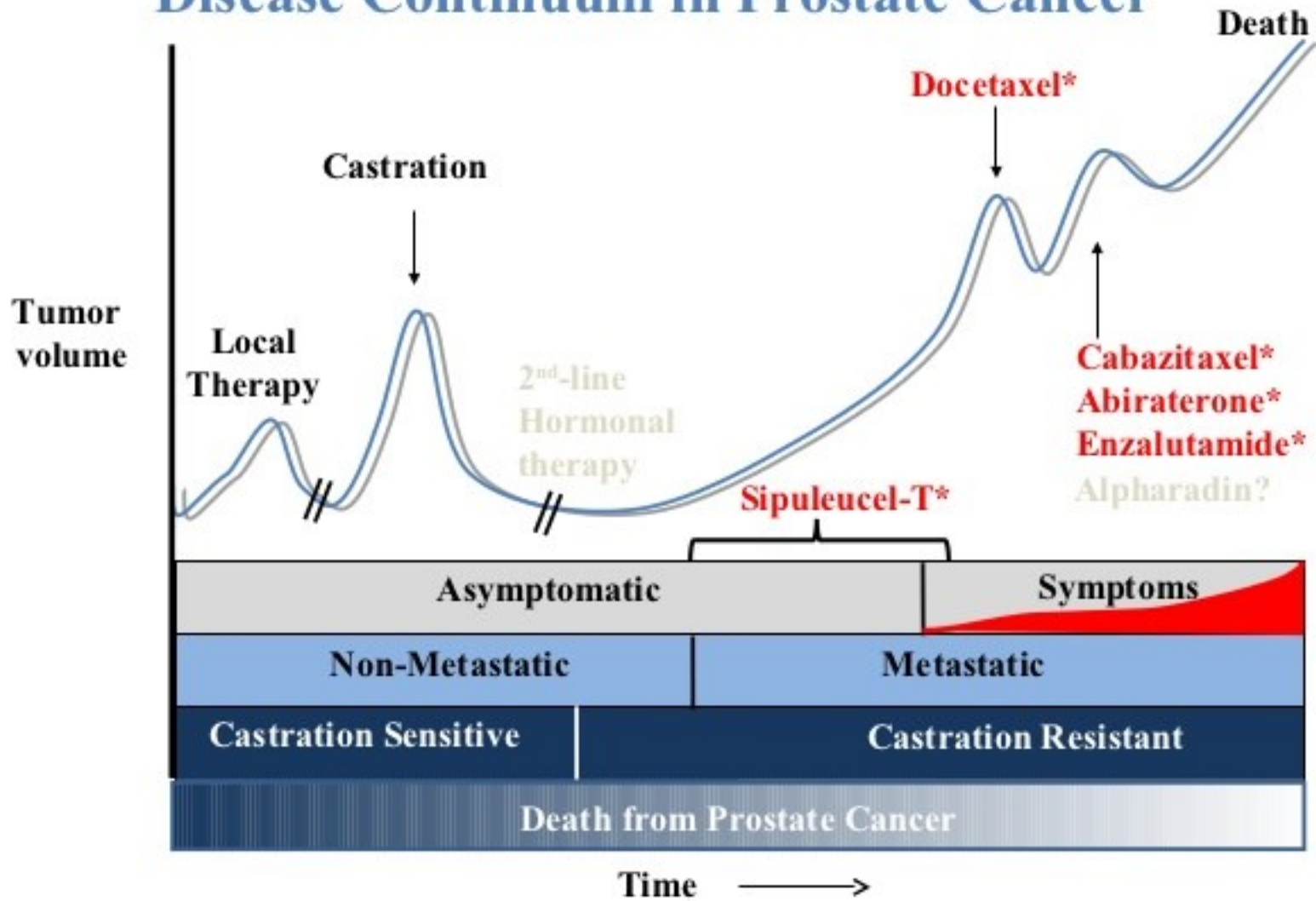
TRANSFAC



CANCER

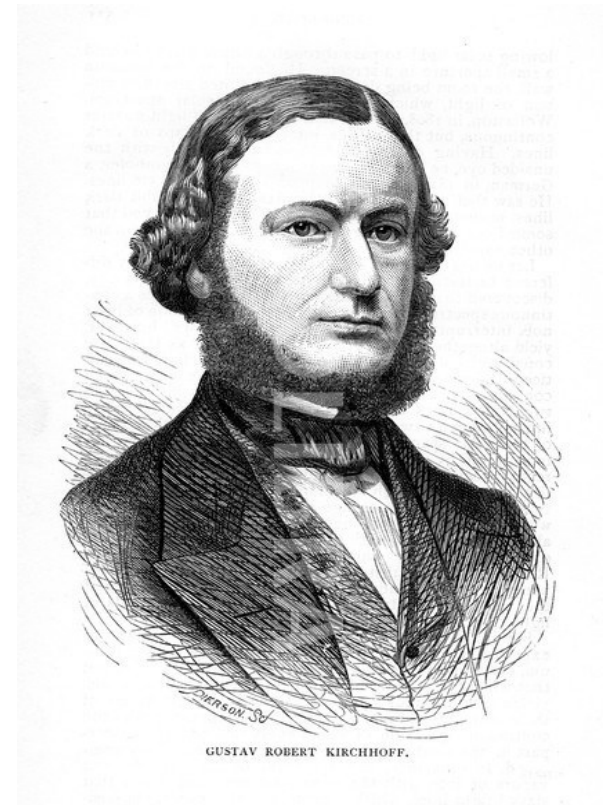
versus

Disease Continuum in Prostate Cancer



**"Eine gute Theorie ist das Praktischste
was es gibt."**

**"A good theory is the most
practical thing"**



Gustav Robert Kirchhoff

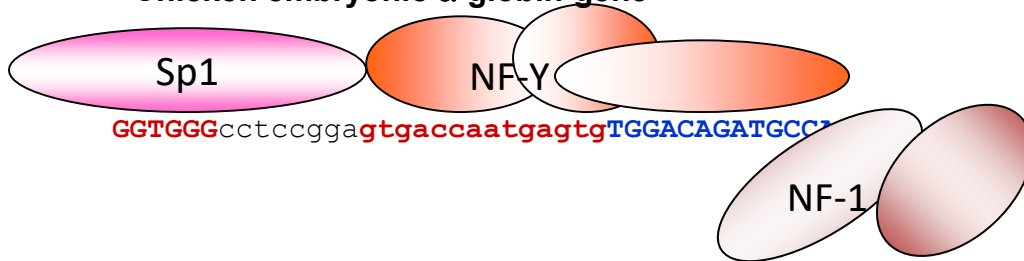
(1824 - 1887), German physicist



Antagonistic composite elements

COMPEL: C00006

Chicken embryonic α -globin gene

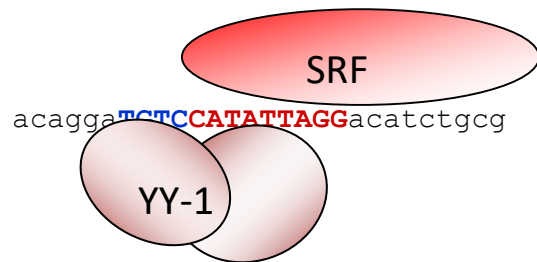


Sp1 cooperatively with **NF-Y** activates transcription in primitive erythroid cells

NF-1 represses transcription in adult cells

COMPEL: C00009

Human *c-fos* protooncogene

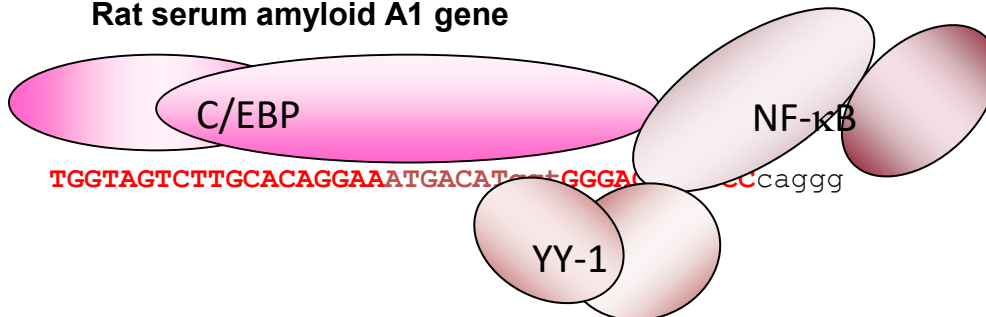


SRF mediates the rapid, transient induction of the *c-fos* protooncogen by serum growth factors.

YY1 diminishes both basal and serum-induced expression of the *c-fos*.

COMPEL: C00054

Rat serum amyloid A1 gene

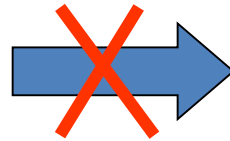


C/EBP and **NF- κ B** synergistically activate transcription in liver cells during acute phase response

YY1 represses inducible transcription of this gene.

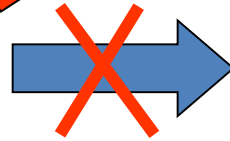
Paradigm shift

Nucleotide



Letter

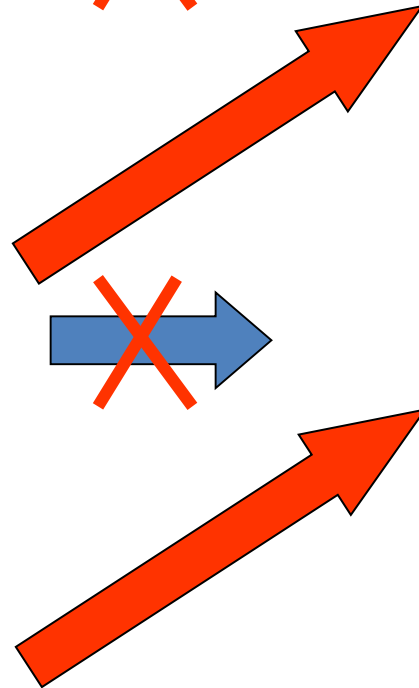
Site



Word

Site

Composition



Composite Modules (CM)

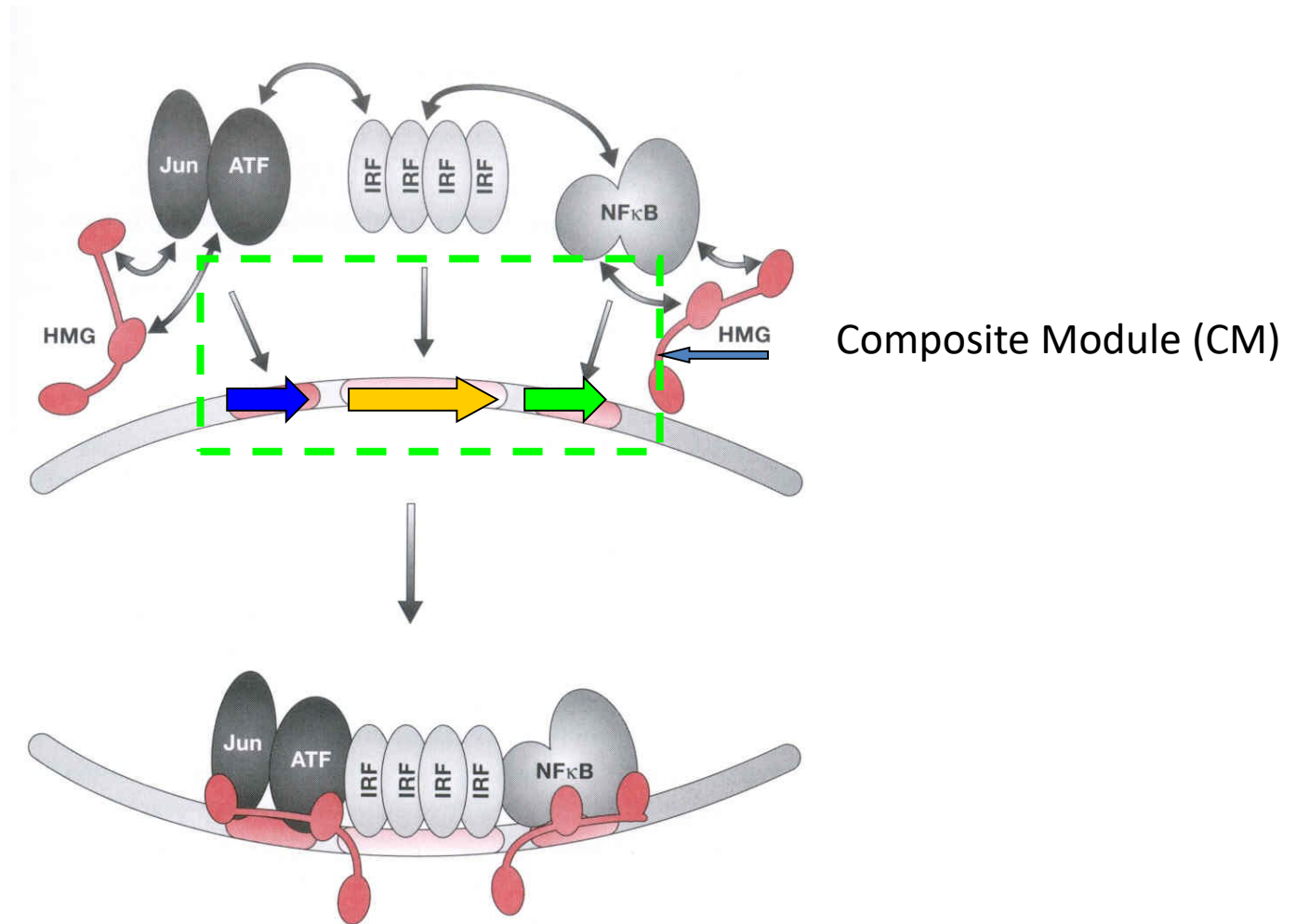
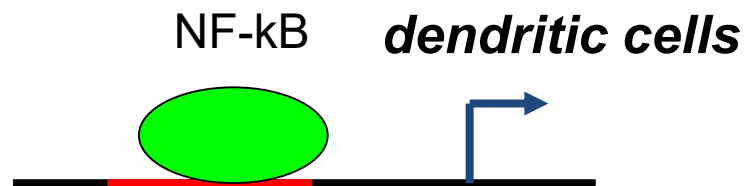
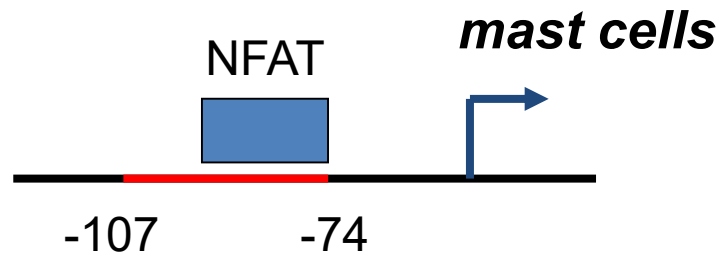
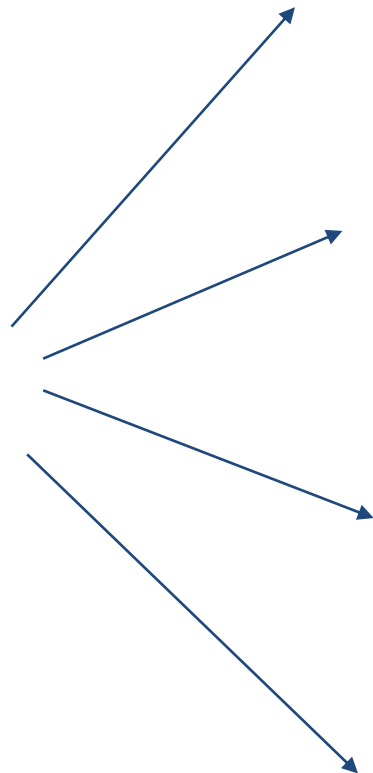


FIGURE 3.3. The human interferon- β enhanceosome. HMG represents HMGI/Y, a ubiquitous protein that binds cooperatively with the three activators. HMGI/Y both bends the DNA and contacts the activators. Each of the transcription factors shown is a member of a family of related activators. (Mark Ptashne, Alexander Gann *Genes and Signals*, 2002)

human TNF α promoter



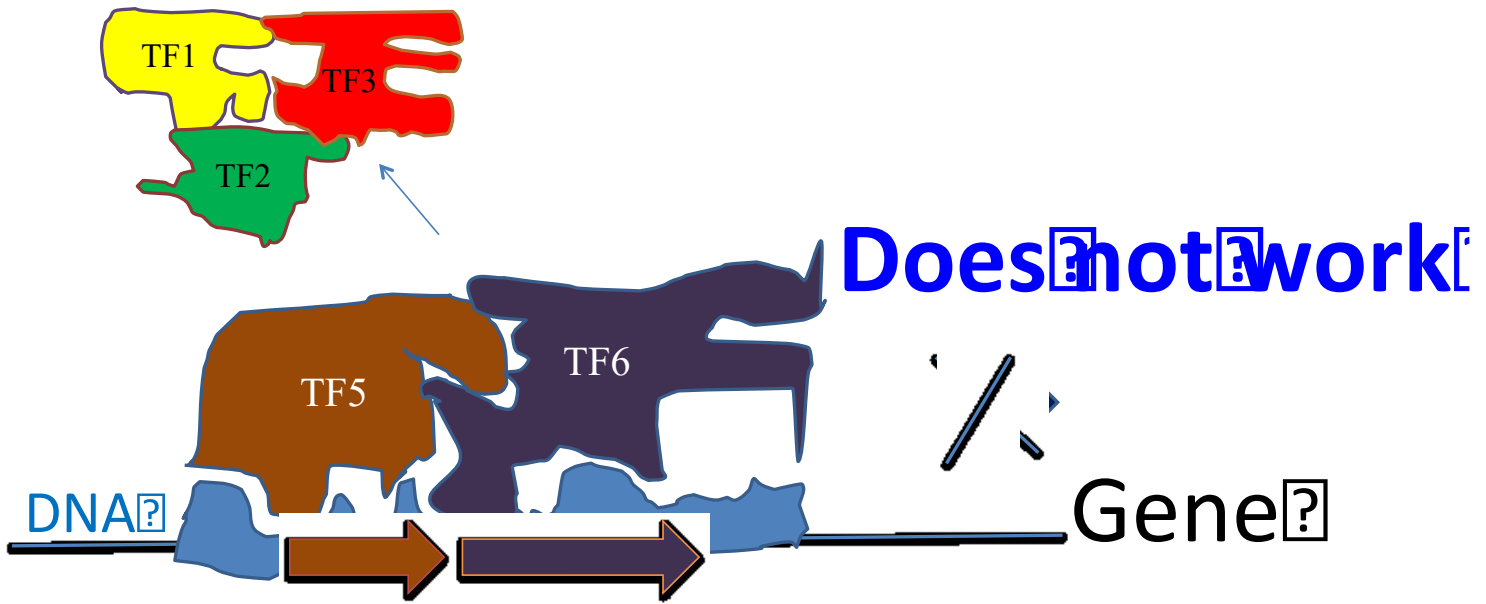
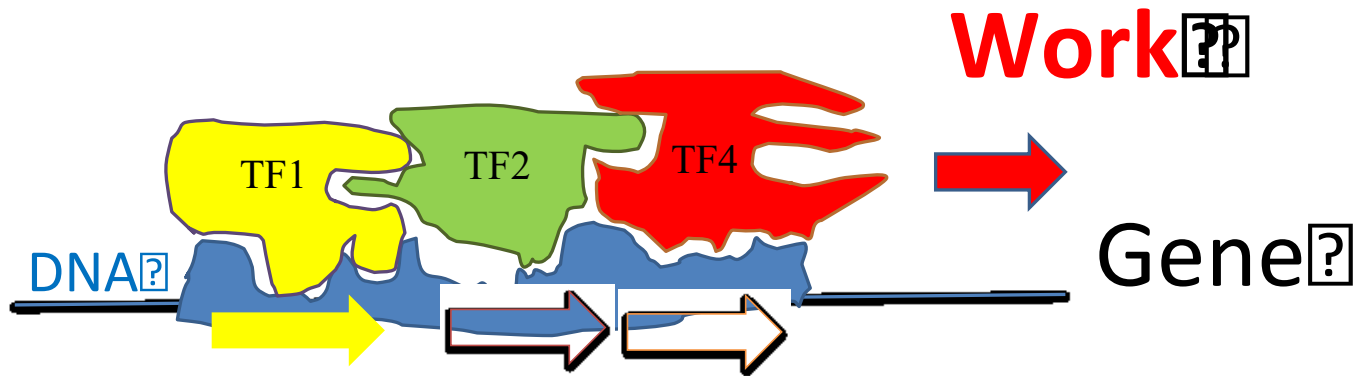
Exact fit

TF



DNA

It's Fuzzy Puzzle!



