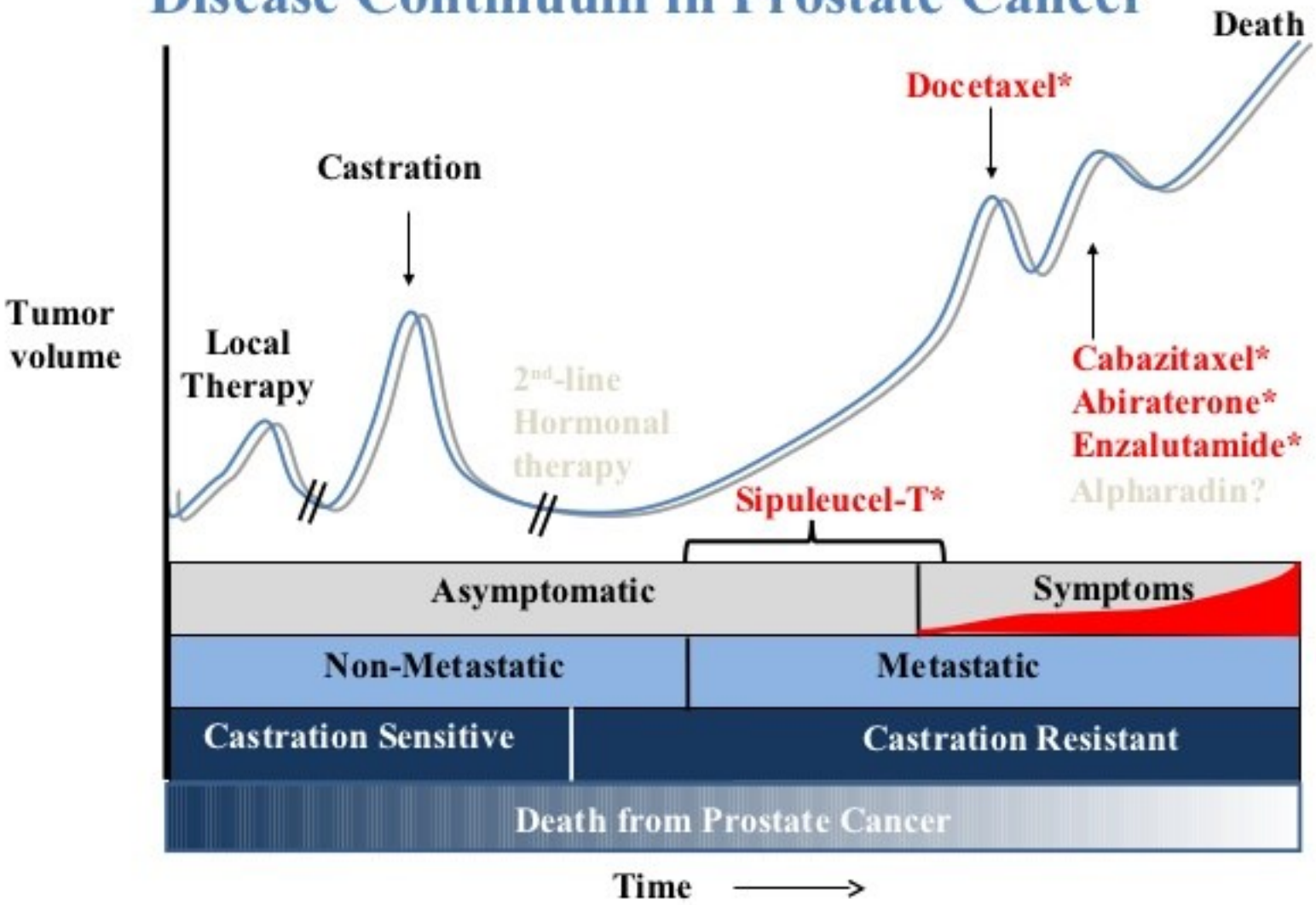
TRANSFAC

CANCER

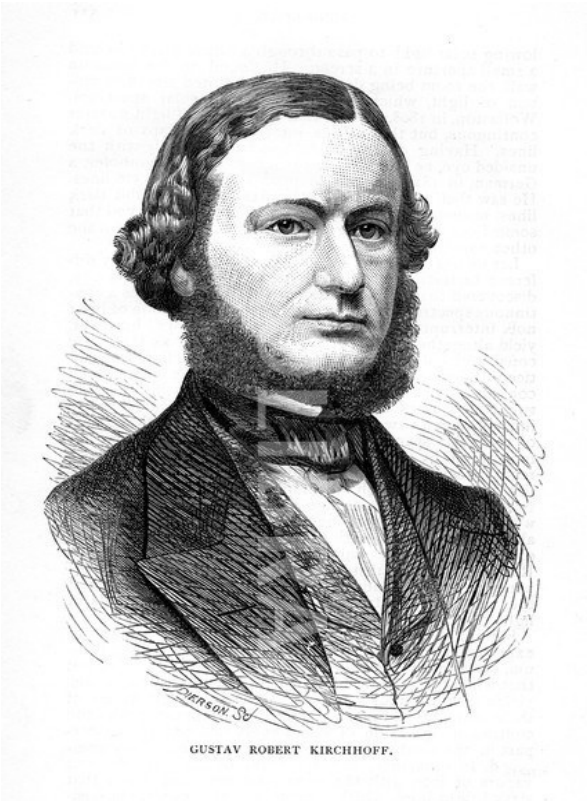
versus

Disease Continuum in Prostate Cancer



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

“A good theory is the most practical thing”

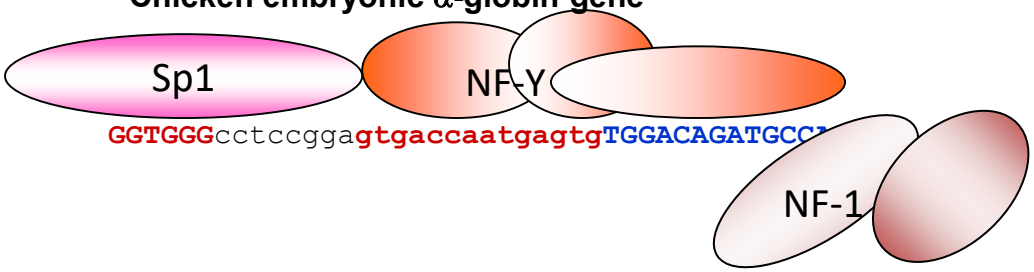


Gustav Robert Kirchhoff

(1824 - 1887), German physicists

Antagonistic composite elements

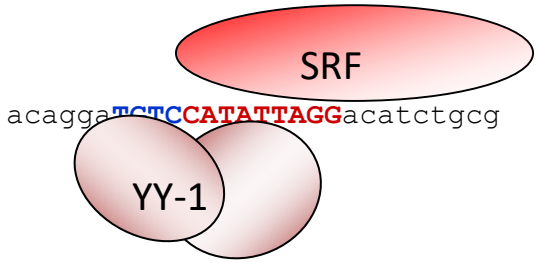
COMPEL: C00006
Chicken embryonic α-globin gene



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

NF-1 represses transcription
in adult cells

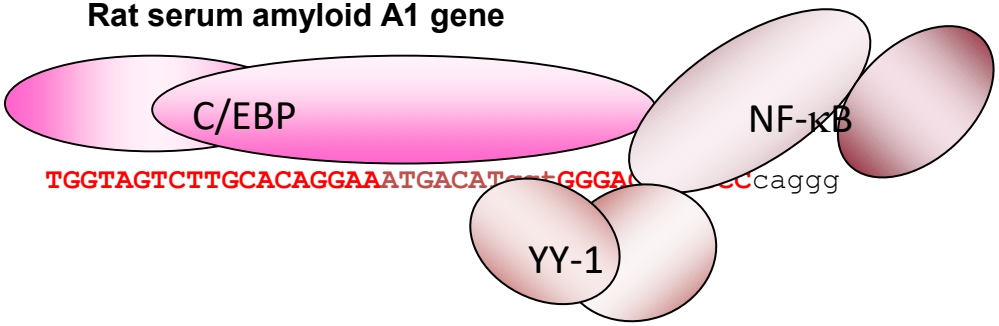
COMPEL: C00009
Human *c-fos* protooncogene



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*.

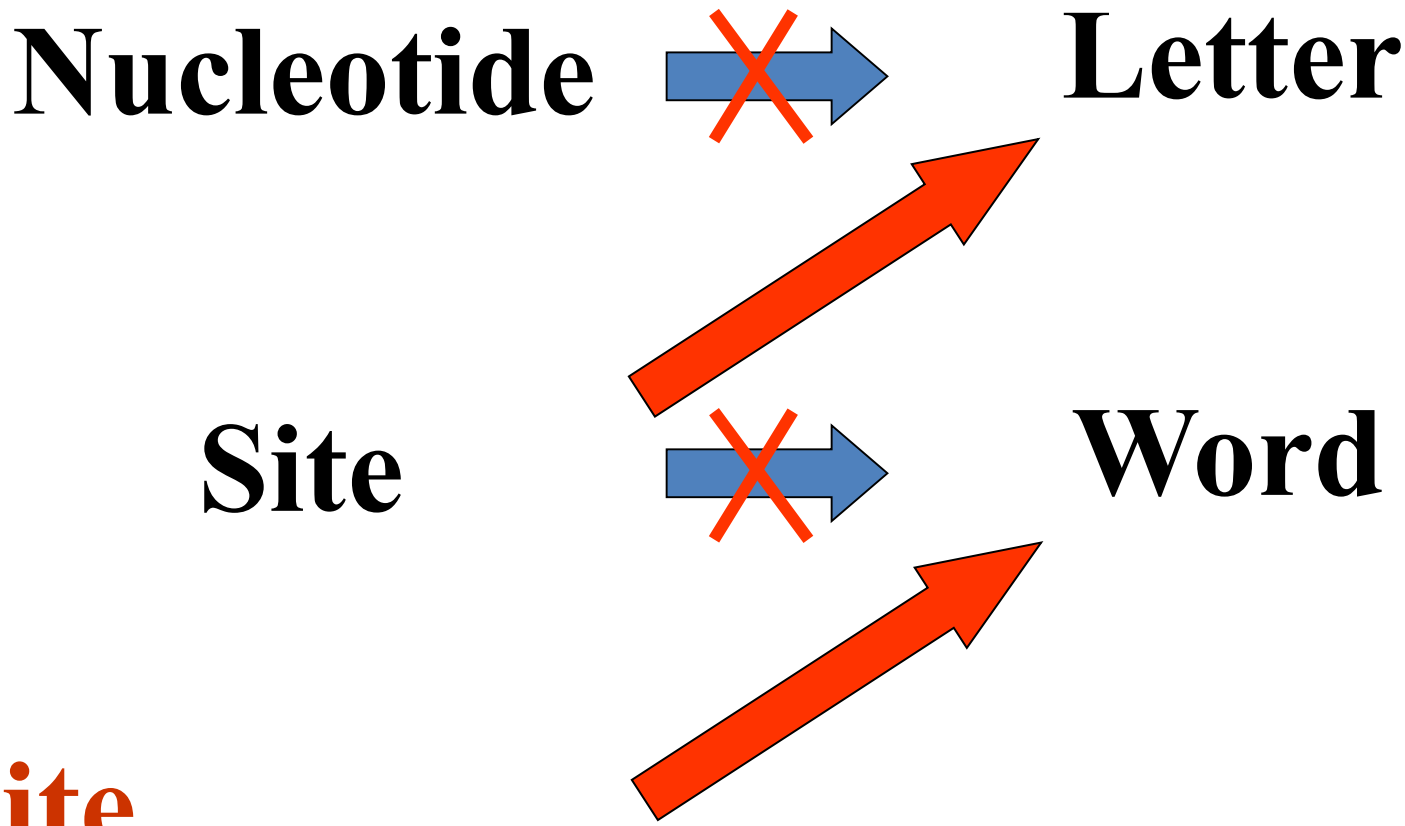
COMPEL: C00054
Rat serum amyloid A1 gene