A Phase Separation Model for Transcriptional Control

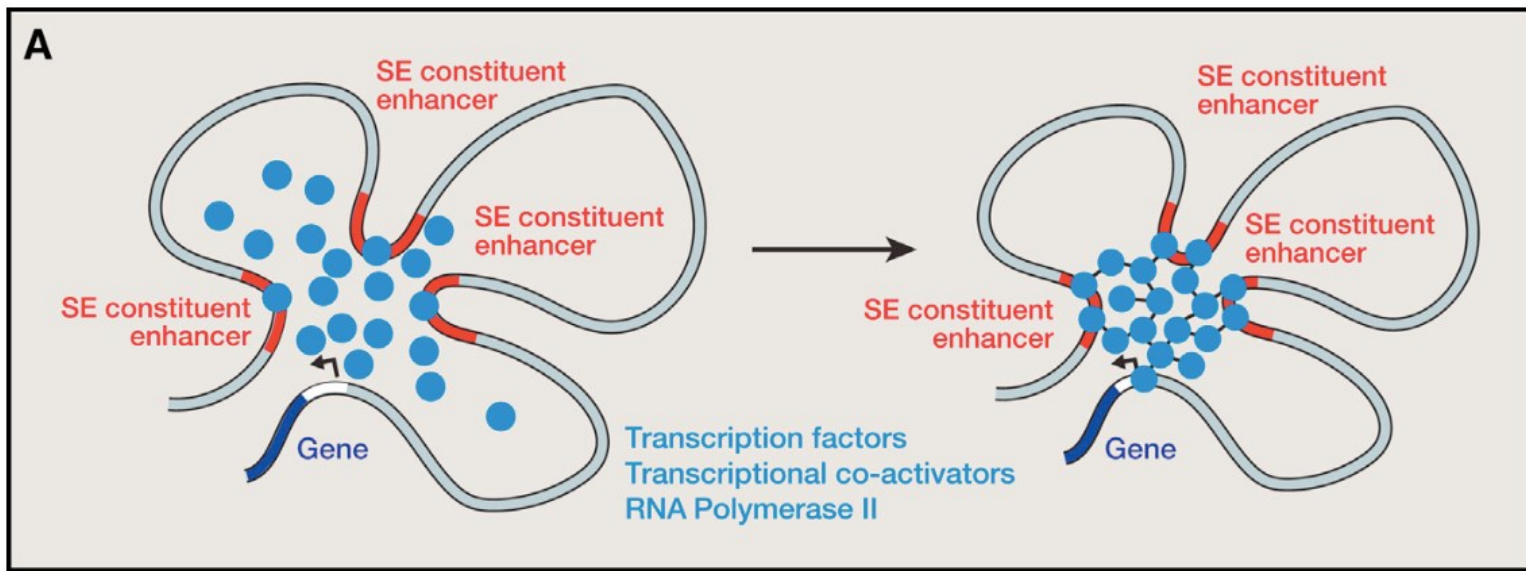
Denes Hnisz,^{1,10} Krishna Shrinivas,^{2,7,8,10} Richard A. Young,^{1,3,*} Arup K. Chakraborty,^{2,4,5,6,7,8,*} and Phillip A. Sharp^{3,9,*}

¹Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

²Department of Chemical Engineering

³Department of Biology

⁴Department of Physics



AP-1

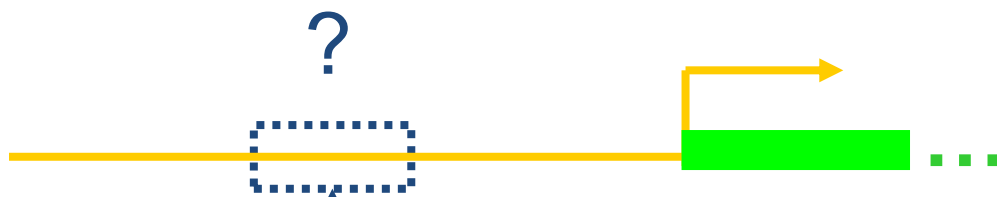
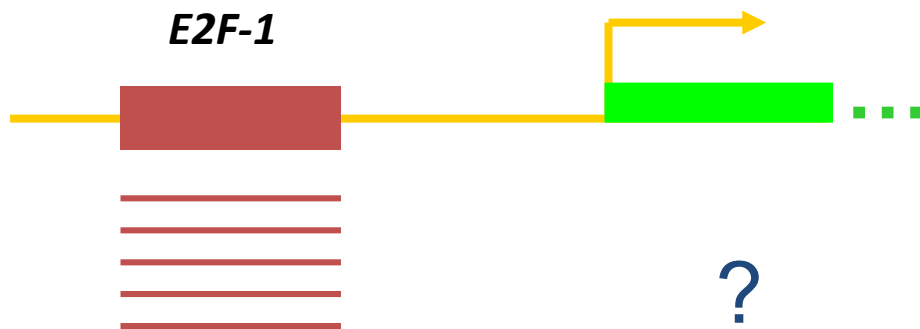
Consensus: TGA_gTCA

Human collagenase (-2013) * * * * *
TGA_GTCA

Mouse IL-2 (-143) * * * * *
TGTGTAA

Mouse IL-2 (-82) * * *
TGTAA_TA

Search for new TF binding sites with PWMs



Matrix length: 8



A	0.212	0.019	0.000	0.000	0.000	0.000	0.000	0.000
C	0.000	0.000	0.000	0.500	0.115	1.000	0.000	0.885
G	0.000	0.000	0.000	0.500	0.885	0.000	1.000	0.115
T	0.788	0.981	1.000	0.000	0.000	0.000	0.000	0.000

$$q = \frac{\sum_{i=1}^l I(i) f(b_i, i) - \sum_{i=1}^l I(i) f^{\min}(i)}{\sum_{i=1}^l I(i) f^{\max}(i)}$$

$$I(i) = \sum_{b \in \{A, T, G, C\}} f(b, i) \ln(4 f(b, i))$$

TRANSFAC®

Mouse c-fos promoter (Matrix search for TF binding sites)

```

1          <-----V$IK1_01(0.86)  -----...V$CREBP1CJUN_01(0.85)
2          <-----V$IK2_01(0.90)  -----...V$CREB_01(0.96)
3          ----->V$AP2_Q6(0.87)  <-----V$GKLF_01(0.87)
4-->V$ATF_01(0.89)  <-----V$MZF1_01(0.99)  -----...V$ELK1_01(0.87)
5          <-----V$AP2_Q6(0.92)  <-----V$SP1_Q6(0.88)
6>V$AP1FJ_Q2(0.89)  <-----V$GKLF_01(0.85)
7>V$AP1_Q2(0.87)  <-----V$GKLF_01(0.86)
8->V$CREB_Q2(0.86)  <-----V$CTS1P54_01(0.90)
9->V$CREB_Q4(0.90)  <-----V$NRF2_01(0.90)
10         <-----V$GC_01(0.88)
11         ----->V$CAAT_01(0.87)
12         <-----V$TCF1_01(0.87)
13         ----->V$AP2_Q6(0.87)
14         <-----V$USF_Q6(0.93)
16         -----...V$ATF_01(0.94)
17         -----...V$AP1FJ_Q2(0.95)
20         -----...V$CREBP1_Q2(0.93)
21         -----...V$CREB_Q2(0.95)
23         ---...V$IK2_01(0.85)
MMCFOF_1  GAGCGCCCGCAGAGGCCCTTGGGGCGCGCTTCCCCCCCCTTCCAGTTCGCCCCAGTGACG  420

1-->V$CREBP1CJUN_01(0.85)  ----->V$BARBIE_01(0.86)
2-->V$CREB_01(0.96)  ----->V$TATA_01(0.95)
3          ----->V$CAAT_01(0.91)  ----->V$AP4_Q5(0.95)
4----->V$ELK1_01(0.87)  ----->V$HEN1_01(0.87)
5          ----->V$AP4_Q5(0.88)  <---...V$CMYB_01(0.93)
6          <-----V$DPCR3HD_01(0.93)  ---...V$VMYB_02(0.89)
7          <-----V$TATA_01(0.88)
8          ----->V$HEN1_02(0.87)
9          <-----V$HEN1_02(0.86)
10         <-----V$AP4_Q1(0.88)
11         ----->V$LMO2COM_01(0.93)
12         <-----V$LMO2COM_01(0.93)
13         <-----V$MYOD_01(0.88)
17---->V$AP1FJ_Q2(0.95)  <-----V$AP4_Q6(0.99)
20---->V$CREBP1_Q2(0.93)  <-----V$MYOD_Q6(0.96)
21---->V$CREB_Q2(0.95)
23----->V$IK2_01(0.85)
24         <===== E2F (0.80)
MMCFOF_1  TAGGAAGTCCATCCATTACAGCGCTTCTATAAAGCGCCAGCTGAGGCGCTACTACTC  480

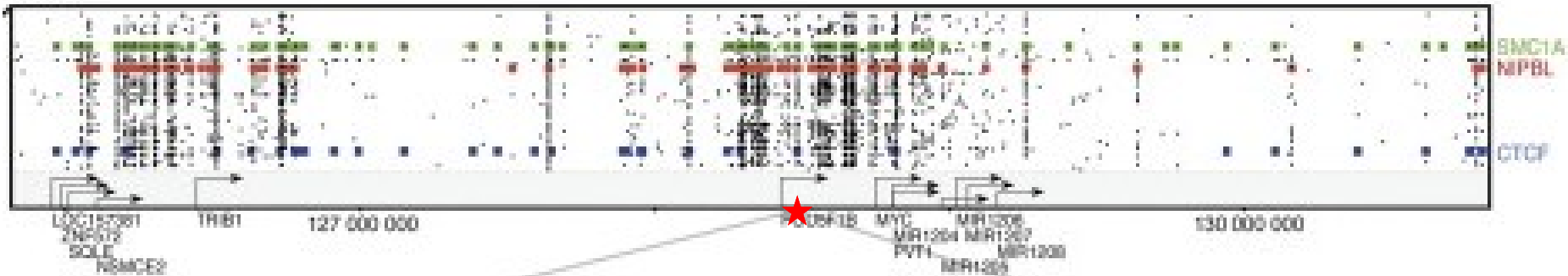
1          <-----V$CMYB_01(0.91)  -----...V$ER_Q6(0.86)
2          <-----V$LMO2COM_01(0.90)  <-----...V$TCF1_01(0.87)
3          ----->V$MYOD_Q6(0.90)  ----->V$STAT_01(0.93)
4          ----->V$VMYB_01(0.89)  <-----V$STAT_01(0.89)
5----->V$CMYB_01(0.93)  ----->V$LMO2COM_02(0.93)
6----->V$VMYB_02(0.89)  <-----V$CAAT_01(0.85)
7          ----->V$VMYB_02(0.88)
8          ----->V$EV1_04(0.86)
9          ----->V$GATA1_02(0.93)
12         <-----V$ZID_01(0.85)
13         <-----V$CP2_01(0.97)
14         ----->V$GATA_C(0.92)
15         ----->V$CMYB_01(0.86)
16         ----->V$CREL_01(0.91)
24         <===== E2F (0.82)
MMCFOF_1  CAACCGCGACTGCAGCGGCAACTGAGAAGACTGGATAGAGCCGCGGTTCCGCGAACGA  540

```

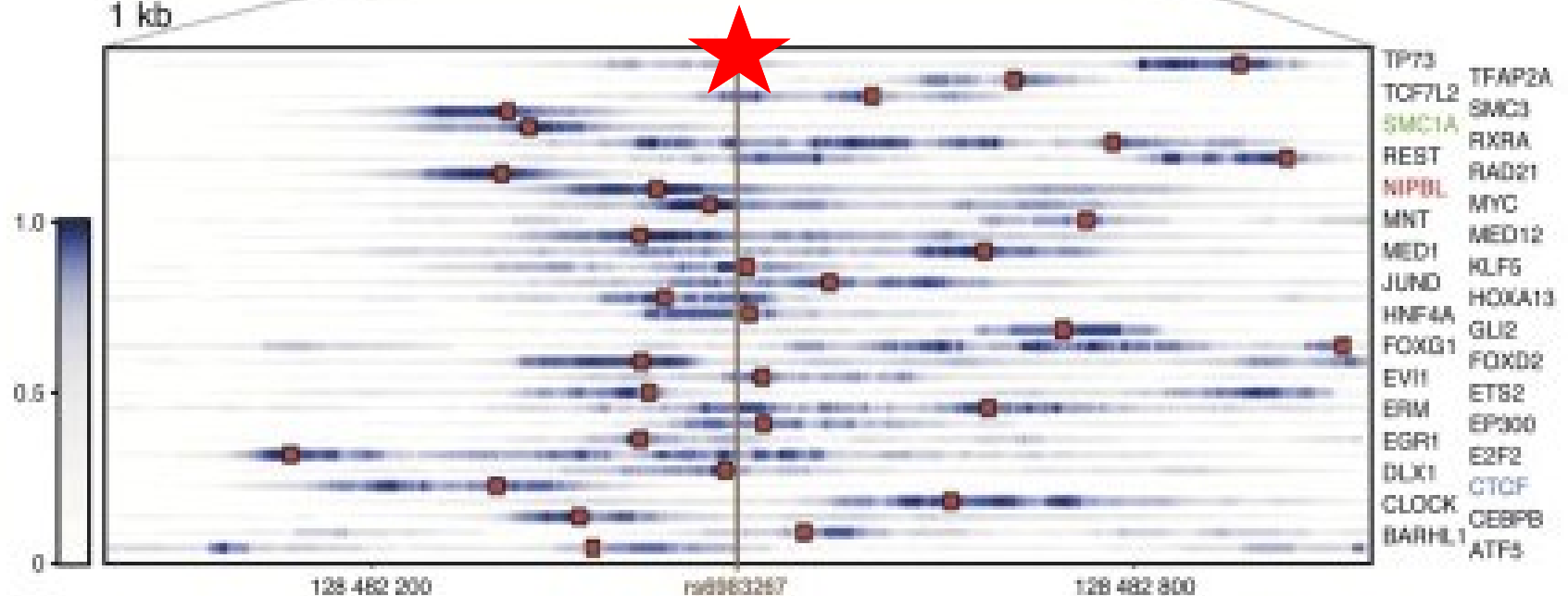
Transcription start

Colorectal cancer: tumor-specific enhancer around a SNP in regulatory region of MYC gene

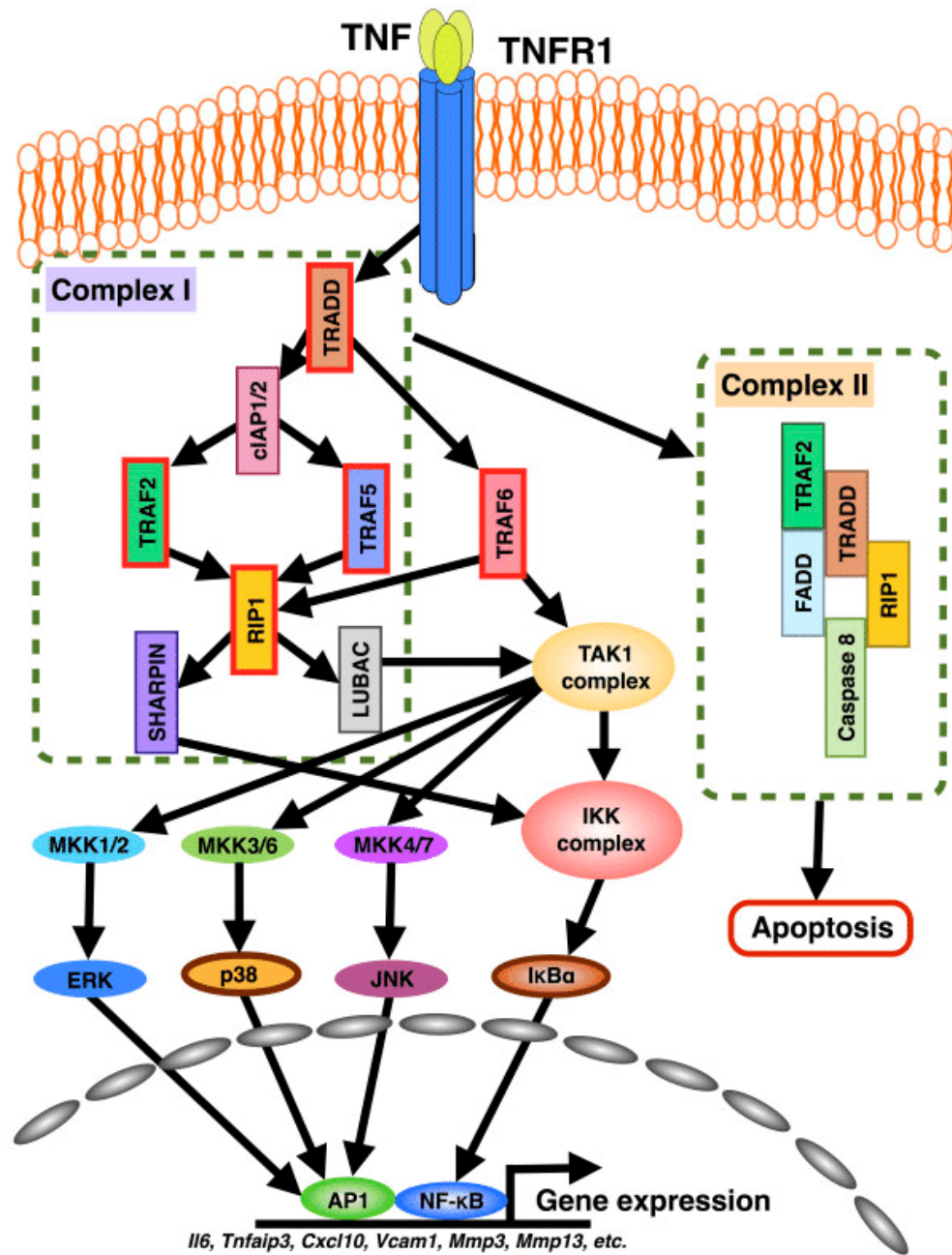
5 Mb



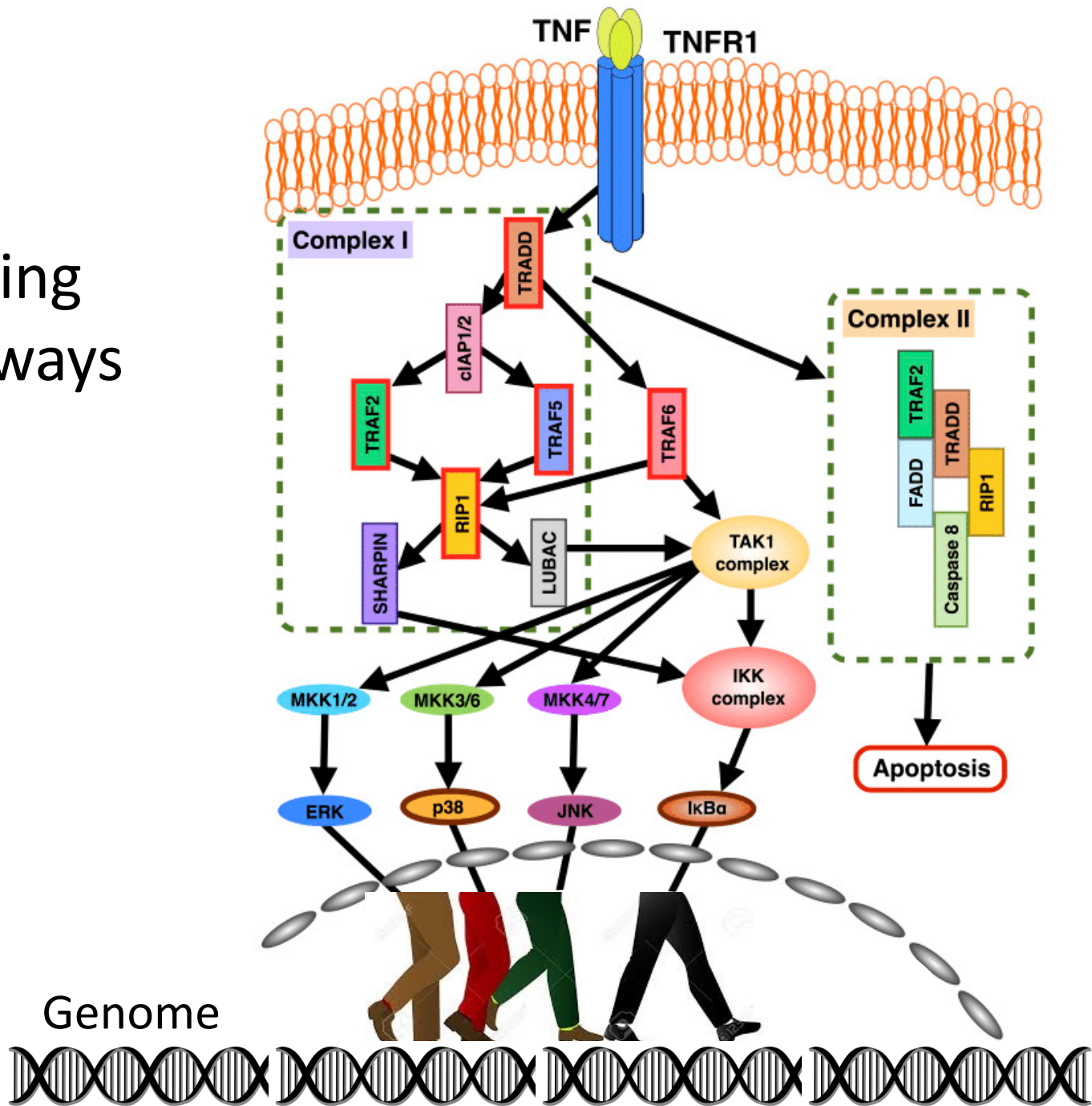
1 kb



Walking pathways



Walking pathways



Characterization of Regulatory Genomic Regions. Development of Databases and Sequence Analysis Tools

TRADAT

1995 - 1998

The concept of the TRADAT consortium included to provide an integrated platform for databases and software tools for the analysis of regulatory genomic regions. According to this concept, a set of databases was established and maintained, mainly EPD (Eukaryotic Promoter Database) and TRANSFAC (Transcription factors and their genomic binding sites). These data sources were successfully linked with each other and with a number of external databases. In addition to this integrated data resource, a series of software tools for the identification of individual regulatory elements and the characterization of their context were developed. These tools as well as the underlying patterns were subjected to systematic evaluation, optimization and experimental verification.

The concept of using weight matrices for the detection of transcription factor binding sites was extended and incorporated into higher order software tools. TRADAT contributed to the further development of MatInspector, ModelGenerator, FastM, and ModelInspector.

Евгений
Multi-omics





Research: SysCol

Default

Databases Data Analyses Users

Start page

- analyses
 - Galaxy
 - JavaScript
 - Methods
 - Data manipulation
 - Data normalization
 - FBC
 - Functional classification
 - Import
 - Molecular networks
 - Add expression values
 - Add reactants
 - Apply state to diagram
 - Cluster by path
 - Cluster by shortest path
 - Effector search
 - Extend network
 - Find longest connected chains
 - Find shortest path between two s
 - Join diagrams
 - Match genes and metabolites
 - Regulator search
 - Save hits
 - Save network
 - Visualize results
 - NGS
 - Optimization
 - Simulation
 - Site analysis
 - Analyze miRNA target enrichmen
 - Apply CMA model to tracks
 - Change profile cutoffs
 - Cluster track
 - Compare TFBS mutations
 - Compute profile thresholds
 - Construct IPS CisModule
 - Construct composite modules
 - Construct composite modules on

Upstream analysis

RNA-seq

Proteomics

Epigenomics

ChIP-seq

Sequence
analysis

miRNA

Microarrays

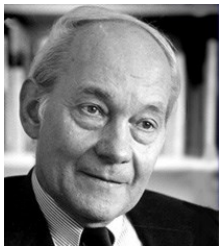
Drug targets

Pathways

NGS

Genomic
variantsPopular
functionsGene or
protein listComplete list of
workflows

Metabolism



Manfred Eigen



Vadim Ratner

(?)

Nikolay Kolchanov

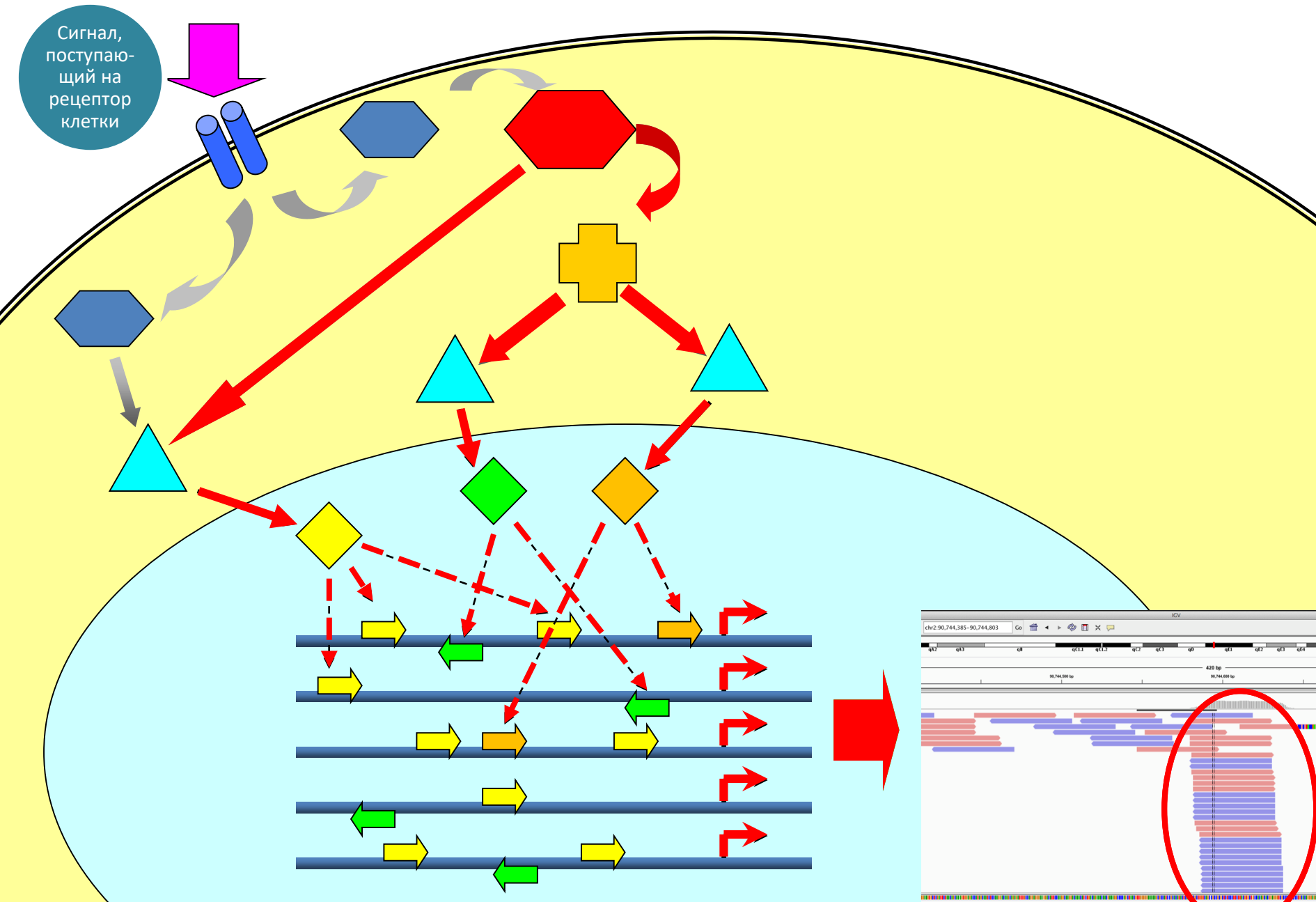


The Edgar

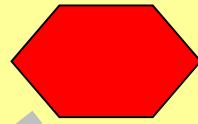


How to search for the causal mechanism of cancer?

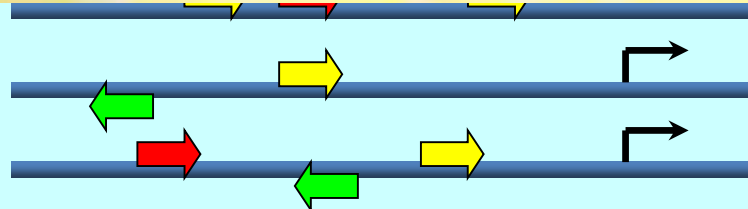
Сигнал, поступающий на рецептор клетки



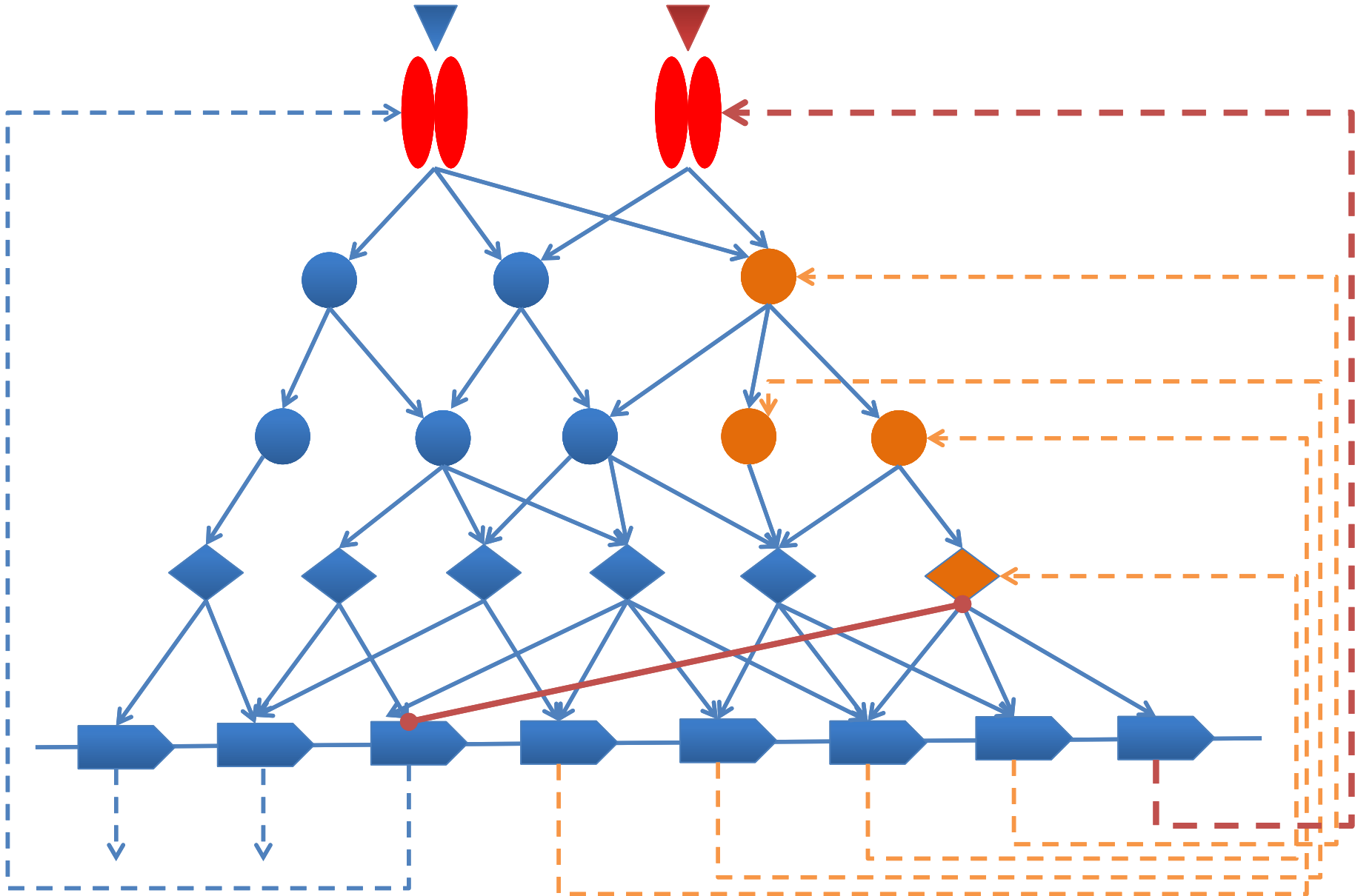
Master-regulator causes a lot of noise in the cell