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

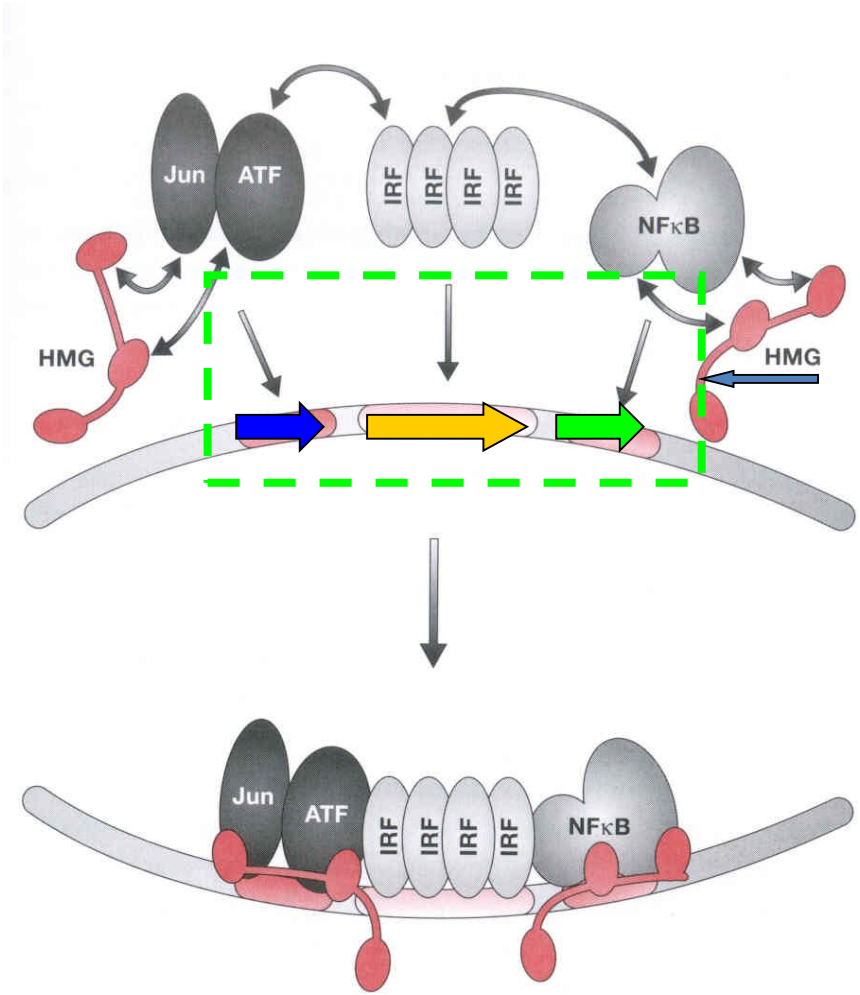
YY1 represses inducible transcription of this gene.

Paradigm shift



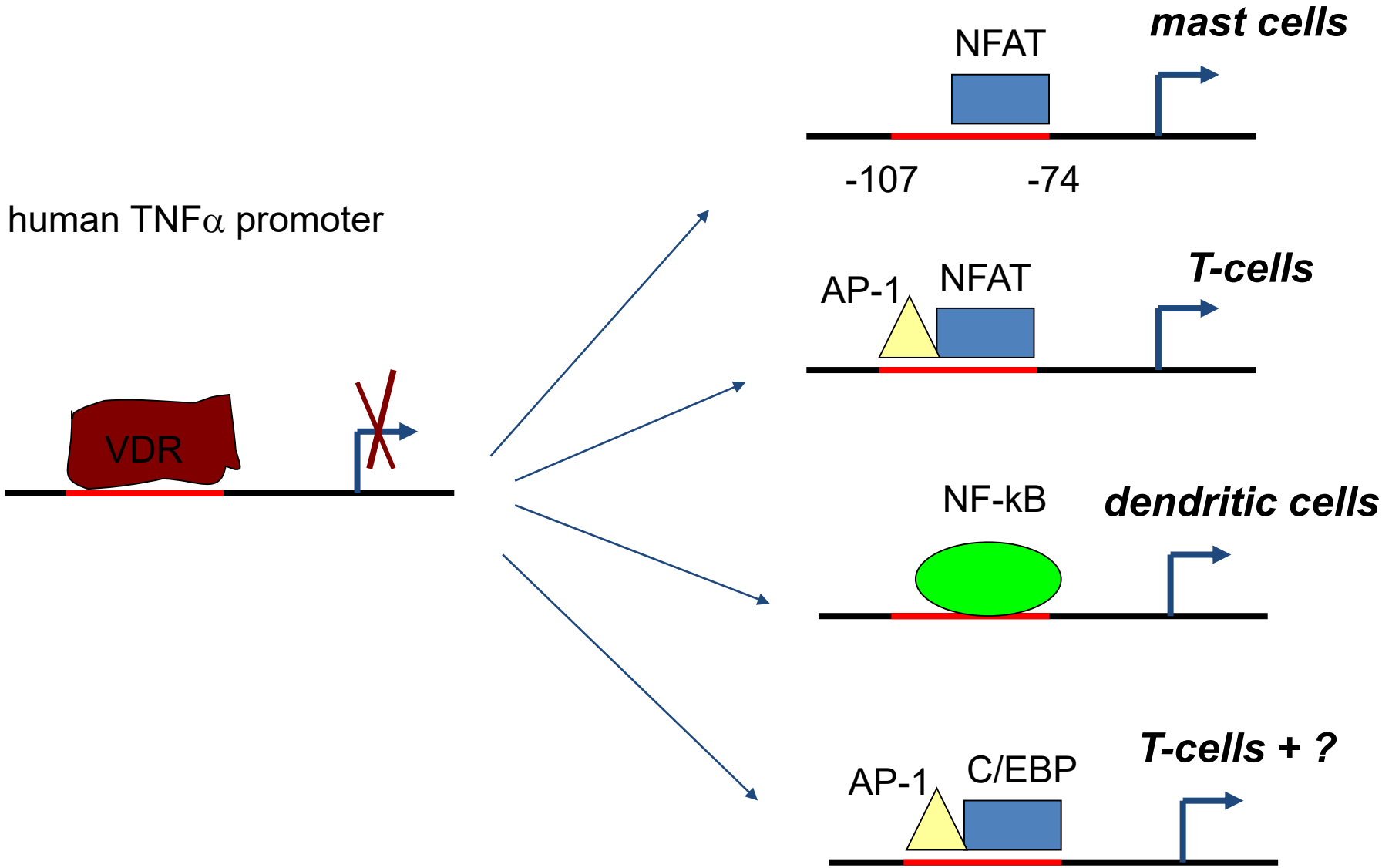
Site Composition

Composite Modules (CM)



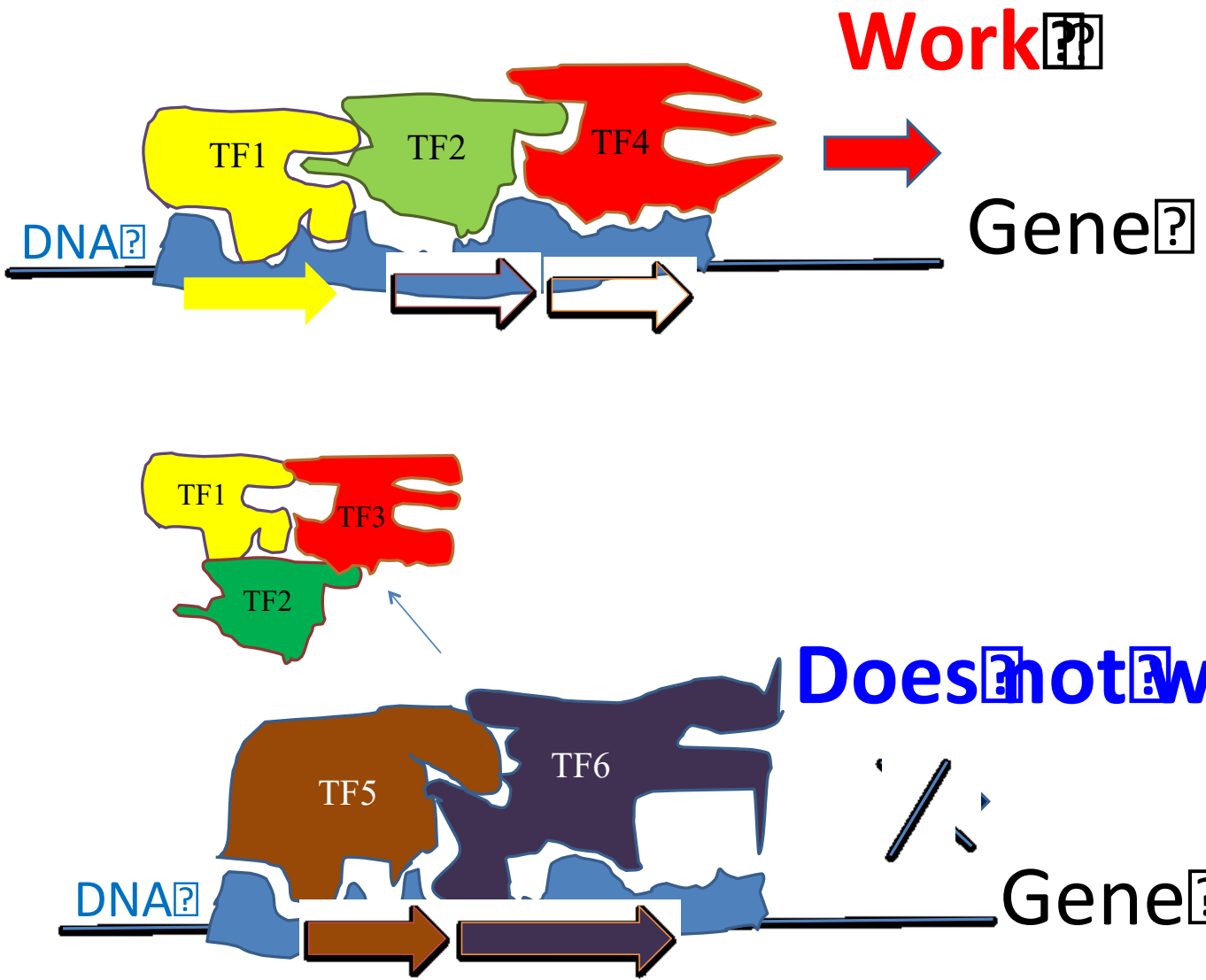
Composite Module (CM)

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 TNFα promoter
VDR
NFAT
mast cells
-107
-74
AP-1
NFAT
T-cells
NF-kB
dendritic cells
AP-1
C/EBP
T-cells + ?

Exact fit

It's Fuzzy Puzzle!!

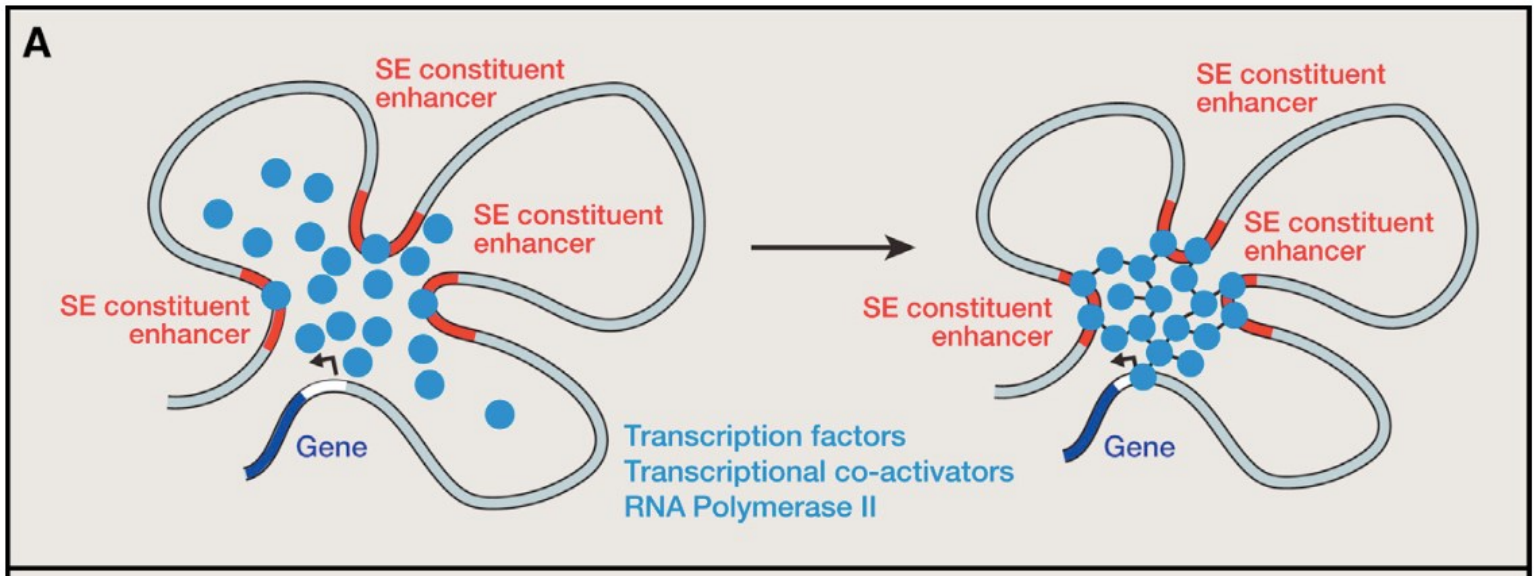


Leading Edge
Perspective

Cell

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,*]
[1]Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA
[2]Department of Chemical Engineering
[3]Department of Biology
[4]Department of Physics