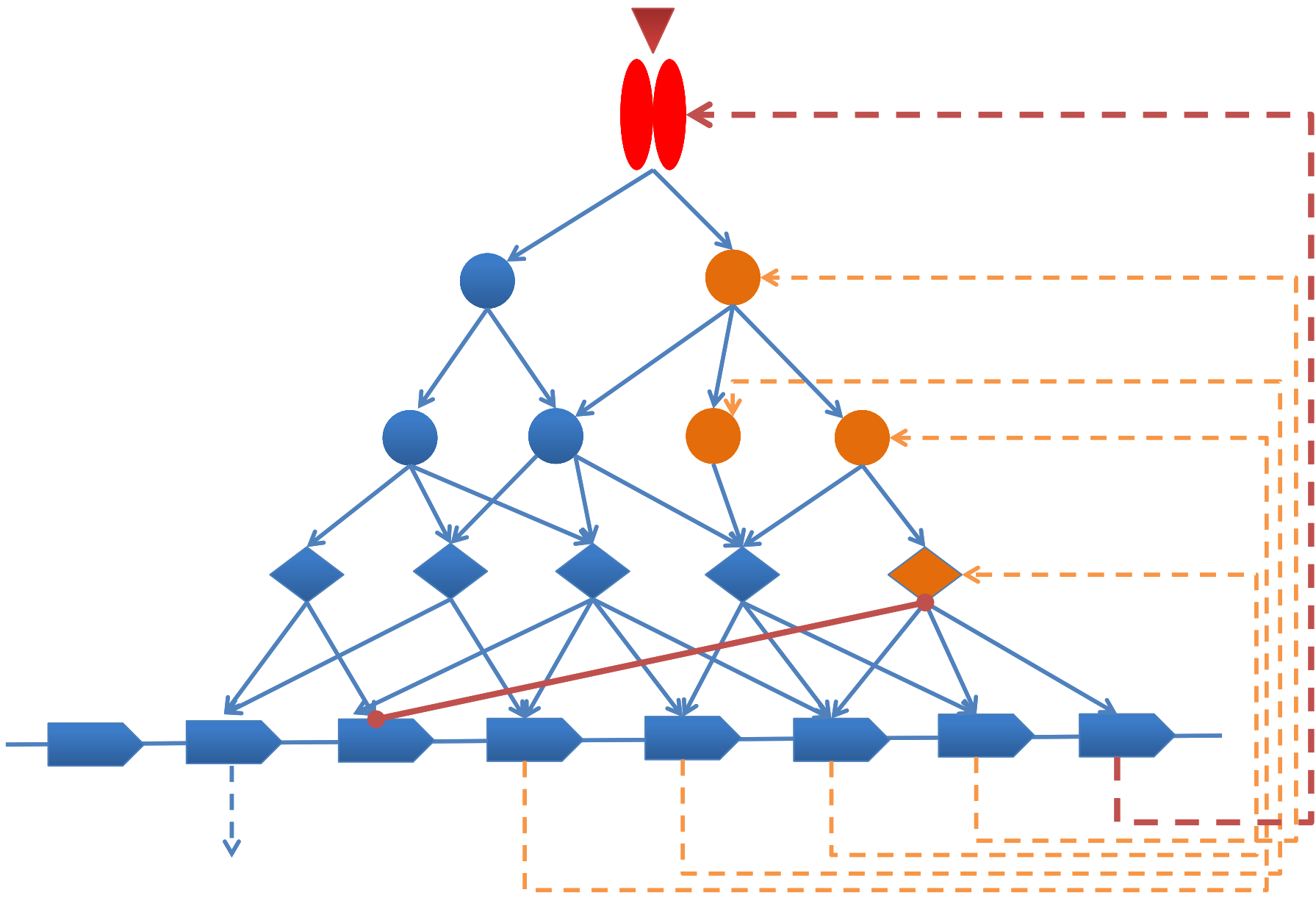


Master regulator



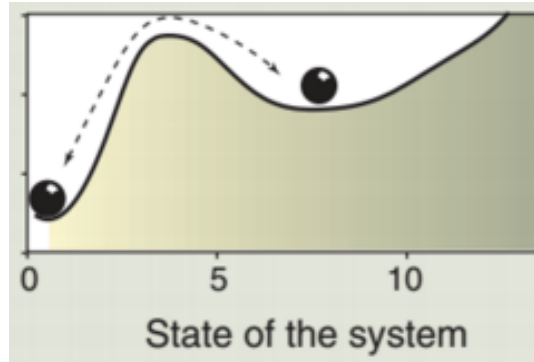
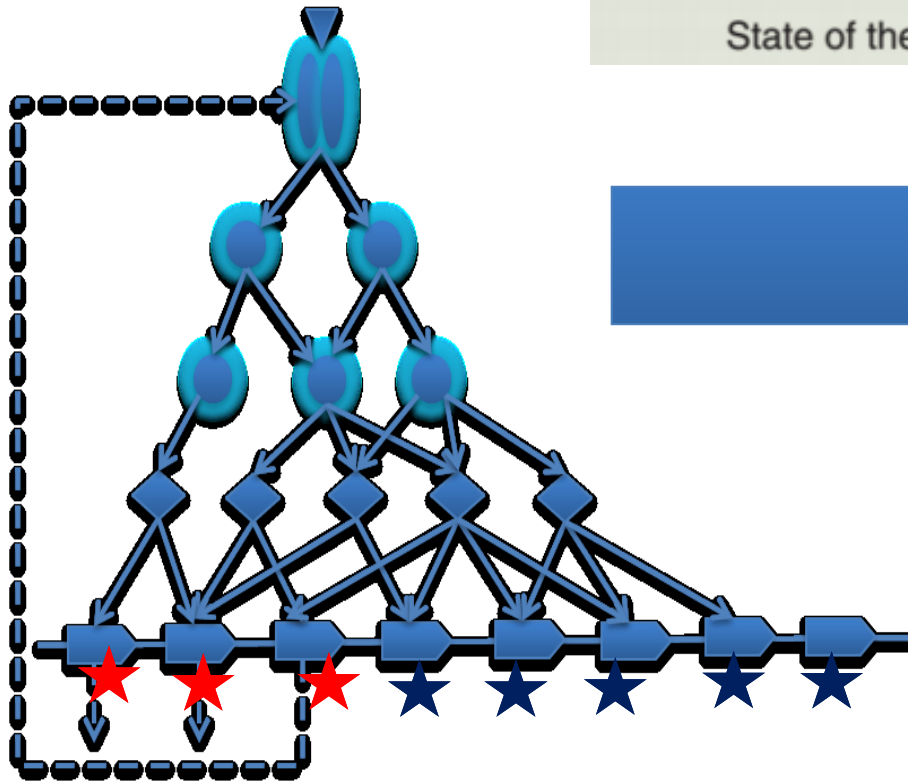
Walking pathways



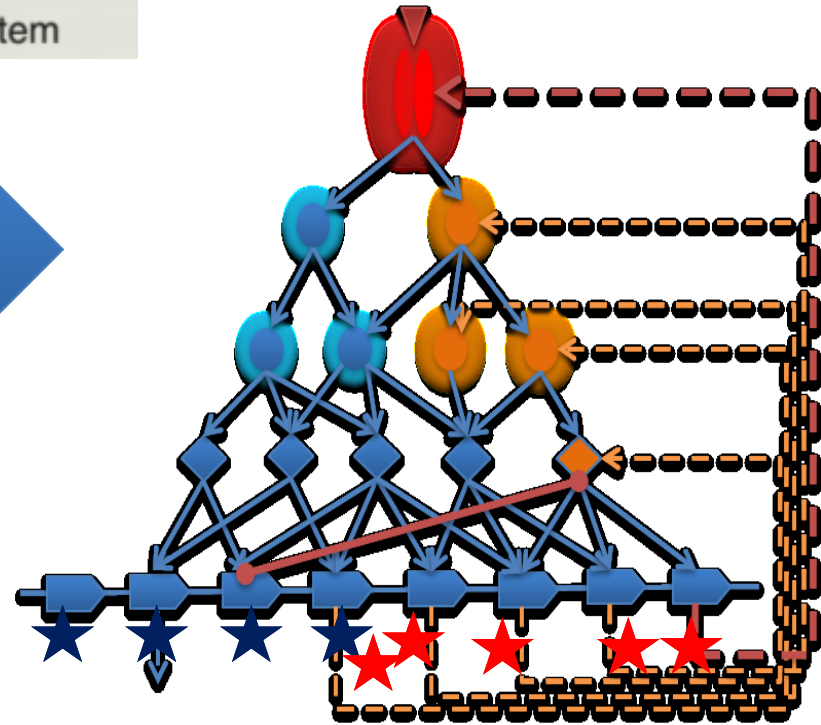


Pathway corruption

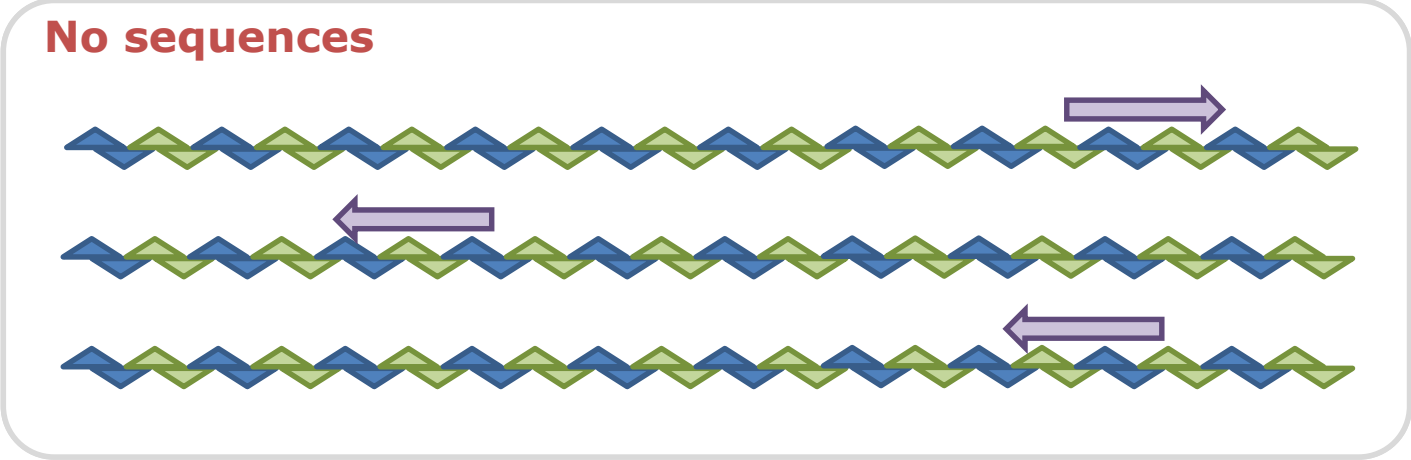
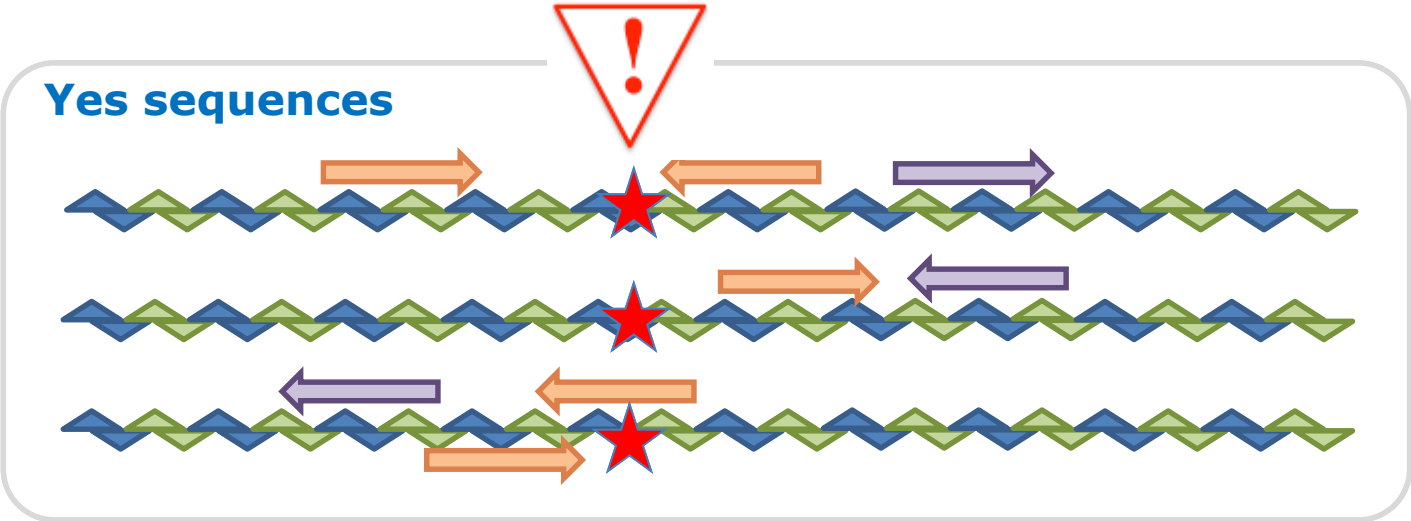
Healthy



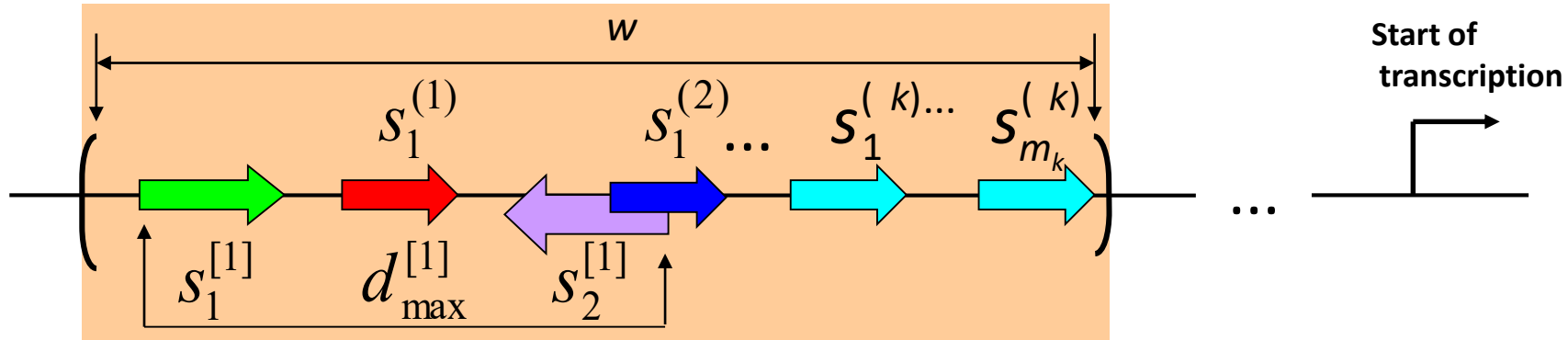
Cancer



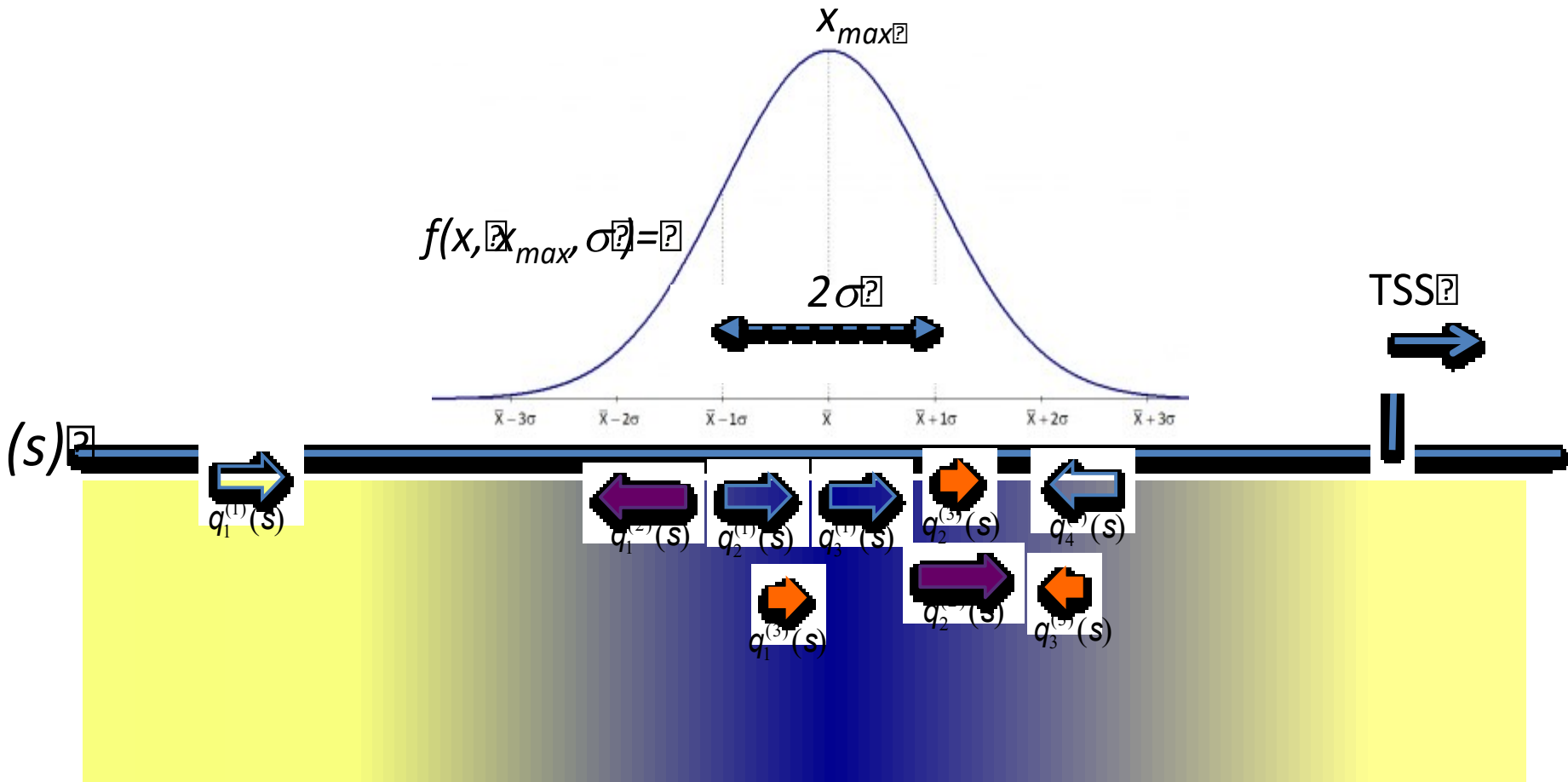
Motif enrichment analysis around regulatory mutations



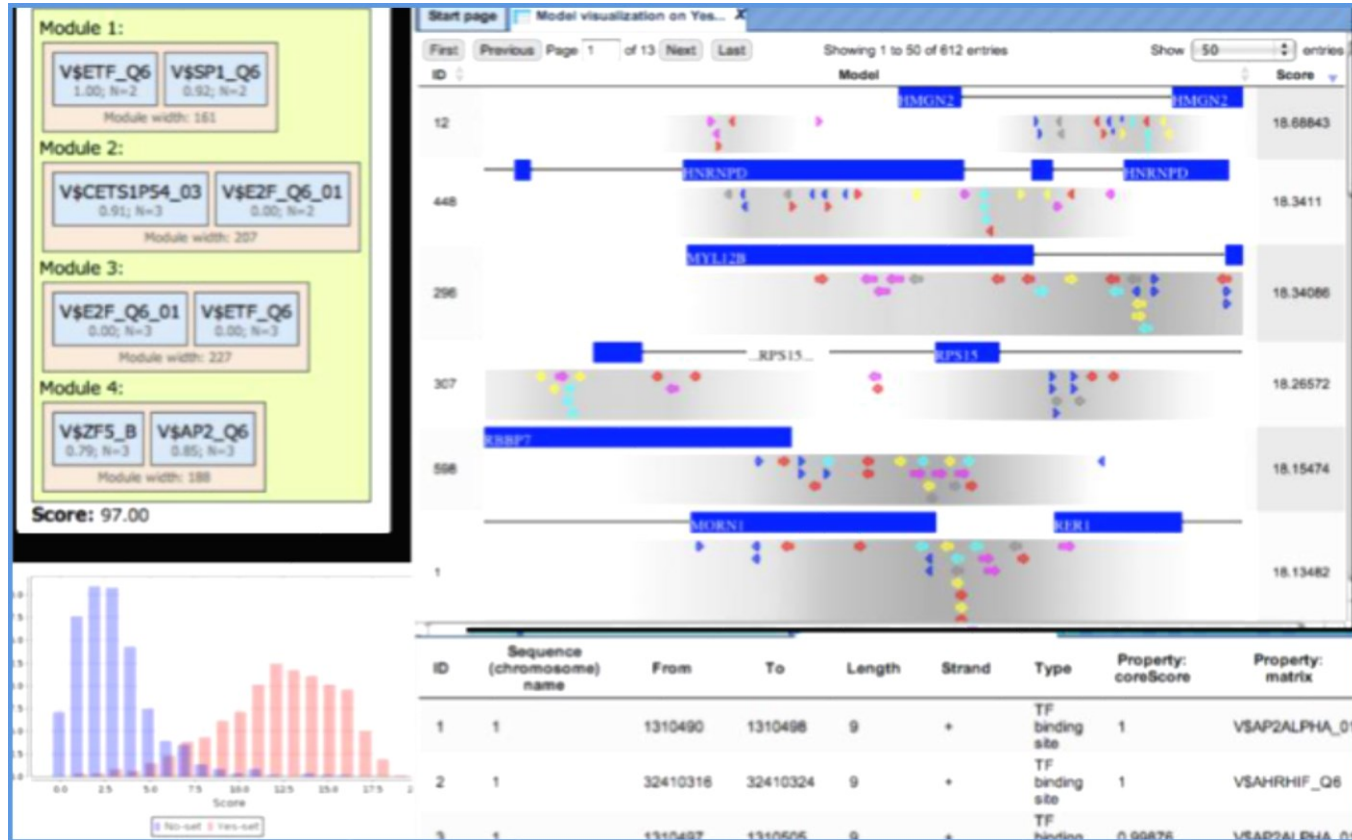
We construct a model of the disease-specific enhancer as a combination of transcription factor binding sites.