AP-1

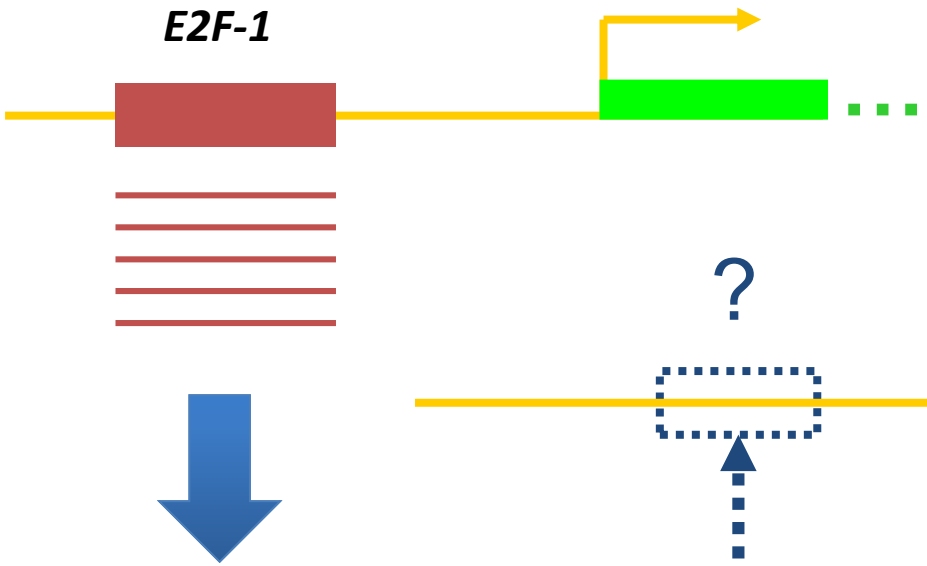
Consensus: TGAgTCA

Human collagenase (-2013) TGAGTCA

Mouse IL-2 (-143) TGTGTAA

Mouse IL-2 (-82) TGTAATA

Search for new TF binding sites with PWMs



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

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

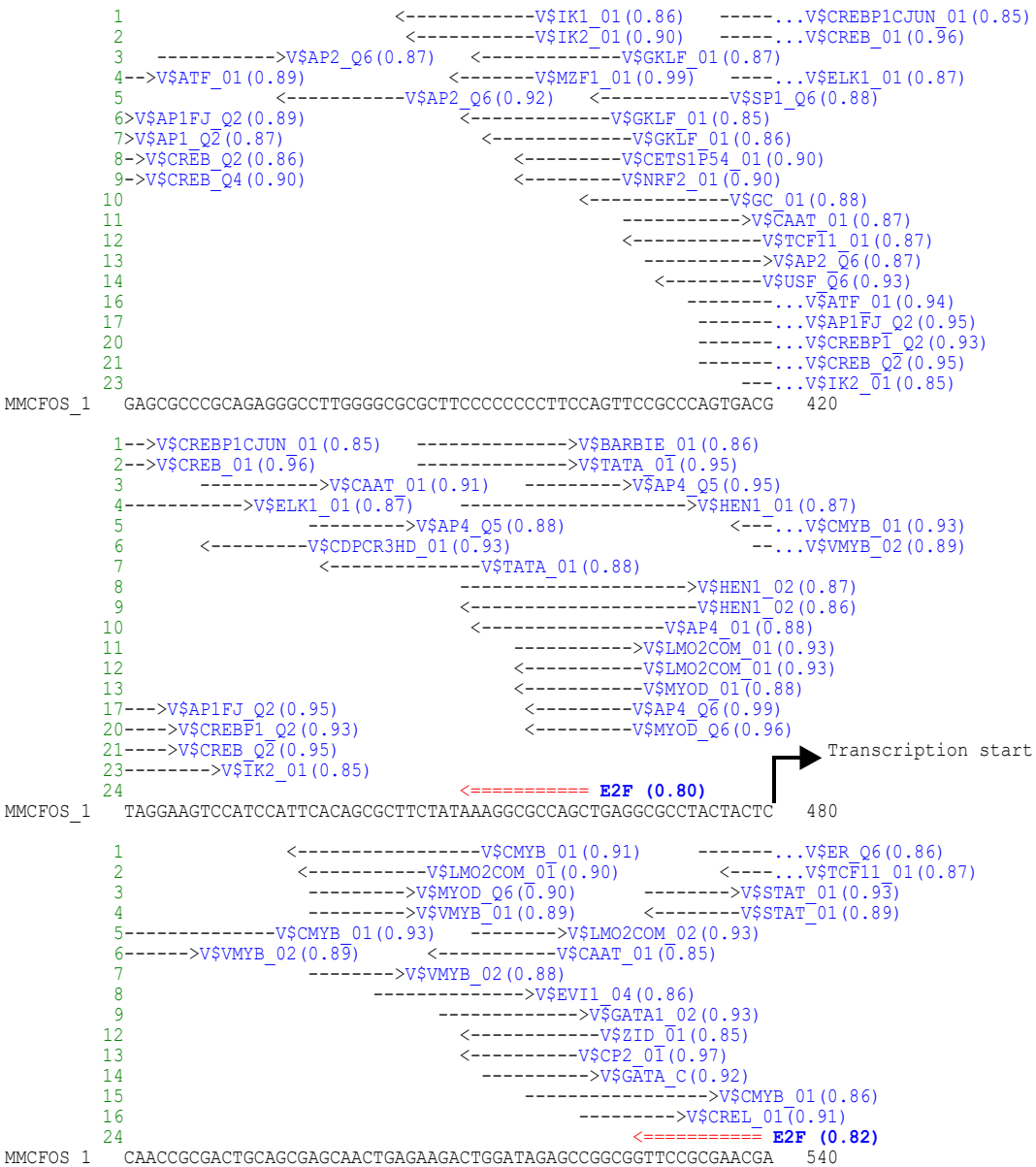
Matrix length: 8

Logo

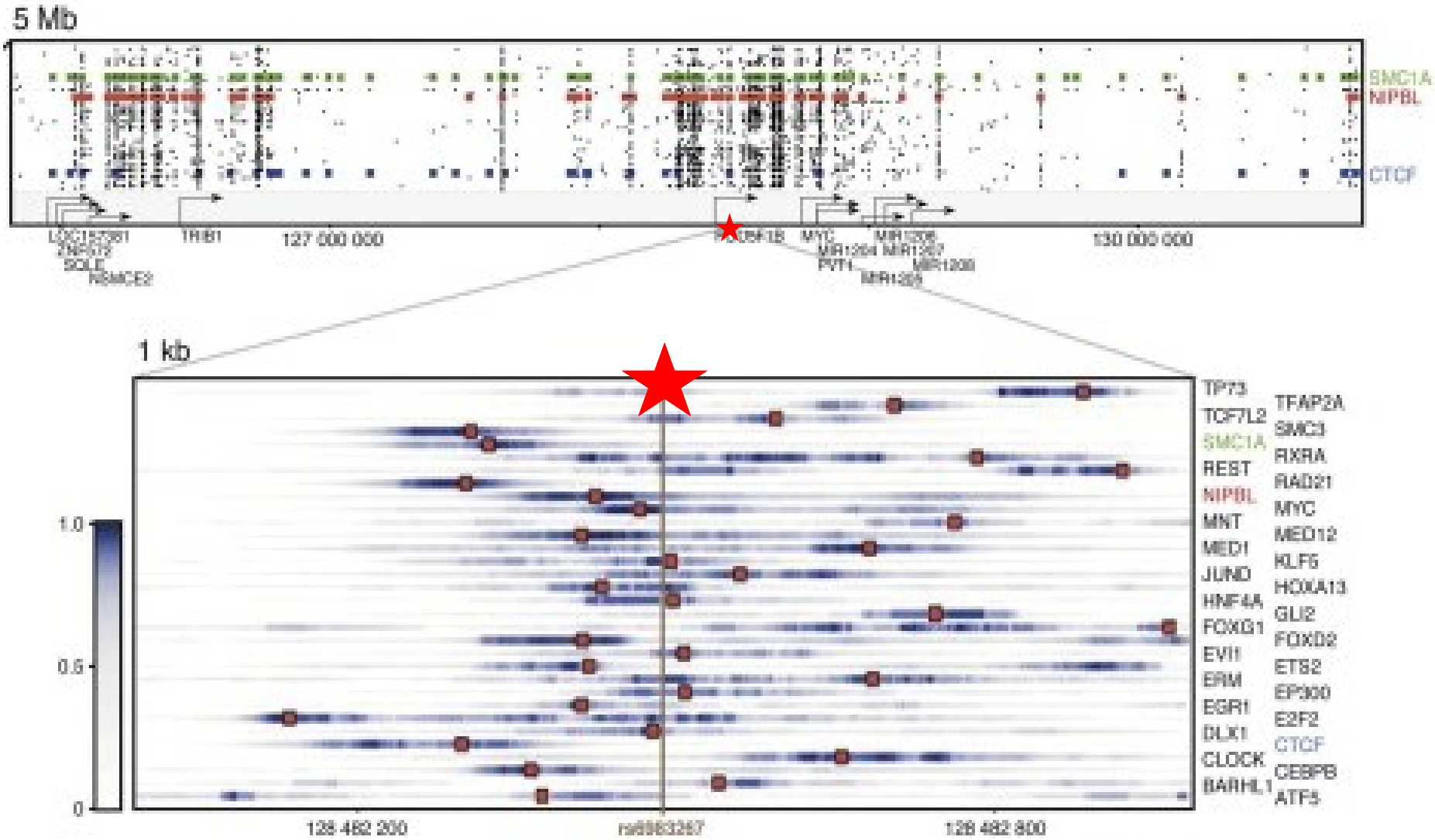
A	0.212	0.019	0.000	0.000	0.000	0.000	0.000	0.000
C	0.000	0.000	0.000	0.500	0.115	1.000	0.000	0.885
G	0.000	0.000	0.000	0.500	0.885	0.000	1.000	0.115
T	0.788	0.981	1.000	0.000	0.000	0.000	0.000	0.000

TRANSFAC®

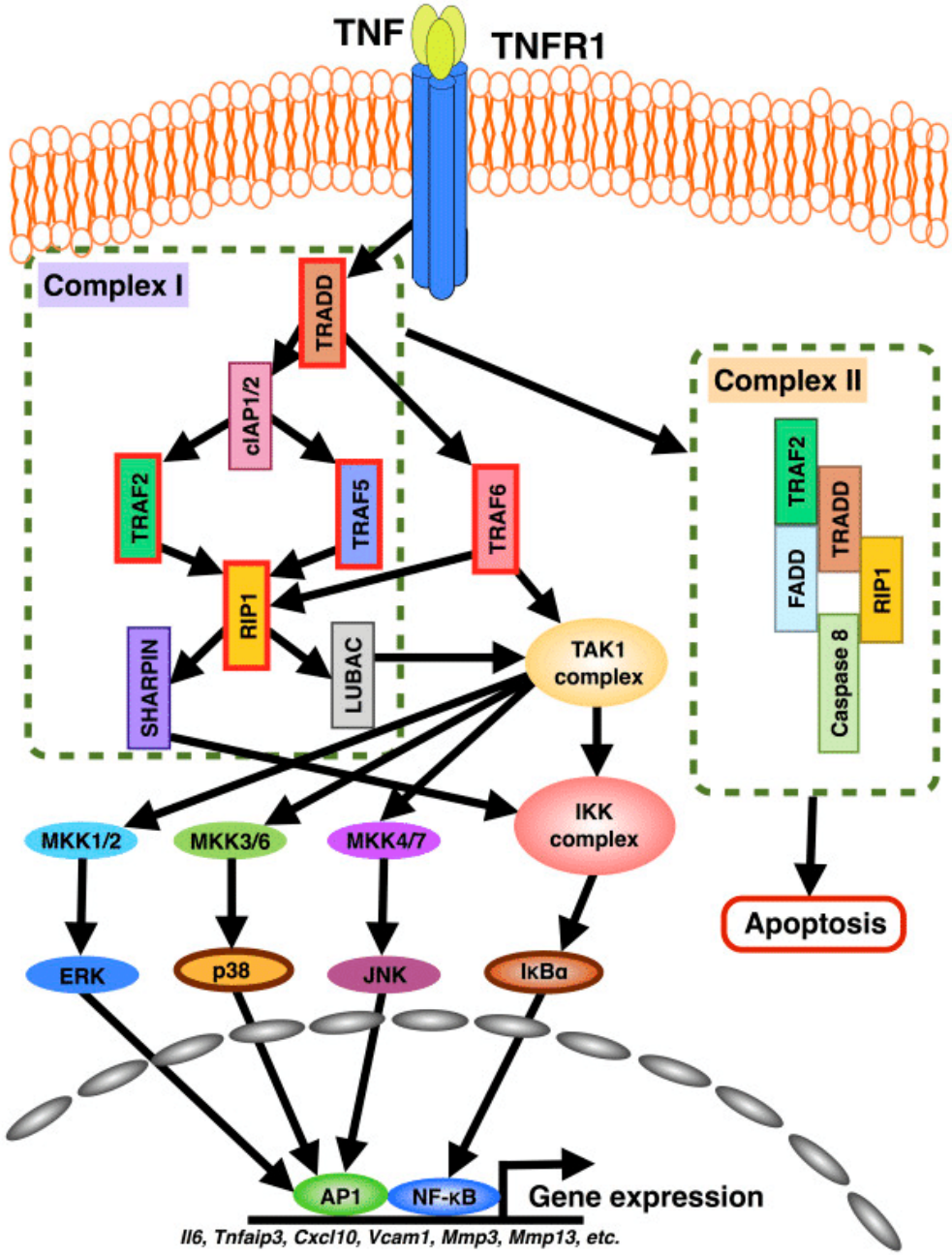
Mouse c-fos promoter
(Matrix search for TF binding sites)



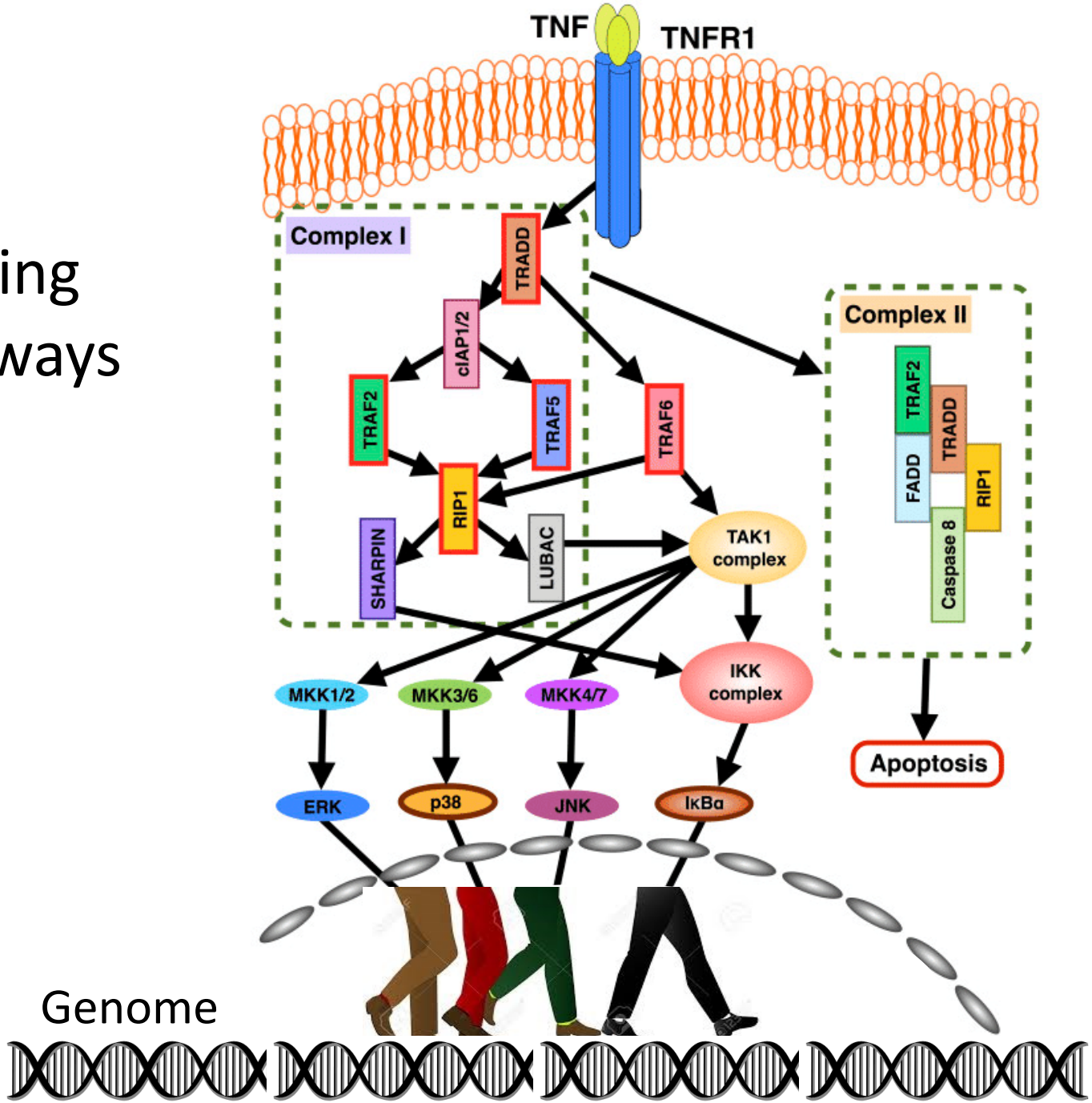
Colorectal cancer: tumor-specific enhancer around a SNP in regulatory region of MYC gene