$d_{\max}^{[1]}$	$d_{\max}^{[1]}$...	$d_{\max}^{[R]}$	} The model parameters are found with the help of genetic algorithm
$q_{\text{cut-off}}^{(1)}$	$q_{\text{cut-off}}^{(2)}$...	$q_{\text{cut-off}}^{(k)}$	
$\phi^{(1)}$	$\phi^{(2)}$...	$\phi^{(k)}$	



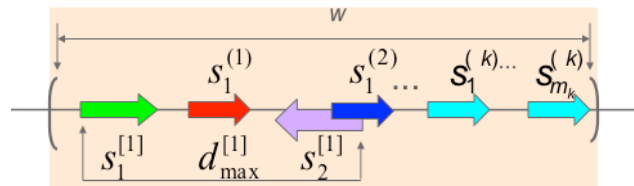
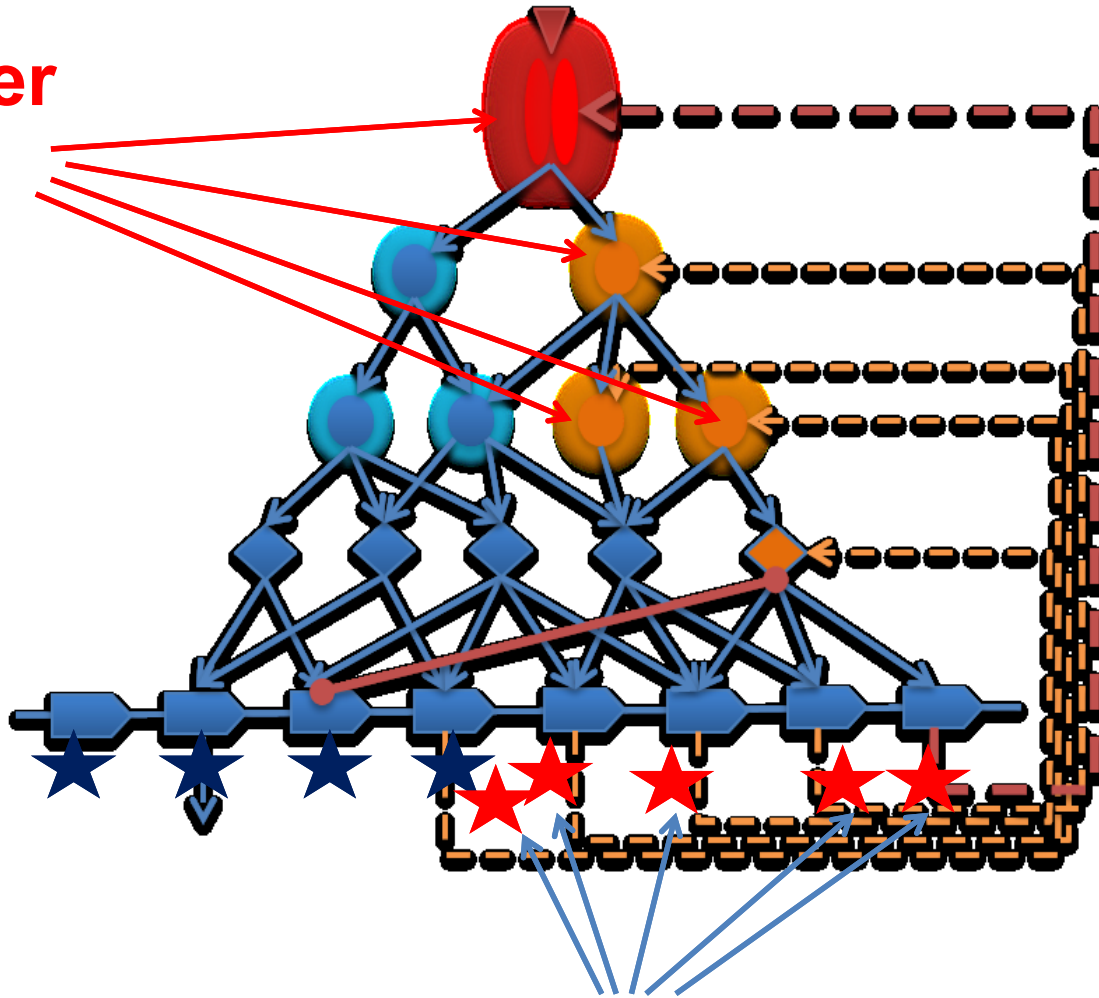
An example of disease-specific enhancers in regulatory regions of oncogenes.



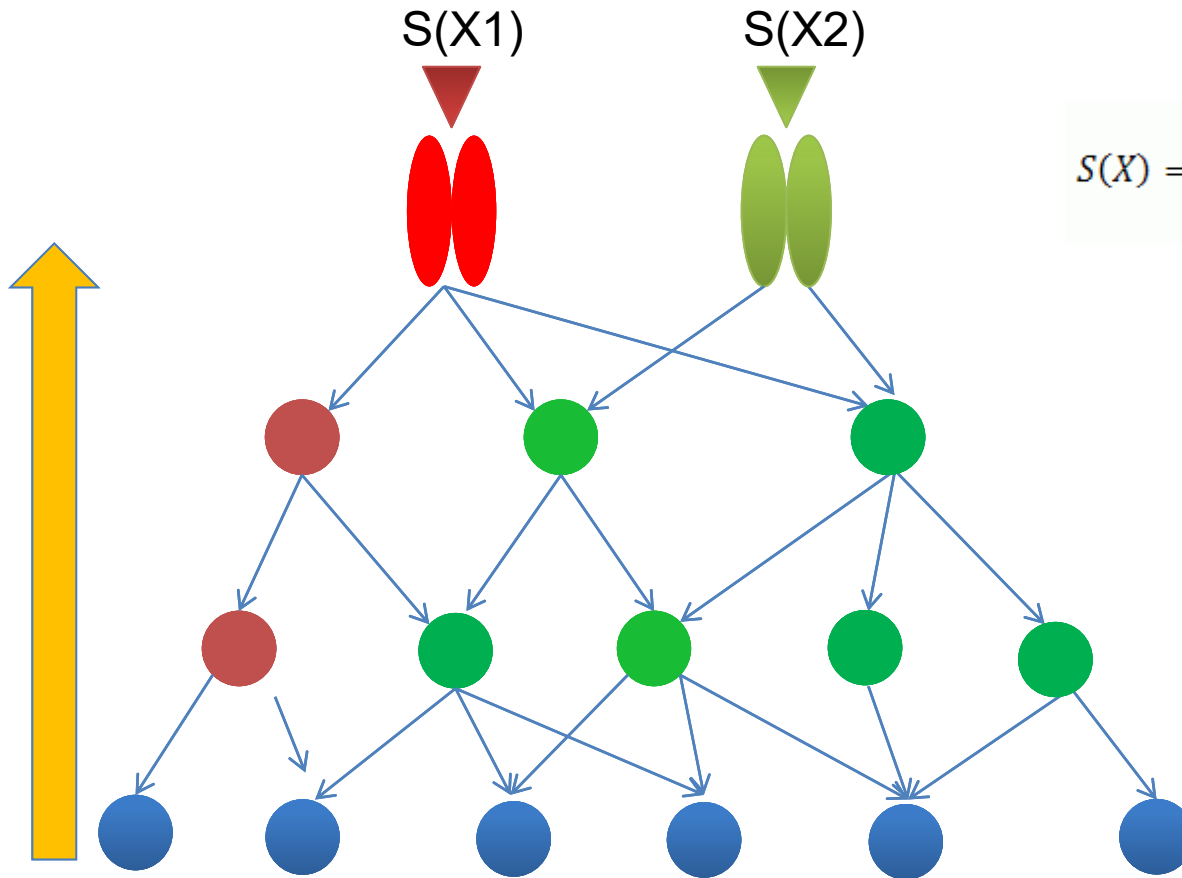
Binding sites for transcription factors that were included into the enhancer model

Cancer

Targets



Search for master regulators with context



$$S(X) = \sum_{r=1}^R \frac{M(X,r)}{M_{max}(r)} \cdot \frac{1}{1 + pN(X,r)/N_{max}(r)}$$

Where:

R - Max radius (input parameter)

p - Penalty (input parameter)

N(X,r) - total number of molecules reachable from key molecule X within the radius r.

N_{max}(r) - maximal value of N(X,r) over all key molecules X found for this radius.

M(X,r) - sum of w(X) for all hits reachable from key molecule X within the radius r, where w(X) - weight of hit X.

M_{max}(r) - maximal value of M(X,r) over all key molecules X found for this radius.

Kel, A., Voss, N., Jauregui, R., Kel-Margoulis, O. and Wingender, E.: Beyond microarrays: Find key transcription factors controlling signal transduction pathways *BMC Bioinformatics* 7(Suppl. 2), S13 (2006).



RESISTANCE

Scope: Format: Amount: GEO accession:

Series GSE11440

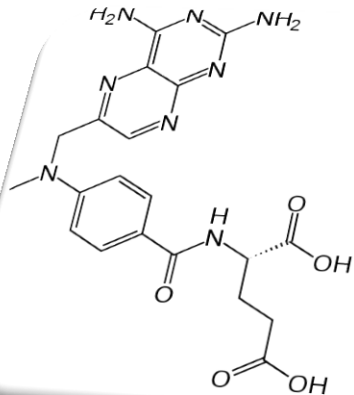
[Query DataSets for GSE11440](#)

Status Public on Sep 08, 2008
 Title Role of Caveolin 1, E-Cadherin, Enolase 2 and PKCa on resistance to methotrexate in human HT29 colon cancer cells
 Organism [Homo sapiens](#)
 Experiment type Expression profiling by array
 Summary A summary of the work associated to these microarrays is the following:

Methotrexate (MTX) is one of the earliest cytotoxic drugs used in cancer therapy, and despite the isolation of multiple other folate antagonists, methotrexate maintains its significant role as a treatment for different types of cancer and other disorders. The usefulness of treatment with methotrexate is limited by the development of drug resistance, which may be acquired through different ways. To get insights into the mechanisms associated with drug resistance and sensitization we have performed a functional analysis of genes deregulated in methotrexate resistant cells, either due to its co-amplification with the DHFR gene or as a result of a transcriptome screening using microarrays. Genes adjacent to dhfr locus and included in the 5q14 amplicon were overexpressed in HT29 MTX-resistant cells. Treatment with siRNAs against those genes caused a slight reduction in cell viability in both HT29 sensitive and resistant cells. On the other hand, microarray analysis of HT29 and HT29 MTX resistant cells unveiled overexpression of caveolin 1, enolase 2 and PKCa genes in treated cells without concomitant copy number gain. siRNAs against these three genes effectively reduced cell viability and caused a decreased MTX resistance capacity. Moreover, overexpression of E-cadherin, which was found underexpressed in MTX-resistant cells, also sensitized the cells toward the chemotherapeutic agent. We provide functional evidences indicating that caveolin 1 and E-cadherin may play a critical role in cell survival and may constitute potential targets for coadjuvant therapy.
 Keywords: DHFR, Methotrexate, drug resistance

Overall design

Two cell lines are compared in the study, which are HT29 colon cancer cells sensitive to methotrexate and HT29 cells resistant to 10e-5M MTX. Six



We took data on 3 MTX resistant patients versus 3 MTX sensitive and loaded them into geneXplain platform.

The screenshot displays the geneXplain platform interface. On the left, a navigation pane shows a tree structure under 'Data' with folders for 'Examples', 'Projects', and 'Public'. The 'Projects' folder is expanded to show a project for 'alexander.ke2@googlemail.com', which contains sub-folders for 'Data', 'Workflows', 'Journal', and 'tmp'. The 'Data' folder is further expanded to show 'Colon_cancer', which contains 'GSE11440_RAW' and 'Workflows'. The 'Workflows' folder is expanded to show 'Net2Drug'.

The main window is titled 'Start page' and 'Normalize Affymetrix exp...'. It contains a table of configuration options for normalizing Affymetrix expression data:

Experiment files	[3] GSM288501.CEL;GSM288502.CEL;GSM288536.C
Control files	[3] GSM288491.CEL;GSM288497.CEL;GSM288499.C
Method	MAS5
Background correction	MAS
Normalization method	quantiles
PM correction	pmonly
Summarization	mas
CDF version	<input type="checkbox"/> (select element)
Output table test data	.../Colon_cancer/Experiment normalized (MAS5) Auto
Output table control data	...ata/Colon_cancer/Control normalized (MAS5) Auto

Below the table is a 'Cancel!' button and a progress bar showing 3% completion. At the bottom, a log window displays the following information:

```
INFO - Normalize files...
INFO - Generating R command...
INFO - Platform detected: HG-U133_Plus_2
INFO - Connecting to R...
INFO - Invoking R command (that will take some time)...
```

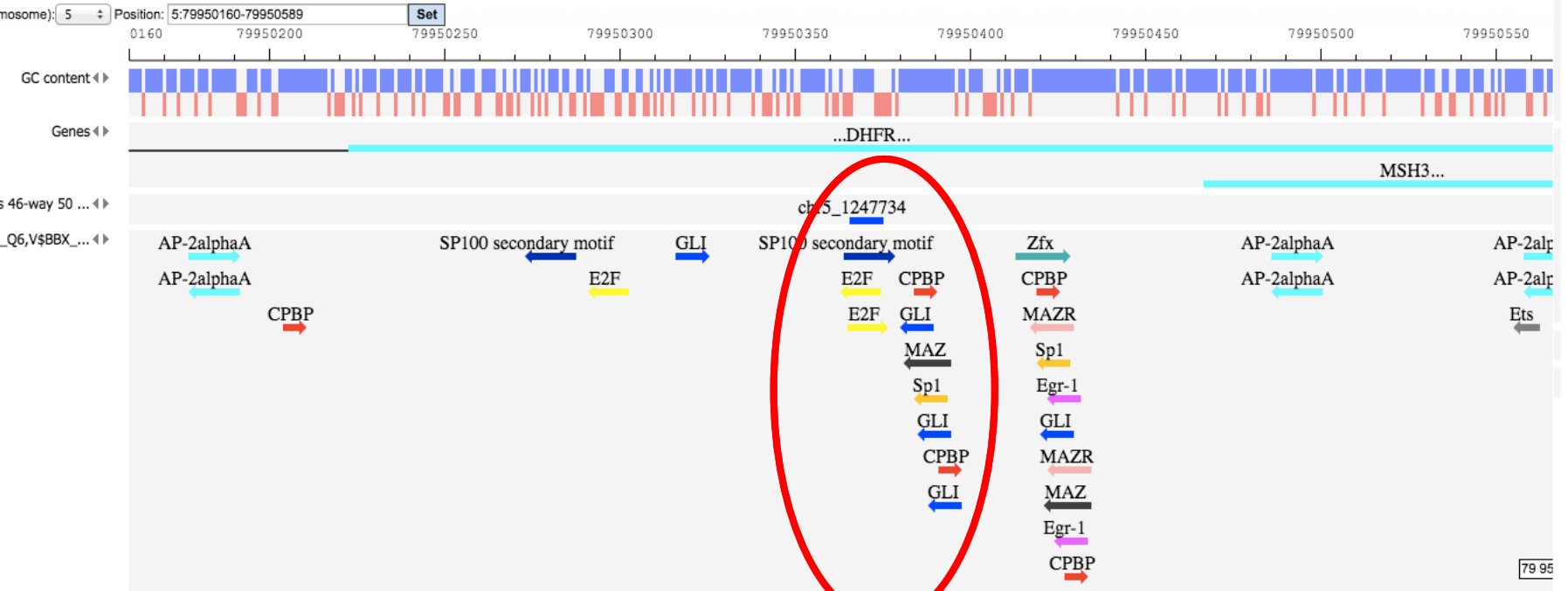
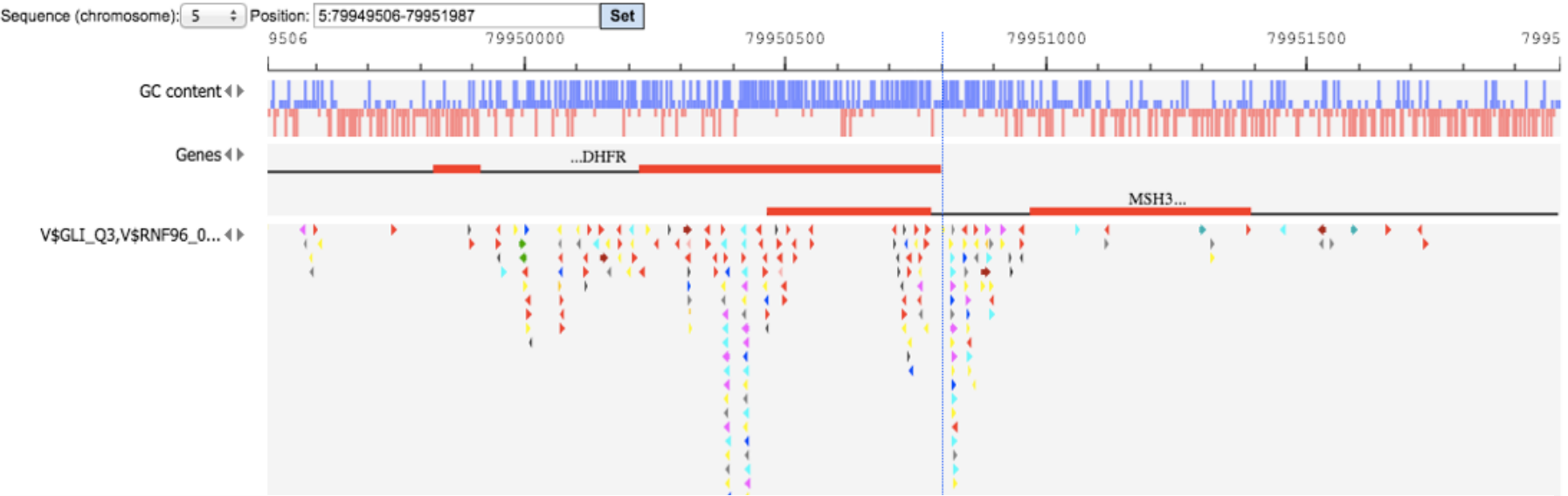
There is a ready pipeline (workflow) in geneXplain for discovery of targets and drug repurposing.

The screenshot displays the geneXplain software interface. At the top, a navigation bar includes 'Start page', '1 X', and 'Net2Drug X'. Below this, the 'Net2Drug' workflow is configured with 'Experiment' and 'Control' data paths set to 'data/Projects/alexander.kel2@googlemail.com'. Buttons for 'Run workflow' and 'Edit workflow' are visible.

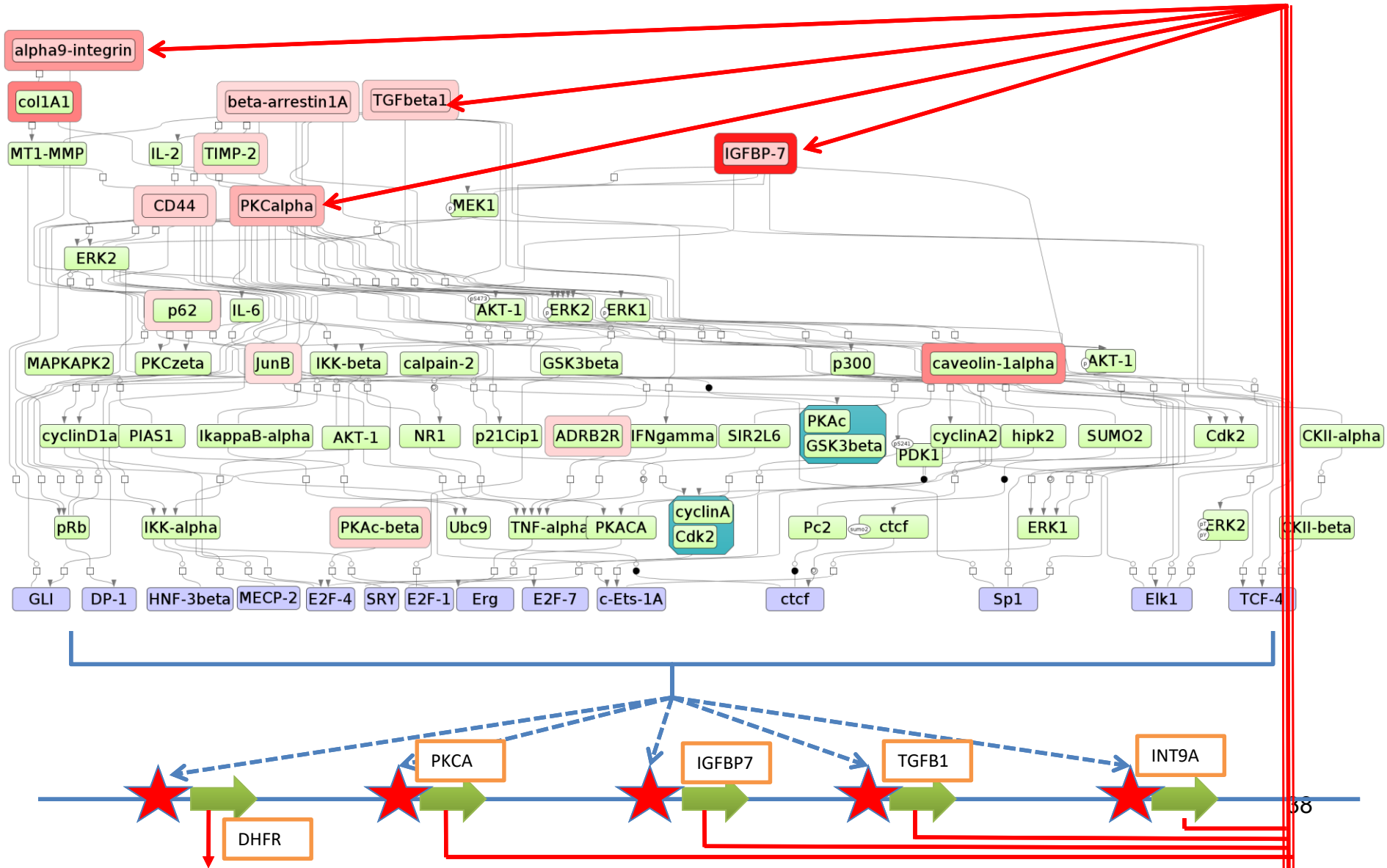
The main workspace is divided into several tabs: 'Overview', 'My description', 'Graph search', 'Script', 'Clipboard', and 'Tasks'. The 'Overview' tab is active, showing a flowchart of the workflow. The process starts with 'Experiment' and 'Control' data inputs. 'Experiment' feeds into 'Fold-Change calculation', which outputs 'FC'. 'Control' feeds into 'Up and Down Identification', which outputs 'p-value'. Both 'FC' and 'p-value' feed into 'Join tables', which outputs 'FC p-value'. 'FC p-value' feeds into 'Filter table(3)', which outputs 'Up'. 'FC p-value' also feeds into 'Filter table(4)', which outputs 'NC'. 'Up' and 'NC' feed into 'Filter table(2)'. 'Filter table(2)' outputs 'Molecules', which feeds into 'Regulator search'. 'Regulator search' outputs 'Molecules Upstream 6', which feeds into 'Filter table(5)'. 'Filter table(5)' outputs 'Molecules Upstream 6 hits 2', which feeds into 'Convert table'. 'Annotate table' outputs 'Gene set', which feeds into 'Search sites on gene set', which outputs 'Sites' and 'summary'. 'Sites' feeds into 'Filter table(2)'. 'summary' feeds into 'Filter table(2)'. 'Filter table(2)' outputs 'Molecules', which feeds into 'Regulator search'. 'Regulator search' outputs 'Molecules Upstream 6', which feeds into 'Filter table(5)'. 'Filter table(5)' outputs 'Molecules Upstream 6 hits 2', which feeds into 'Convert table'. 'Convert table' outputs 'Rank column'.

On the left side, a sidebar shows a file tree under 'Projects' for 'alexander.kel2@googlemail.com', including 'Data', 'Colon_cancer', 'GSE11440_RAW', and 'Control normalized (MASS)'. Below the file tree, a search bar is set to 'Default'. The 'Info' tab is active, displaying the following details:

- ID:** Net2Drug
- Title:** Regulator analysis
- Size:** 94
- Complete name:** data/Projects/alexander.kel2@googlemail.com/Data/Workflows/Net2Drug



Resistance to MTX



Prediction & Interpretation - C:\KEL\MTX\Top_200_Drugs_2010-153-selected_MTX.SDF. 39/151

Divalproex

Donepezil Doxazosin Doxycycline Duloxetine

Prediction & Interpretation - C:\KEL\MTX\prestwick_chemical_library (PASS2014).SDF. 744/1074

Zardaverine Memantine Hydrochloride Ozagrel hydrochloride Piribedil hydrochloride Nitrocaramiphen hydrochloride Nandrolone

Save TXT Save SD Clipboard Exclude Pa Pi Predicted value descending Show non predicted activities

0,925	0,003	Phosphodiesterase inhibitor	Metabolism: 19	Transport: 2	Gene Expression:
0,899	0,005	Antineoplastic (liver cancer)	Effect: 66	Mechanism: 312	Toxicity: 89

Pa Pi <GENERIC_NAME>

0,153	0,040	Divalproex
0,143	0,060	Ibuprofen
0,139	0,070	Naproxen
0,137	0,078	Gabapentin
0,131	0,097	Clonidine

Pa Pi <chemical_name>

0,867	0,002	Zardaverine
0,217	0,009	guaiacol
0,144	0,018	Guaifenesin
0,108	0,027	Trimetazidine dihy
0,058	0,056	Trimethoprim

Prediction & Interpretation - C:\KEL\MTX\HMDB metabolites_MTX.SDF. 1563/15866

Nicotinamide N-oxide Fucoxanthin cis-5,6-Dihydro-5,6-dihydroxy-car Prostaglandin A₂ Myricetin

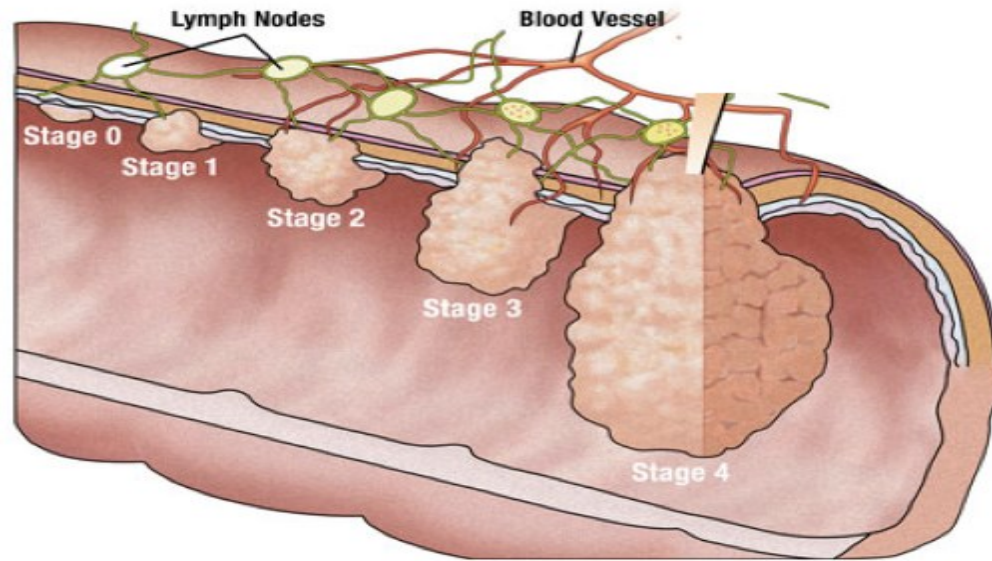
1) Divalproex, which is also known as valproic acid, is an old drug primarily used to treat epilepsy and bipolar disorder and to prevent migraine headaches. Recently a number of clinical trials were performed with this drug and they confirmed its efficacy for treatment of **Acute Myeloid Leukaemia, Cervical cancer and Breast cancer**.

2) There is a number of recent studies confirming the potential use of **zardaverine** in cancer therapy, against **hepatocellular carcinoma and against Chronic Lymphocytic Leukemia**.

3) Nicotinamide is known to **sensitize** a number of rodent tumors to single dose of **radiation**. Its combination with carbogen results in large **enhancement of tumor response** to certain treatment and it was confirmed in a **clinical trials**.



Colorectal cancer



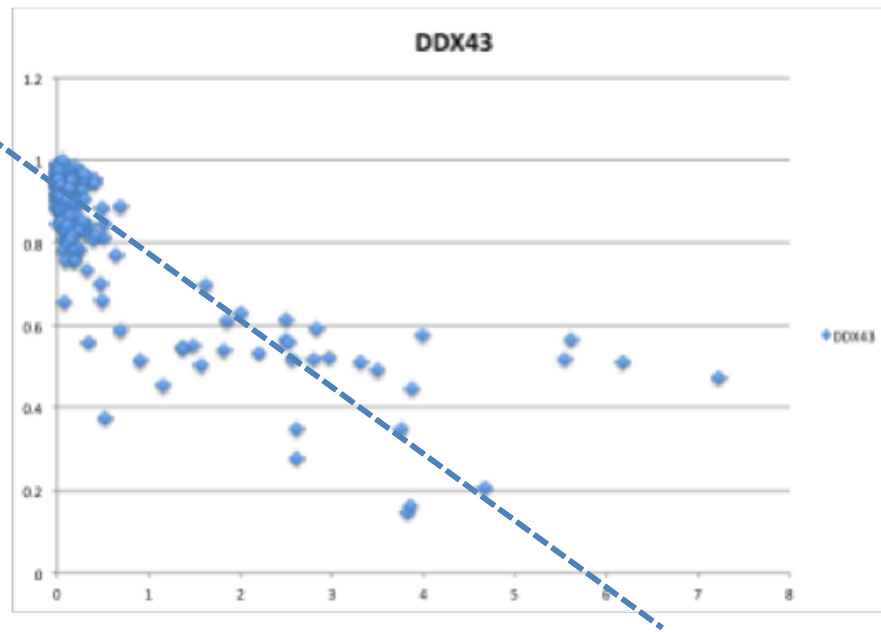
SYSCOL

Systems Biology of Colorectal Cancer

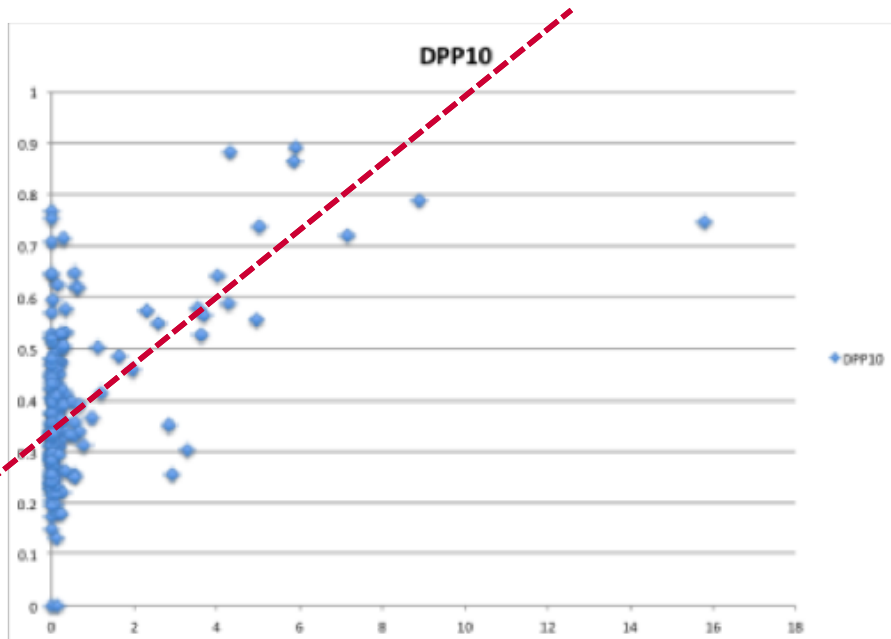


Gene symbol	Methylation vs Expression metastasis fc04 corr03: Count	Methylation vs Expression metastasis fc04 corr03: Count in exons	Methylation vs Expression metastasis fc04 corr03: Count in introns	Methylation vs Expression metastasis fc04 corr03: Count in 5'	Methylation vs Expression metastasis fc04 corr03: Count in 3'	Methylation vs Expression metastasis fc04 corr03: Schematic
GFRA1	14	6	2	6	0	
TENM3	12	3	9	0	0	
TBX15	11	0	11	0	0	
TIAM1	11	1	10	0	0	
DPP10	11	4	7	0	0	
AP000251.2	11	1	4	6	0	
RP11-60L3.1	9	1	3	5	0	
DDX43	8	4	0	4	0	
STRA6	8	6	2	0	0	
GNG3	8	2	0	0	6	
AC004696.2	8	7	0	1	0	
BSCL2	8	6	2	0	0	
ZNF667	8	2	5	1	0	
OOEP	8	7	0	1	0	
RP11-831H9.16	8	0	8	0	0	
snoU13	8	0	0	8	0	
ADGB	7	4	0	3	0	
ERBB2	7	3	1	2	1	
SYCE2	7	1	0	6	0	
ZCCHC24	7	6	0	0	1	
AC066593.1	7	1	0	6	0	
RP11-342M3.5	7	2	4	1	0	
RP11-715G15.1	7	2	1	4	0	
MIR5695	7	0	0	7	0	
PCDHB3	6	3	0	3	0	

**Meth.
vs
Expr.**



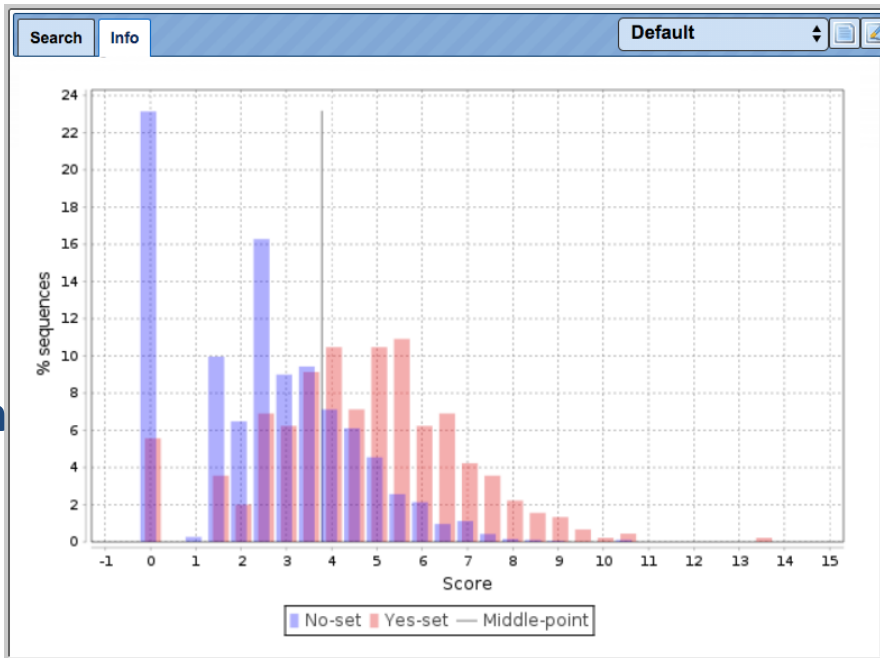
**Negative
correlation**



**Positive
correlation**

Meth.
vs
Expr.