Walking pathways



Walking pathways



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

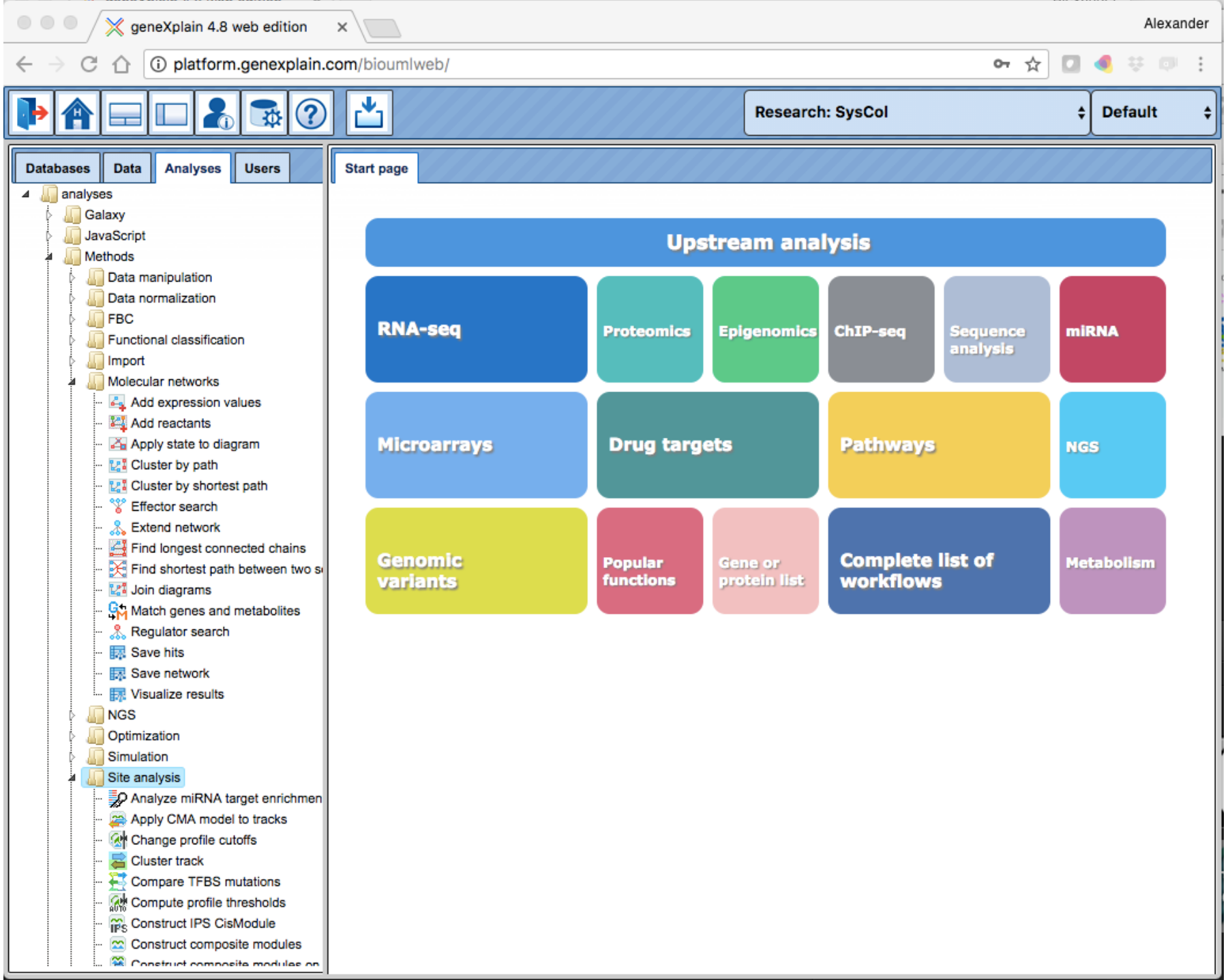
TRADAT

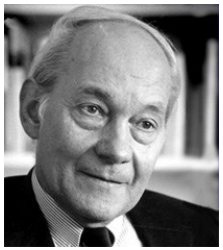
1995 - 1998

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.

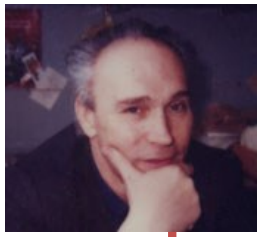
The concept of using weight matrices for the detection of transcription factor binding sites was extended and incorporated into higher order software tools. TRADAT contributed to the further development of MatInspector, ModelGenerator, FastM, and ModelInspector.

Multi-omics


geneXplain 4.8 web edition
Alexander
platform.genexplain.com/bioumlweb/
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
Upstream analysis
RNA-seq
Proteomics
Epigenomics
ChIP-seq
Sequence analysis
miRNA
Microarrays
Drug targets
Pathways
NGS
Genomic variants
Popular functions
Gene or protein list
Complete list of workflows
Metabolism



Manfred
Eigen



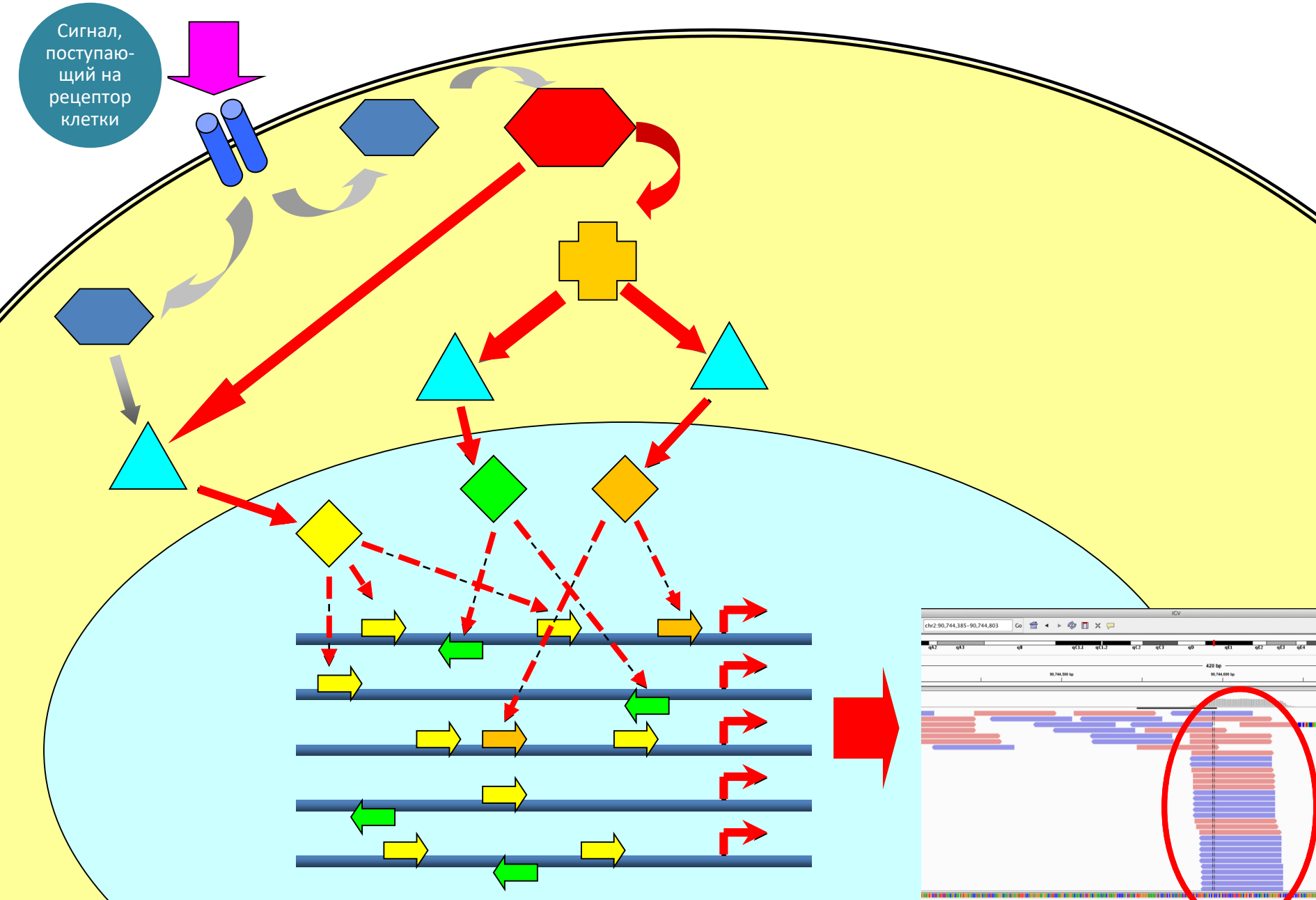
Vadim
Ratner

(?)

Nikolay
Kolchanov

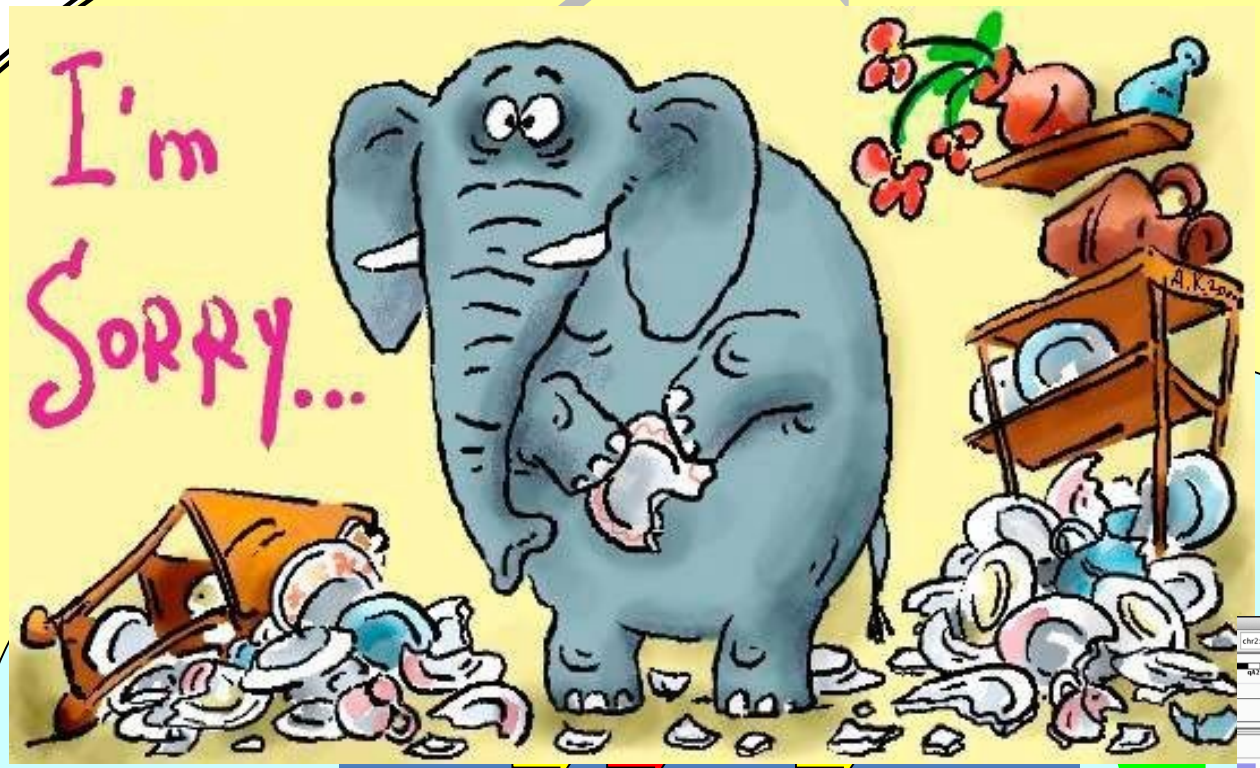
The
Edgar

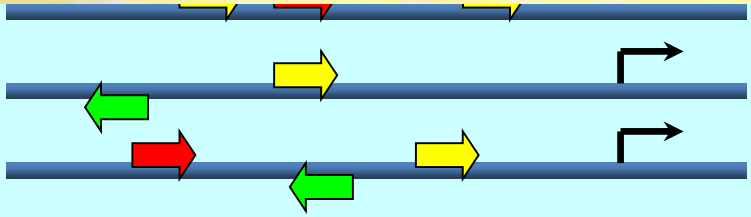
How to search for the causal mechanism of cancer?