Negative
correlation



Site colors My description

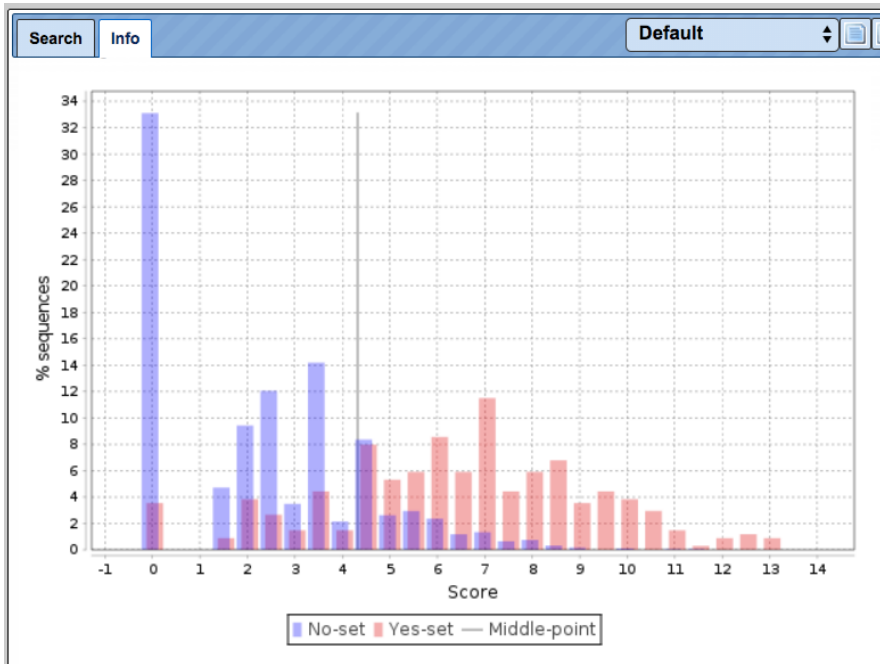
information is shown:
-PWMs producing matches,
-number of individual matches for each PWM,
-score of the best match.

Module 1:

V\$E2F6_01 0.00; N=3	V\$ZNF625_02 0.82; N=3	V\$HNF4_Q6 0.92; N=2	V\$HMGYI_01 0.95; N=3
V\$HNF3_Q6_01 0.85; N=2	V\$ZNF43_01 0.77; N=2	V\$AML_Q6 0.82; N=2	V\$GR_Q6 0.87; N=2
V\$FOSL2_02 0.83; N=3	V\$TGIF_02 0.78; N=3	Module width: 104	

Model score (-p*log10(pval)): 35.15
 Wilcoxon p-value (pval): 7.46e-71
 Penalty (p): 0.501
 Average yes-set score: 4.62
 Average no-set score: 2.56
 AUC: 0.77
 Middle-point: 3.78
 False-positive: 25.07%
 False-negative: 33.41%

Positive
correlation



Site colors My description

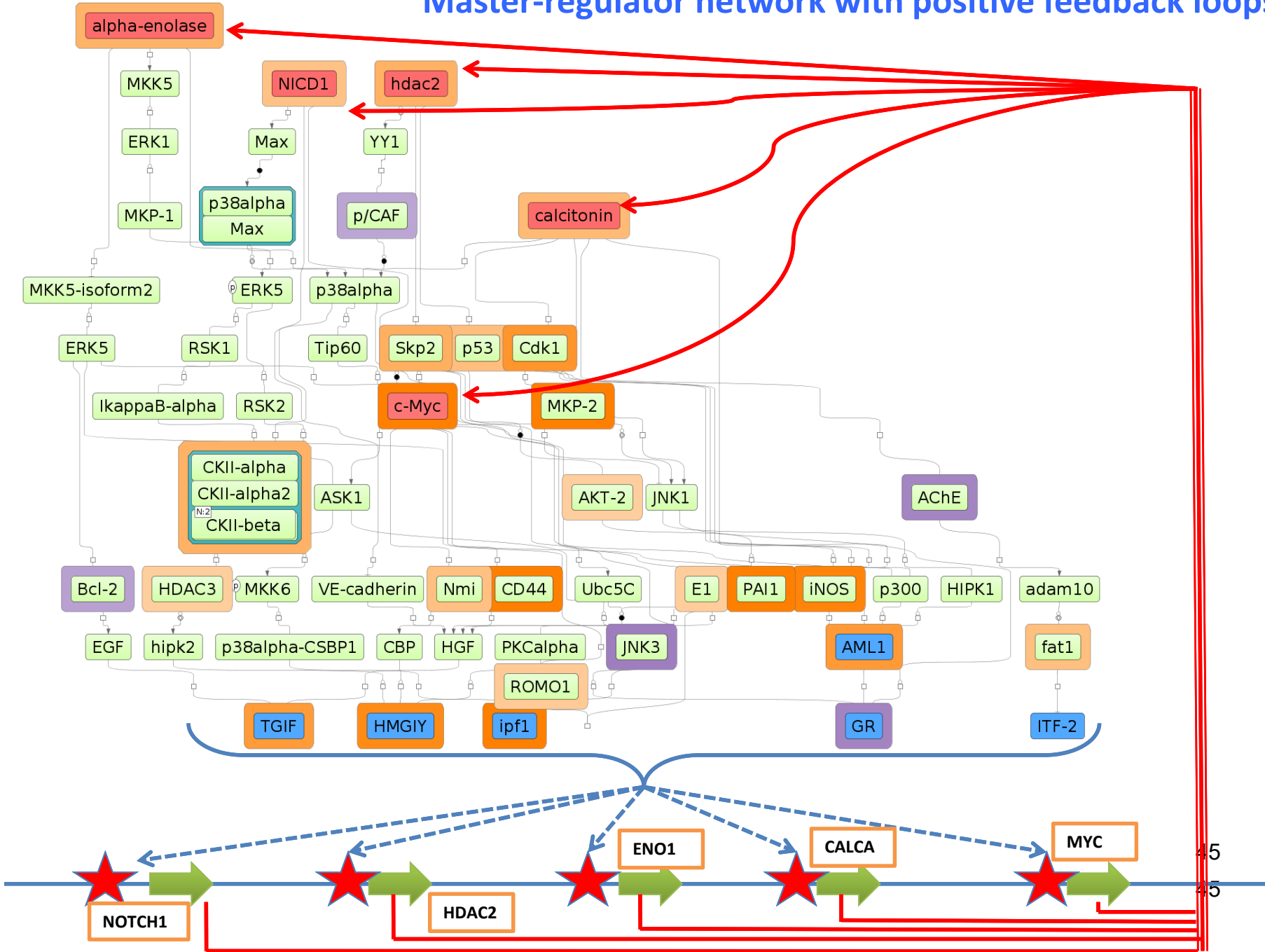
information is shown:
-PWMs producing matches,
-number of individual matches for each PWM,
-score of the best match.

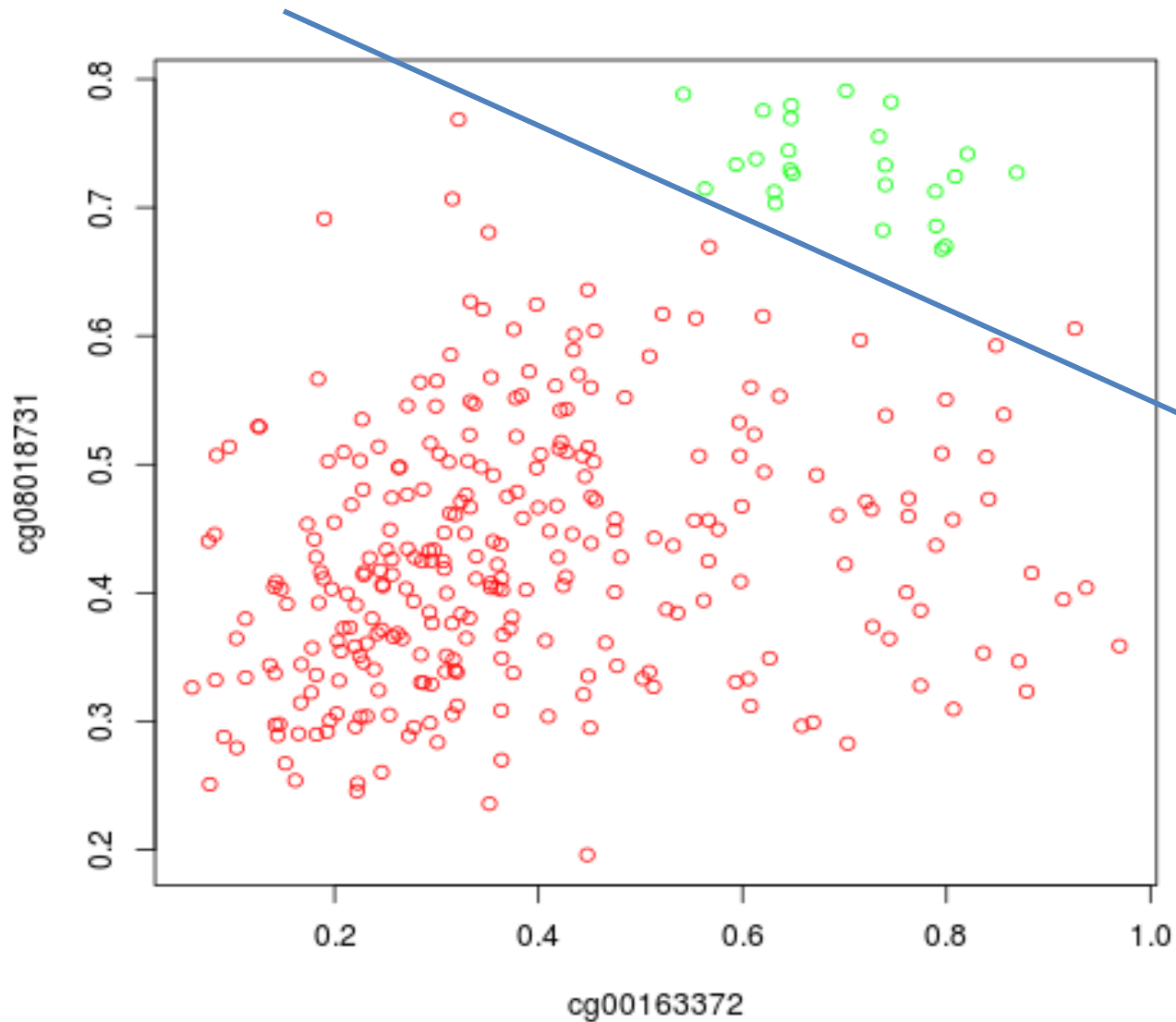
Module 1:

V\$IRX5_01 0.80; N=3	V\$GR_Q6_01 0.95; N=3	V\$TAL1BETAITF2_01 0.78; N=3	
V\$HNF4_Q6 0.90; N=3	V\$HMGYI_Q3 0.00; N=3	V\$MAFB_Q4 0.91; N=3	V\$HNF3_Q6_01 0.85; N=3
V\$BATF_03 0.80; N=2	Module width: 165		

Model score (-p*log10(pval)): 56.15
 Wilcoxon p-value (pval): 1.63e-105
 Penalty (p): 0.536
 Average yes-set score: 6.45
 Average no-set score: 2.41
 AUC: 0.87
 Middle-point: 4.32
 False-positive: 19.23%
 False-negative: 18.29%

Master-regulator network with positive feedback loops





DNA methylation values of two markers **cg00163372** (in gene *MYC*) and **cg08018731** (in gene *NOS3*). The red dots shows values obtained in tumor samples, the green dots shows values for the normal samples.

6 CpG methylation biomarkers for yearly detection of CRC

Probe ID	Chromos
cg01421342	
cg06972019 (CpG No.3)	
cg00163372	
cg02991571	
cg24093411	
cg02612618	

?

93% success rate on independent cohort (from Siberia)



Signatures for diagnosis of colorectal cancer

Diagnosis of colorectal cancer on the base of DNA methylation markers.

1. Info about markers

The set of markers consists of the six CpG methylation markers. Markers » Histogram »

2. Upload data

The file should contain a table with the values of the markers. Markers should be placed in the rows of the table, and samples should be in the columns.

You have uploaded Example_markers6CpG

3. Select samples for analysis

Sample3
 Sample4
 Sample5
 Sample6

Completed

Is that all...?

Hm, but..how to cure
my patient?





Right ovarian tumor

UP 1162 genes

DOWN: 2060 genes

Number of "mutations": 80,449

LogFC Endom	LogFC OverLinks	LogFC OverRight	Gene symbol	SNP_indels.vcf different from Normal_from_exome: Schematic
8.92186	0	9.22477	ABHD15-AS1	
5.12912	6.67793	9.06892	TRBV11-1	
5.90125	5.75969	9.01443	TRBV5-2	
4.26481	7.92993	8.82873	TRBV10-2	
5.30409	8.99687	8.76248	MUC16	
2.73426	5.19135	8.69899	RPL23AP39	
2.25818	7.99442	8.59259	CTC-501O10.1	
7.18713	9.85593	8.50615	LINC01508	
2.69963	8.22939	8.41096	TRBV10-1	
4.24128	7.59892	8.39401	TRBV5-3	
5.392	7.33392	8.37958	FGL1	
0	6.0722	8.27947	RP1-177I10.1	
0	5.32324	8.27573	RPL21P17	
5.84437	7.26608	8.25135	TRBV7-3	
1.52818	6.66159	8.10189	IMPG2	
7.5239	7.14924	8.06829	RP11-105N14.1	
3.69525	6.02003	7.97286	CP	
3.01768	7.72384	7.97073	LINC01033	
5.74077	4.56477	7.90199	RP11-339D20.1	
0	0	7.7779	RP11-277J6.2	
0	0	7.69918	RP11-757O6.4	
8.33921	0	7.67474	OSBPL9P5	

LogFC Endom	LogFC OverLinks	LogFC OverRight	Gene symbol	SNP_indels.vcf different from Normal_from_exome: Schematic
0.90363	0.00686	-10.17248	RNU6-331P	
-10.14128	-10.10633	-10.07692	Y_RNA	
-6.26598	-9.84722	-9.82177	IGHV3-20	
-5.66689	-9.82102	-9.79517	LINC01589	
-9.83347	-6.28776	-9.77982	ABCC8	
-9.61872	-9.58082	-9.55237	RNU6-157P	
-9.30059	-10.76241	-9.5176	HPSE2	
-9.48781	-9.45797	-9.43238	UBE2Q2P6	
-2.4301	-8.97405	-8.94721	RP11-330C7.4	
-6.76285	-8.95668	-8.93133	RP11-246A10.1	
-8.95974	-3.27728	-8.90503	RP5-965F6.2	
-7.80183	-6.2067	-8.89261	HTR2A	
-8.83414	-8.80332	-8.77838	SNORA76	
-3.62202	-1.85801	-8.76298	FTLP8	
-8.81165	-8.78394	-8.76	IGHV6-1	
-1.21351	-0.55545	-8.75963	RP11-1029M24.4	
-8.76935	-8.73709	-8.71226	RP11-731D1.3	
0.94001	1.32524	-8.51803	snoU109	
-9.80146	-7.07864	-8.48572	CNN1	
-4.07047	-8.46437	-8.43826	LINC01214	
-0.68333	-8.08378	-8.05691	RNU2-65P	



65 HIGH impact mutations in proteins.

[Edit](#)

[First](#)
[Previous](#)
Page of 2
 [Next](#)
[Last](#)
 Showing 1 to 50 of 65 entries

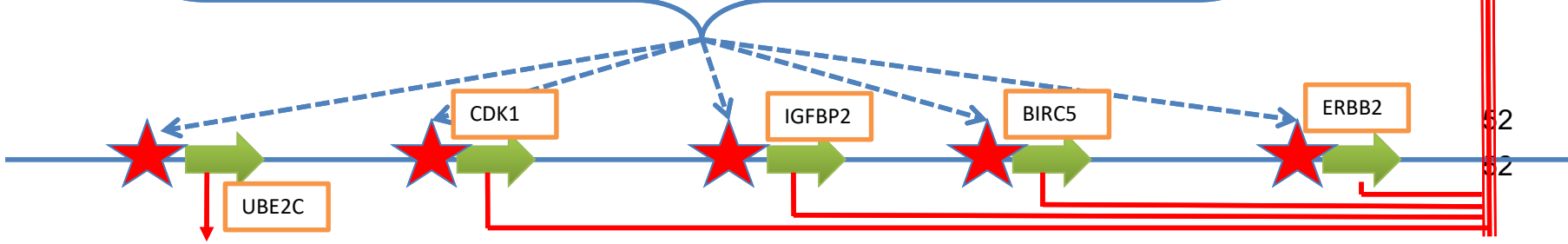
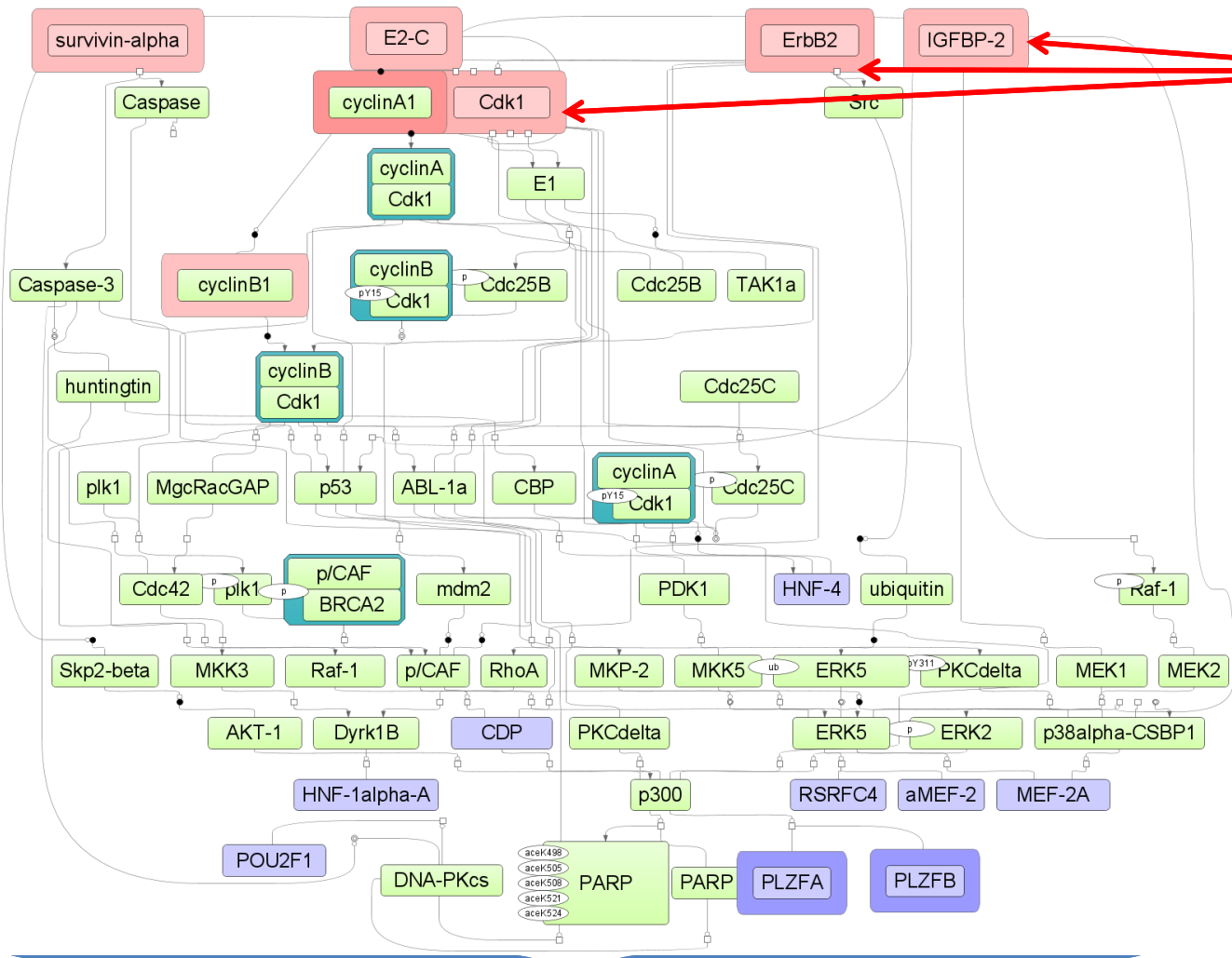
ID	Gene description	Gene symbol	Location	Allele	Consequence	IMPACT
ENSG00000005187	acyl-CoA synthetase medium-chain family member 3	ACSM3	16:20790685-20790685	G	stop_gained	HIGH
ENSG000000031691	centromere protein Q	CENPQ	6:49492214-49492214	G	stop_gained	HIGH
ENSG000000055609	lysine (K)-specific methyltransferase 2C	KMT2C	7:152247986-152247986	T	NMD_transcript_variant,frameshift_variant,stop_gained	HIGH
ENSG00000100225	F-box protein 7	FBXO7	22:32479203-32479203	A	start_lost	HIGH
ENSG00000104938	C-type lectin domain family 4 member M	CLEC4M	19:7766042-7766042	T	stop_gained	HIGH
ENSG00000107651	SEC23 interacting protein	SEC23IP	10:119932315-119932315	T	stop_gained	HIGH
ENSG00000108828	vesicle amine transport 1	VAT1	17:43015940-43015940	A	NMD_transcript_variant,splice_donor_variant	HIGH
ENSG00000111424	vitamin D (1,25-dihydroxyvitamin D3) receptor	VDR	12:47879112-47879112	G	NMD_transcript_variant,start_lost	HIGH
ENSG00000113013	heat shock protein family A (Hsp70) member 9	HSPA9	5:138556493-138556493	A	stop_gained	HIGH
ENSG00000114544	solute carrier family 41 member 3	SLC41A3	3:126006425-126006429	-	frameshift_variant	HIGH
ENSG00000116039	ATPase, H+ transporting, lysosomal 56/58kDa, V1 subunit B1	ATP6V1B1	2:70935956-70935956	C	NMD_transcript_variant,start_lost	HIGH
ENSG00000116809	zinc finger and BTB domain containing 17	ZBTB17	1:15945140-15945140	A	stop_gained	HIGH
ENSG00000117676	ribosomal protein S6 kinase, 90kDa, polypeptide 1	RPS6KA1	1:26553429-26553429	C	NMD_transcript_variant,stop_lost	HIGH
ENSG00000121052	eosinophil	EBV	17:58105140-58105140	A	stop_gained	HIGH