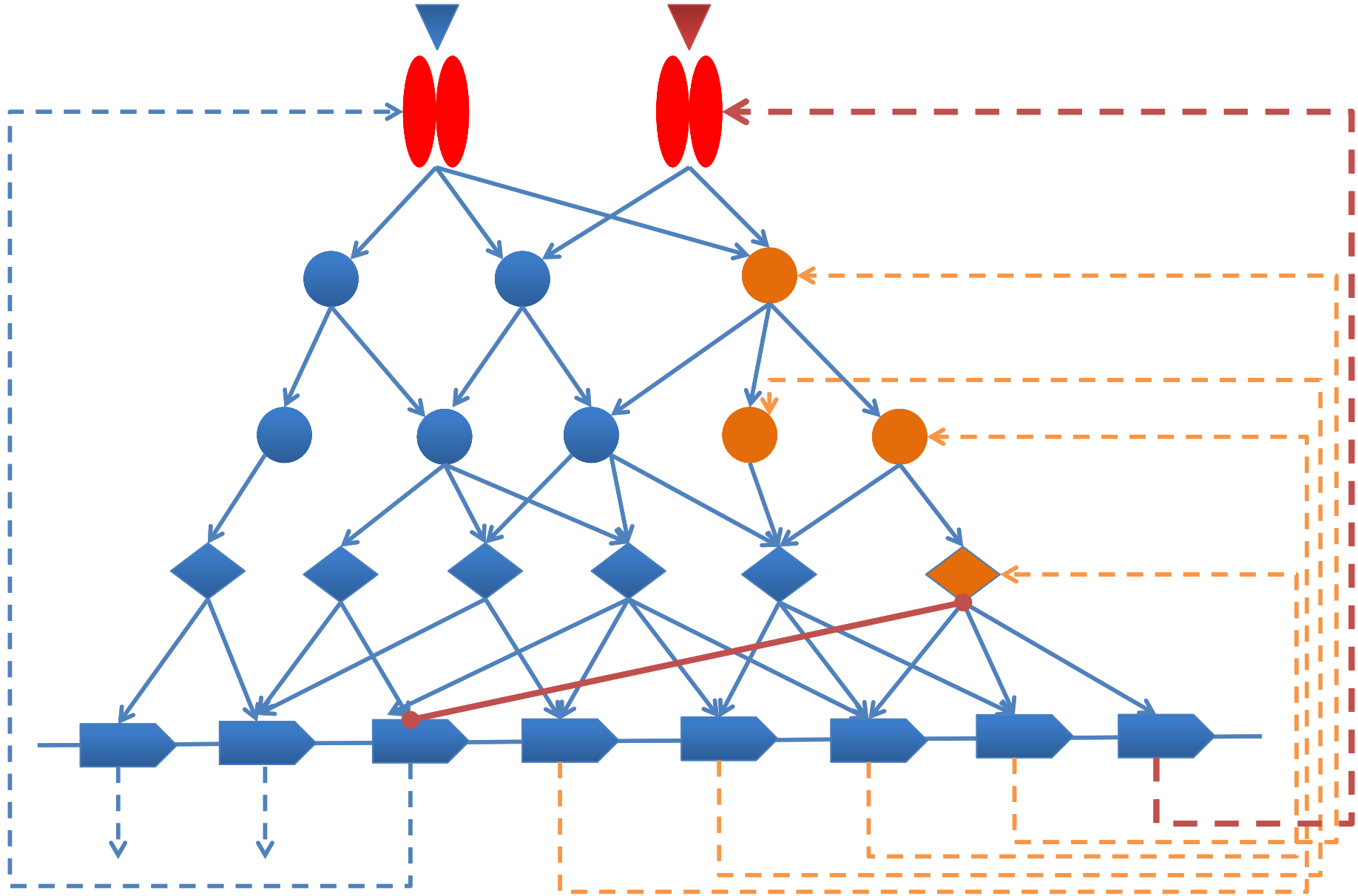
Master-regulator causes a lot of noise in the cell