Mutations in promoters of up-regulated genes in Right Ovarian Tumor

ID	LogFC Endom	LogFC OverLinks	LogFC OverRight	gene	SNP_indels.vcf different from Normal_from_exome in UP_0.05_4100: Count	SNP_indels.vcf different from Normal_from_exome in UP_0.05_4100: Schematic	SNP_indels.vcf different from Normal_from_exome in UP_0.05_4100: Structure
ENSG00000205174	0	4.2646	7.3861	C7orf66	1		5'
ENSG00000136574	0	7.90306	7.2932	GATA4	1		5'
ENSG00000236407	0	4.40921	7.15112	HMGB1P18	3		5' x 3
ENSG00000211716	6.10002	6.21719	7.14532	TRBV9	5		5' x 5
ENSG00000225155	4.16448	9.19819	7.11554	TOMM22P5	2		5' x 2
ENSG00000274874	4.71019	5.67645	7.02952	RP11-214K3.25	1		Intron
ENSG00000258897	0	5.3355	7.01461	EGLN3-AS1	1		5'
ENSG00000248176	4.0058	4.05574	6.61439	RP11-472K22.1	1		5'
ENSG00000275666	0	5.15062	6.61067	KCNQ1OT1_1	7		5' x 7
ENSG00000259138	5.94961	6.95123	6.59956	RP11-950C14.7	1		5'
ENSG00000102854	3.37773	3.47242	6.5814	MSLN	1		5'
ENSG00000161905	-0.67665	5.47409	6.56401	ALOX15	2		5' x 2
ENSG00000130720	5.80285	4.01709	6.358	FIBCD1	1		5'
ENSG00000259687	4.59765	4.79779	6.35417	LINC01220	3		5' x 3
ENSG00000230968	6.01051	5.36605	6.25494	AC084149.1	1		5'
ENSG00000254370	5.78995	4.81077	6.23202	RP11-181B11.1	5		5' x 5
ENSG00000278035	6.40422	5.57588	6.23069	RP11-234K24.6	2		5' x 2
ENSG00000260240	0	0	6.16965	APOOP5	1		5'
ENSG00000258864	8.09974	5.84099	6.16307	CTC-554D6.1	3		5' x 3
ENSG00000233975	4.27087	7.0631	6.155	RP11-288L9.1	1		Intron
ENSG00000257316	6.1845	4.04627	5.81065	RP11-267D19.2	2		5' x 2
ENSG00000225598	4.37132	5.39961	5.78362	RP11-339D23.1	7		5' x 7
ENSG00000174844	3.0074	1.84431	5.4673	DNAH12	1		5'
ENSG00000234965	0	2.64519	5.4605	SHISA8	1		5'

Master-regulator in Right Ovarian Tumor





**STOP
CANCER!**

We learned....

Promoters do not exist !

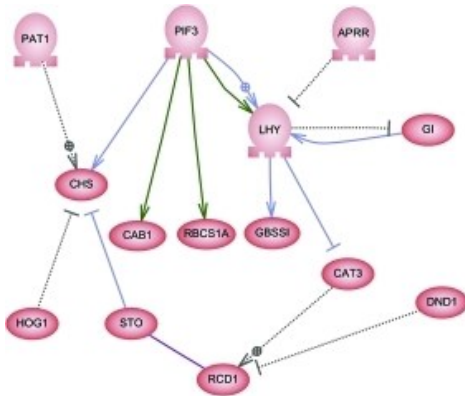
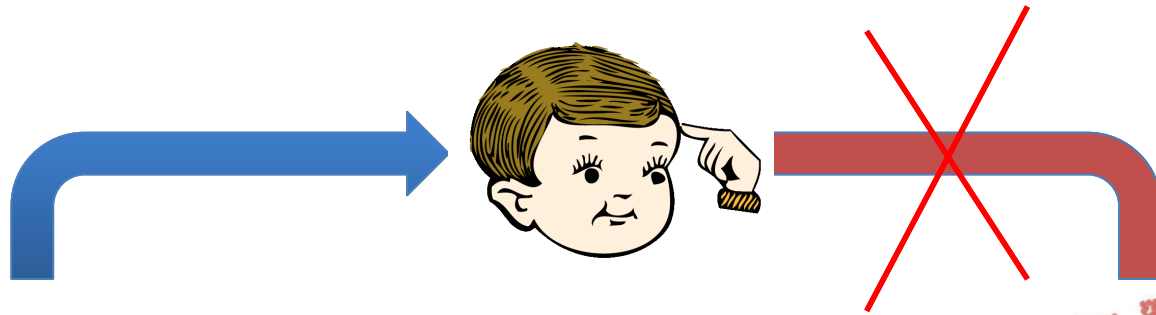
Sites do not exist !

Pathways do not exist !

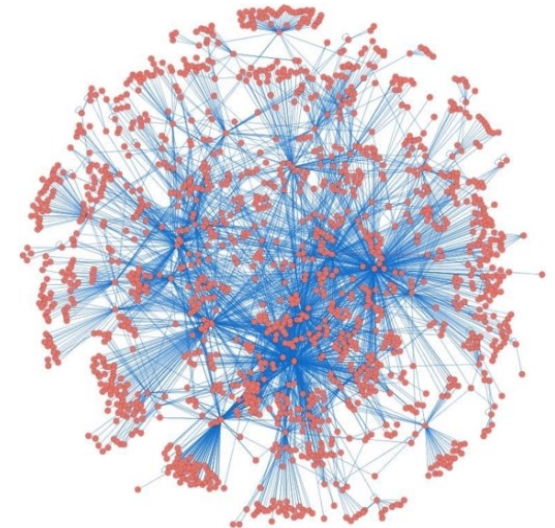
Things are very complex....So what?

Simplified model
fits into our brain!

Realistically
big and complex
model doesn't
fit into our brain

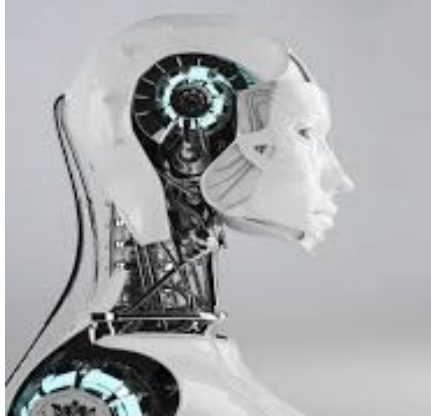
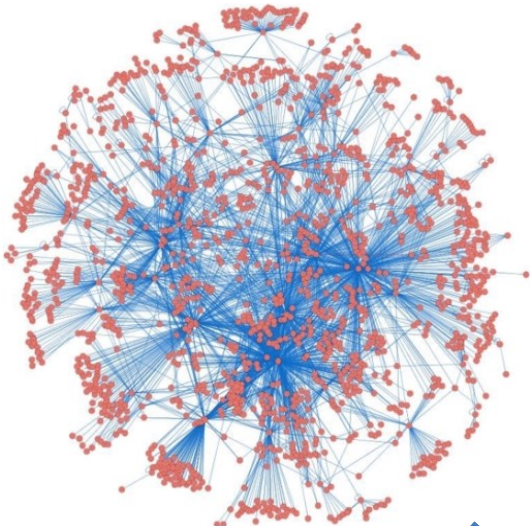


Model

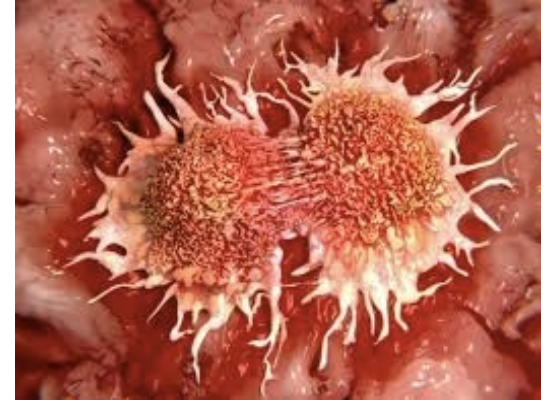


Reality

How to treat cancer?



Cancer



IN A HUGE BREAK GOOGLE'S AI BEATS HUMAN PLAYERS AT THE GO

THEIR BOUT

News & Analysis

Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

DeepMind

MUST READ **NEW DOCUMENTS REVEAL FBI PAID GEEK SQUAD REPAIR STAFF AS INFORMANTS**

Gaming AI beats human top scores by cheating

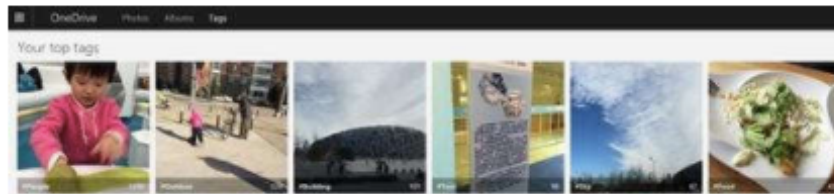
The artificial intelligence system had no qualms about exploiting ancient bugs to win.



By [Charlie Osborne](#) for [Between the Lines](#) | March 2, 2018 -- 11:12 GMT (11:12 GMT) | Topic: [Innovation](#)

See a stream of papers claiming they have one-upped humans too.

For instance, only 6 days after Microsoft announced it had beat the human benchmark of 5.1% errors with a 4.94% error grabbing neural network, Google announced it had one-upped Microsoft by 0.04%.



▲ AlphaZero's victory is just the latest in a series of computer triumphs over human players since Computer programs have been able to beat the best IBM's Deep Blue defeated Garry Kasparov in 1997. Photograph: 18percentgrey / Alamy/Alamy

Genome enhancer

Enhance drug discovery with genomics data.

Please sign in

Sign In

Request access



Wizard Description X

Project: Demo - Transcriptomics with Epigenomics

Enhance your genome

Drop files here

+ Add files...

Remote upload

< Prev

Create my
project

Upload
my data

Describe
my data

Start
analysis

See
report

Next >

Identification of master-regulators in gene regulation transduction pathways.

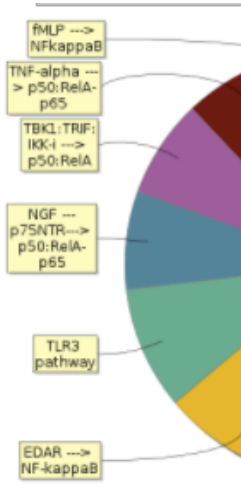
Alexander geneXplai
Data received on

Data

For this study the following

Table 1. Experimental datasets used

File name
Normalized (RMA) (1)



Summary

In this repository we automatized the goal of targets for factors (TFs) search for the identification of the pathway.

Here we present a transcription involved in CXCL8, RPS us to better an approach with true po

Figure 1. Annotation diagram of the data that are compared in our analysis.

Results

We have run the workflow to

Figure 3. Enriched TRANSPATH Pathway. [Full classification →](#)

Identification of target

HumanPSD(TM) disease

In the first step of the analysis we applied the Limma tool (F expression) in the following: on the basis of 2 of the fold (corrected to the multiple test) ([Supplementary table 1](#)) a LogFC>0 for up-regulated a genes according to their p-values downregulated genes

Table 2. Identified upregulated genes. [See full table →](#)

ID	Gene symbol
ENSG00000162692	VCAM1
ENSG0000007908	SELE
ENSG00000090339	ICAM1
ENSG00000100906	NFKBIA
ENSG00000110848	CD69
ENSG00000104312	RIPK2

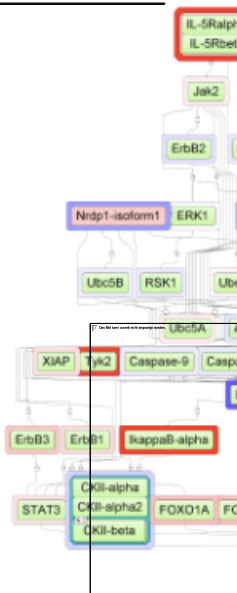
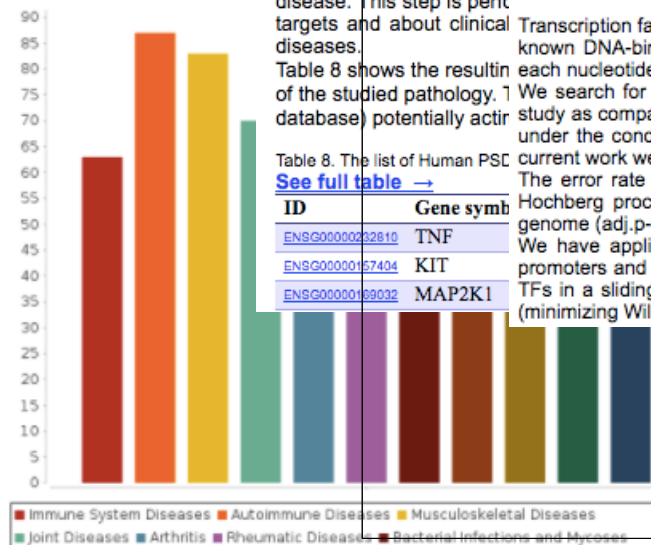


Figure 10. Diagram of intracellular signaling pathways. [See full diagram →](#)

Conclusion

At the last step of the analysis we identified the activation of the identified disease. This step is performed on the basis of clinical targets and about clinical diseases.

Table 8 shows the results of the studied pathology. 1 database) potentially active

Table 8. The list of Human PSC. [See full table →](#)

ID	Gene symbol
ENSG00000032810	TNF
ENSG00000057404	KIT
ENSG00000059032	MAP2K1

Methods

Methods for Analysis of Enriched Transcription Factor Binding Sites and composite modules

Transcription factor binding sites in promoters and enhancers of differentially expressed genes were analyzed using known DNA-binding motifs. The motifs are specified using position weight matrices (PWMs) that give weights to each nucleotide in each position of the DNA binding motif for a transcription factor or a group of them.

We search for transcription factor binding sites (TFBS) that are enriched in the promoters and enhancers under study as compared to a background sequence set such as promoters of genes that were not differentially regulated under the condition of the experiment. We denote study and background sets briefly as Yes and No sets. In the current work we used a workflow considering promoter sequences of a standard length of 1100 bp (-1000 to +100). The error rate in this part of the pipeline is controlled by estimating the adjusted p-value (using the Benjamini-Hochberg procedure) in comparison to the TFBS frequency found in randomly selected regions of the human genome (adj.p-value < 0.01).

We have applied the CMA algorithm (Composite Module Analyst) for searching composite modules [7] in the promoters and enhancers of the Yes and No sets. We searched for composite module consisting of a cluster of 10 TFs in a sliding window of 200-300 bp that statistically significantly separates sequences in the Yes and No sets (minimizing Wilcoxon p-value).

PC:17903182	1-[5-methyl-2-(trifluoromethyl)furan-3-yl]-3-[(2Z)-5-(2-[[6-(1H-1,2,4-triazol-3-ylamino)pyrimidin-4-...		NEK7, MAPK6, TRAF6, CDK1, CCNA1, BARD1, RPS6KA5...	4.64	0.14
PC:17902747	N-(CYCLOPROPYLMETHYL)-4-(METHYLOXY)-3-({5-[3-(3-PYRIDINYLPHENYL)-1,3-OXAZOL-2-YL]AMINO)BENZENESULFO...		NEK7, MAPK6, TRAF6, CDK1, CCNA1, DUSP4, BARD1...	4.19	0.16

Robot-scientist

Is this our future?