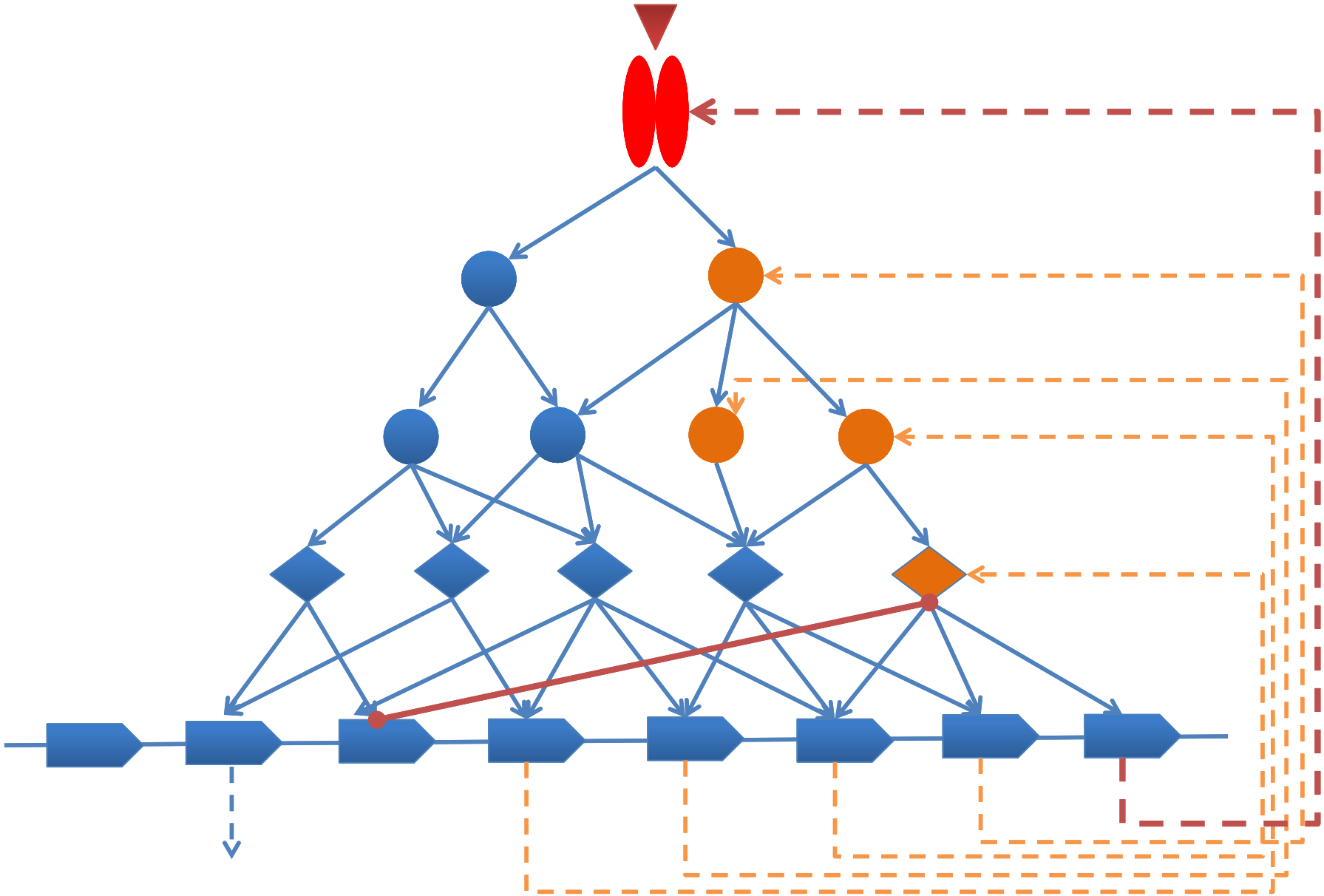
Master regulator



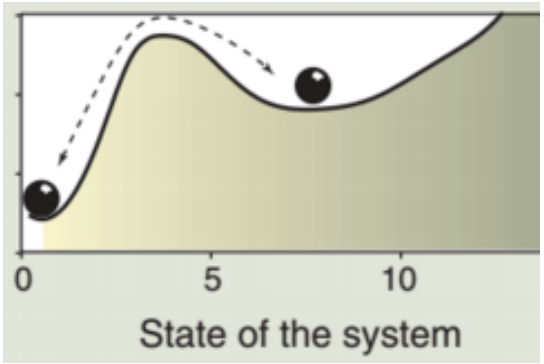


Walking pathways



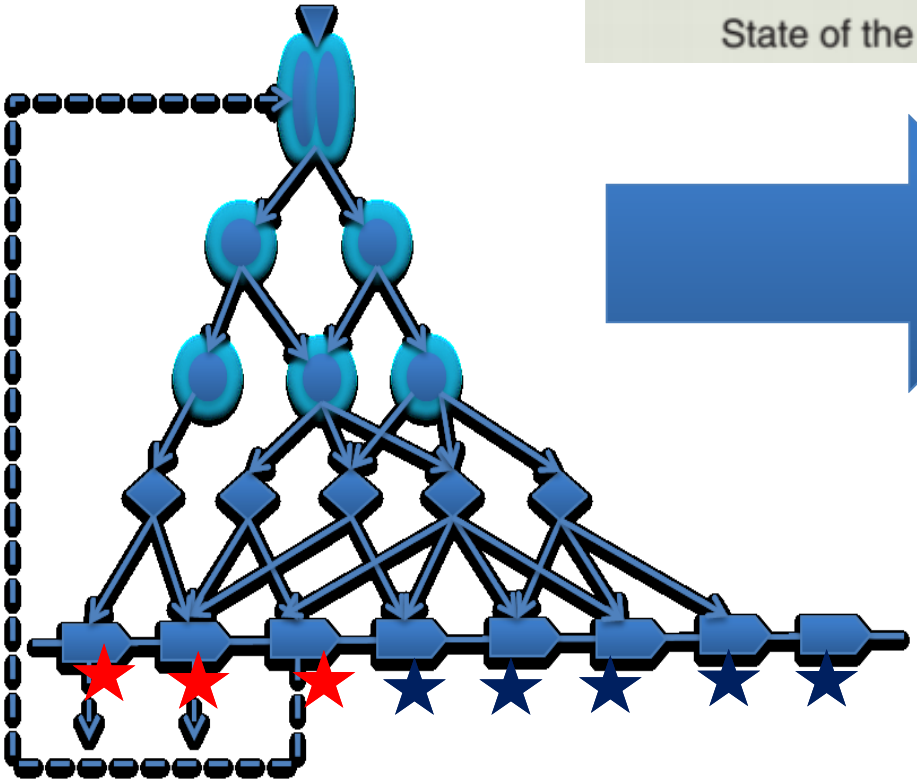


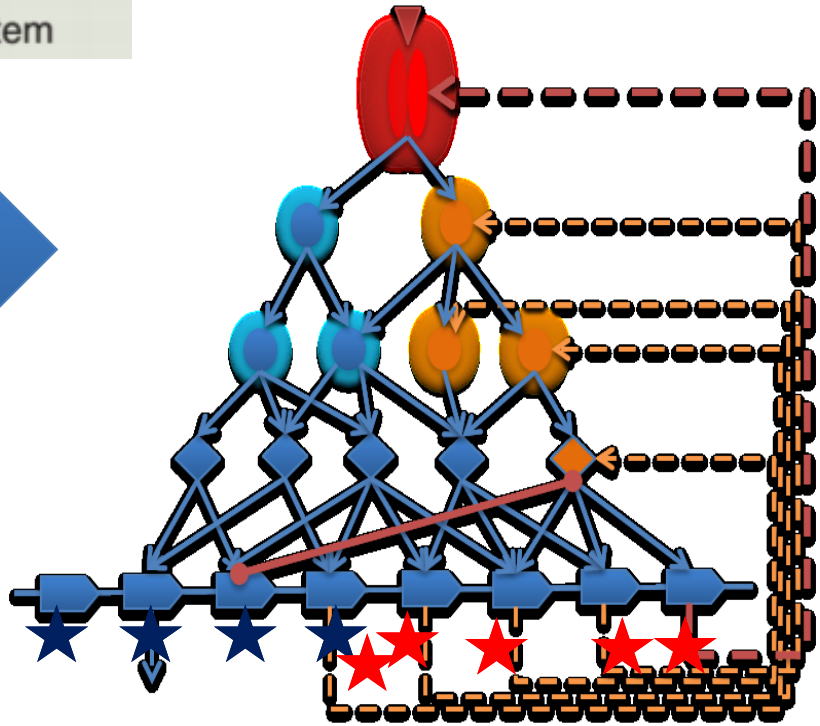
Pathway corruption



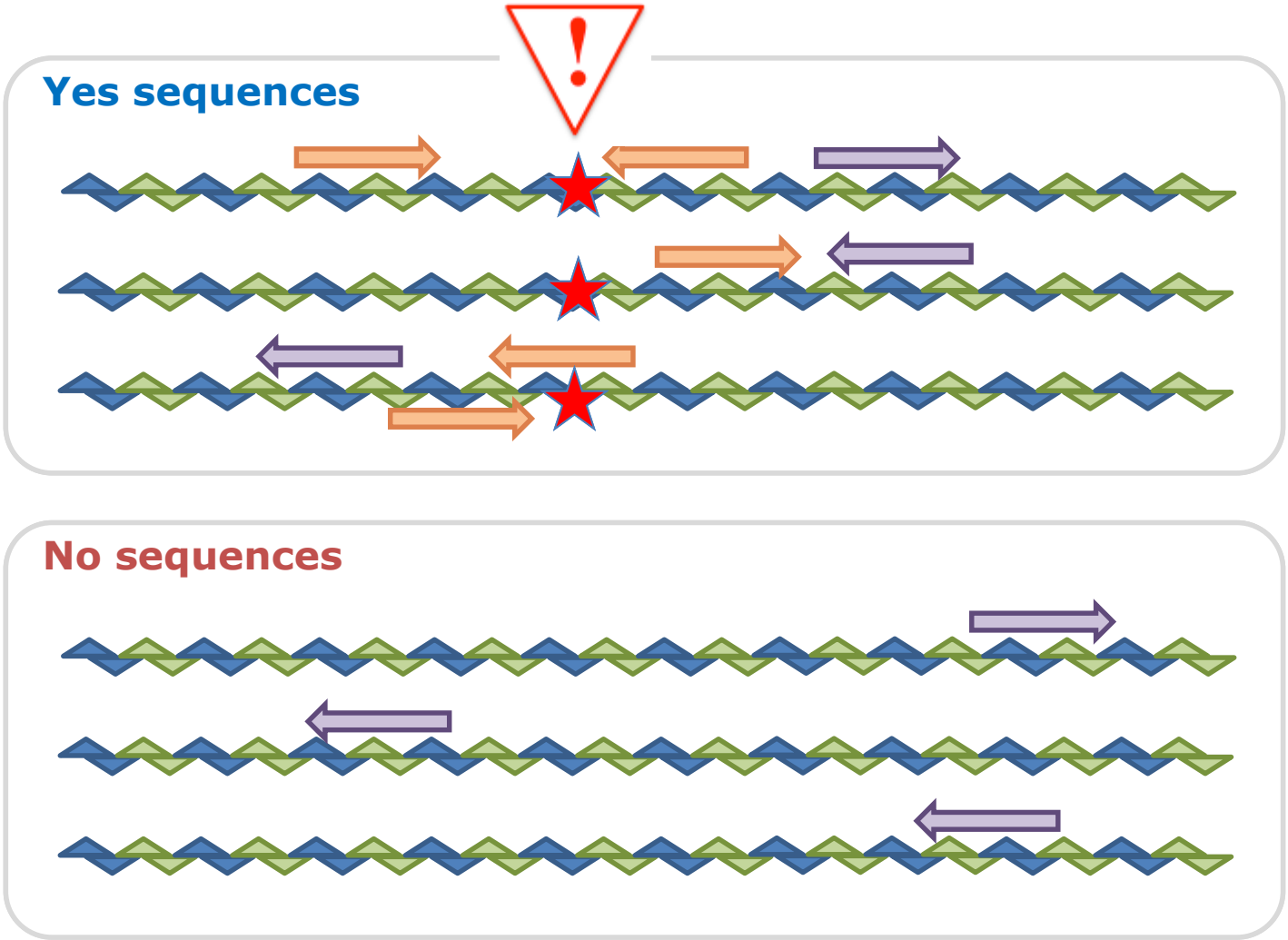
Healthy

Cancer

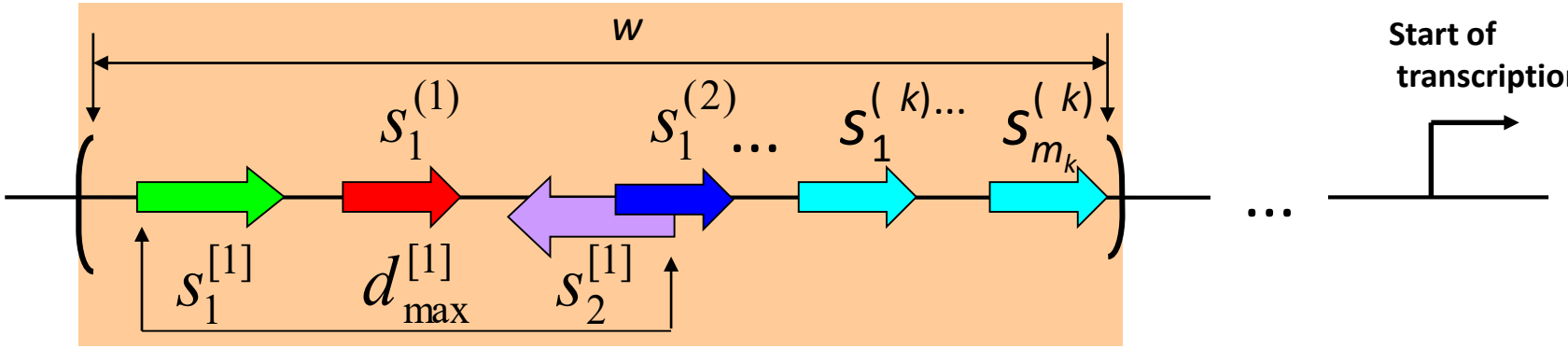




Motif enrichment analysis around regulatory mutations

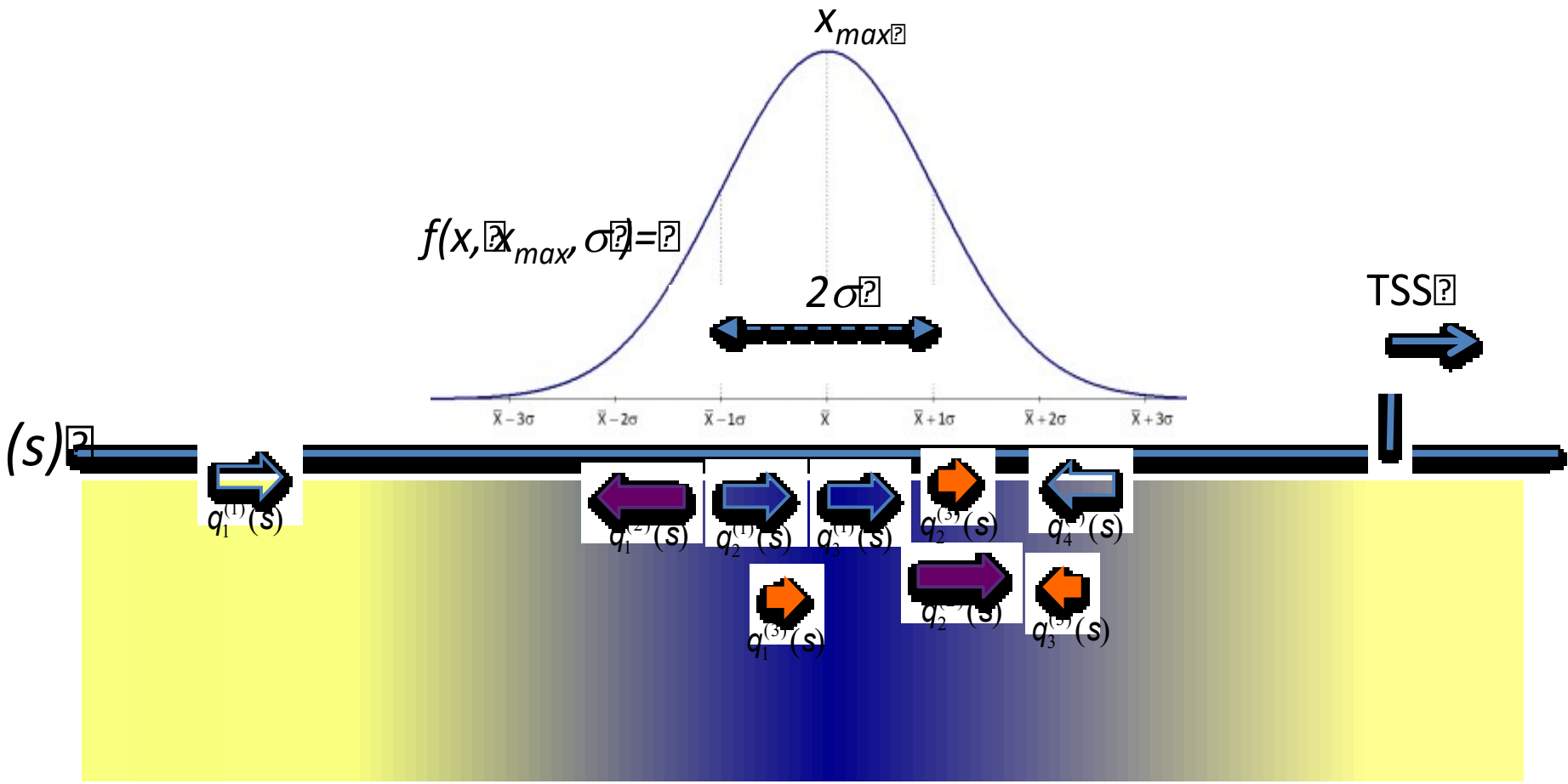


We construct a model of the disease-specific enhancer as a combination if transcription factor binding sites.

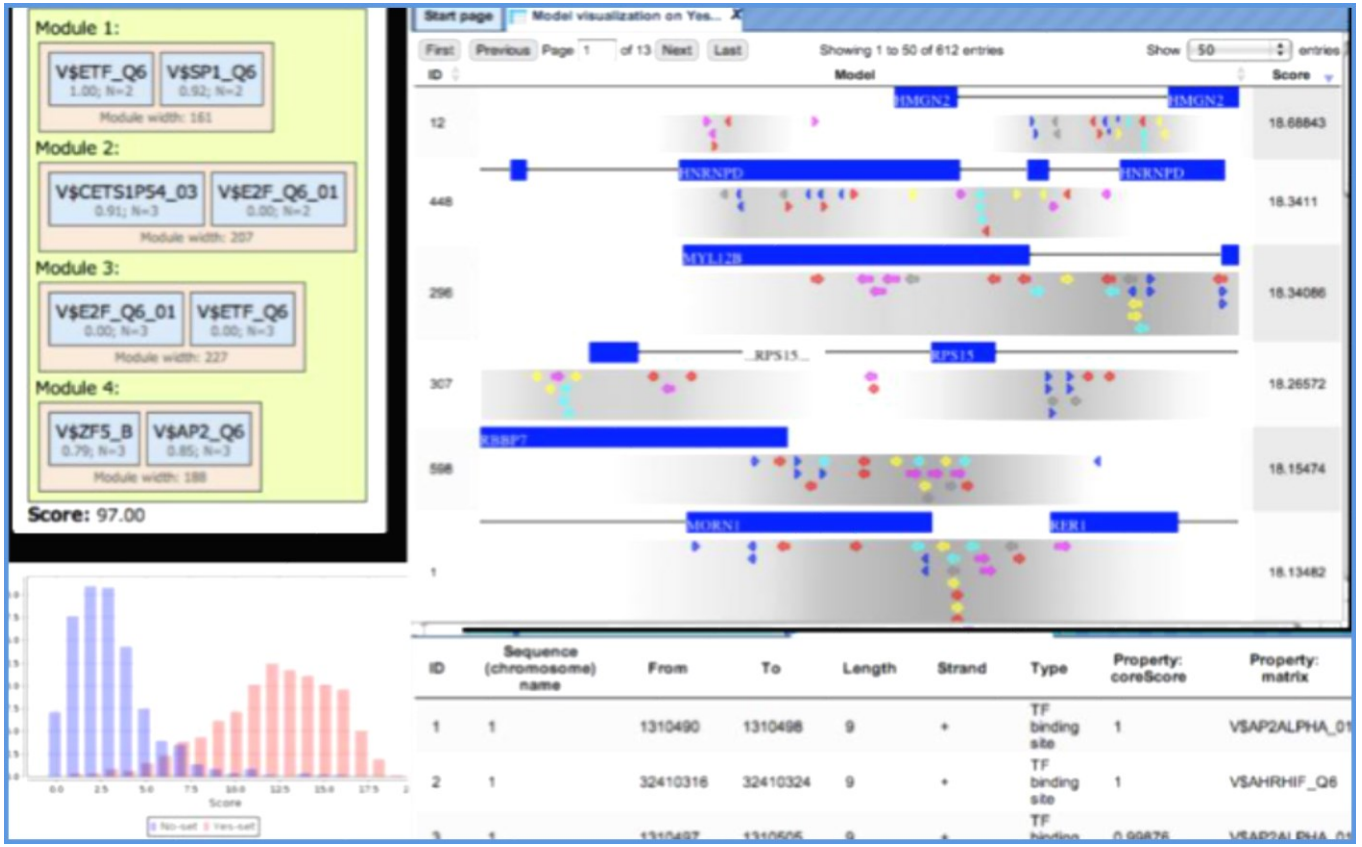


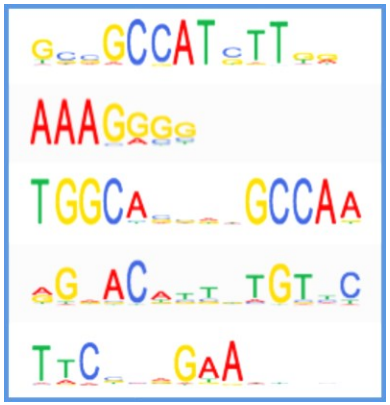
$$\begin{matrix} d_{\max}^{[1]} & d_{\max}^{[1]} & \dots & d_{\max}^{[R]} \\ q_{cut-off}^{(1)} & q_{cut-off}^{(2)} & \dots & q_{cut-off}^{(k)} \\ \phi^{(1)} & \phi^{(2)} & \dots & \phi^{(k)} \end{matrix}$$

The model parameters are found with the help of ***genetic algorithm***


x_{max}
$f(x, x_{max}, \sigma) =$
2σ
TSS
$\bar{X}-3\sigma$
$\bar{X}-2\sigma$
$\bar{X}-1\sigma$
$\bar{X}$
$\bar{X}+1\sigma$
$\bar{X}+2\sigma$
$\bar{X}+3\sigma$
$q_1^{(1)}(s)$
$q_1^{(2)}(s)$
$q_2^{(1)}(s)$
$q_3^{(1)}(s)$
$q_2^{(3)}(s)$
$q_4^{(2)}(s)$
$q_1^{(3)}(s)$
$q_2^{(2)}(s)$
$q_3^{(2)}(s)$

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

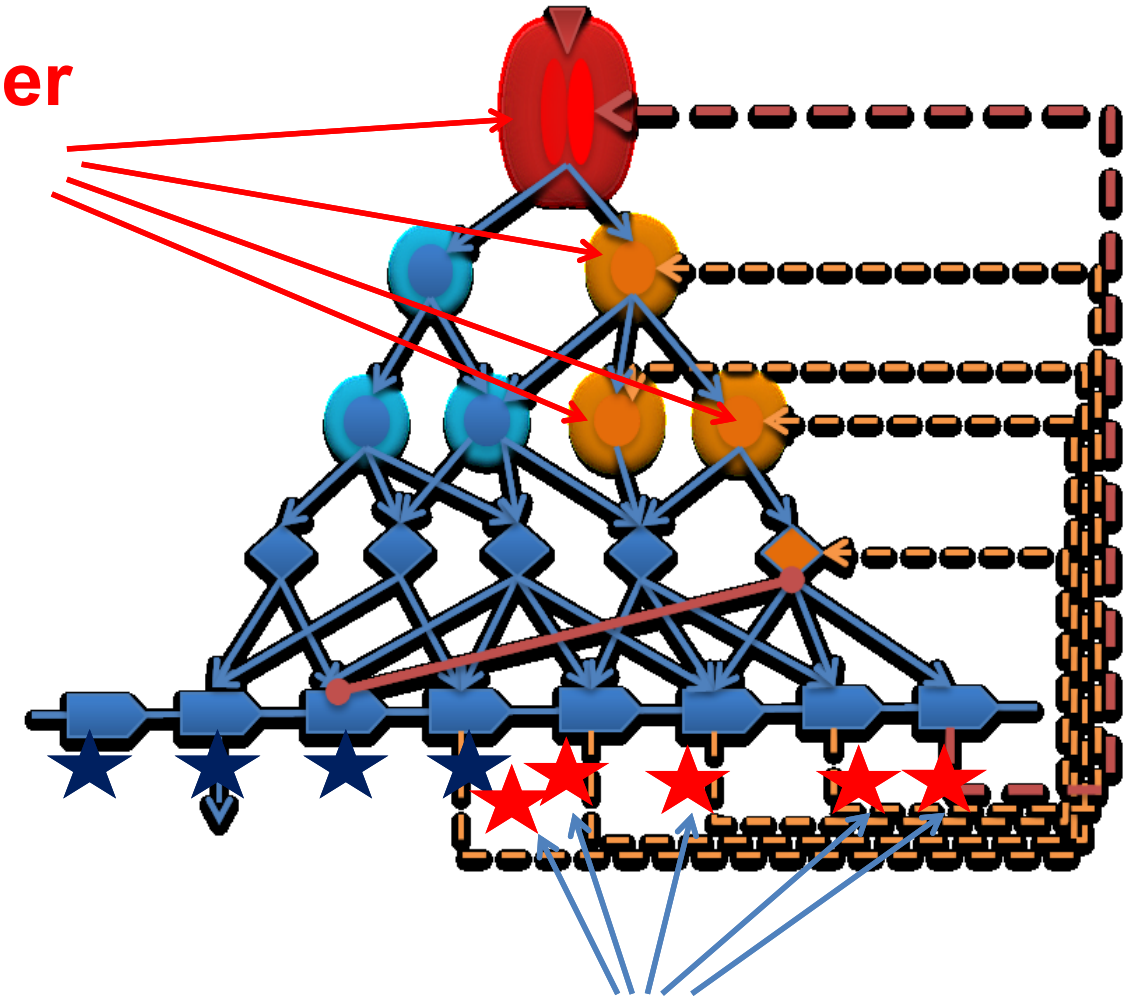


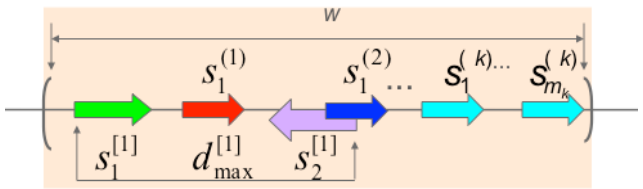


Binding sites for transcription factors that were included into the enhancer model

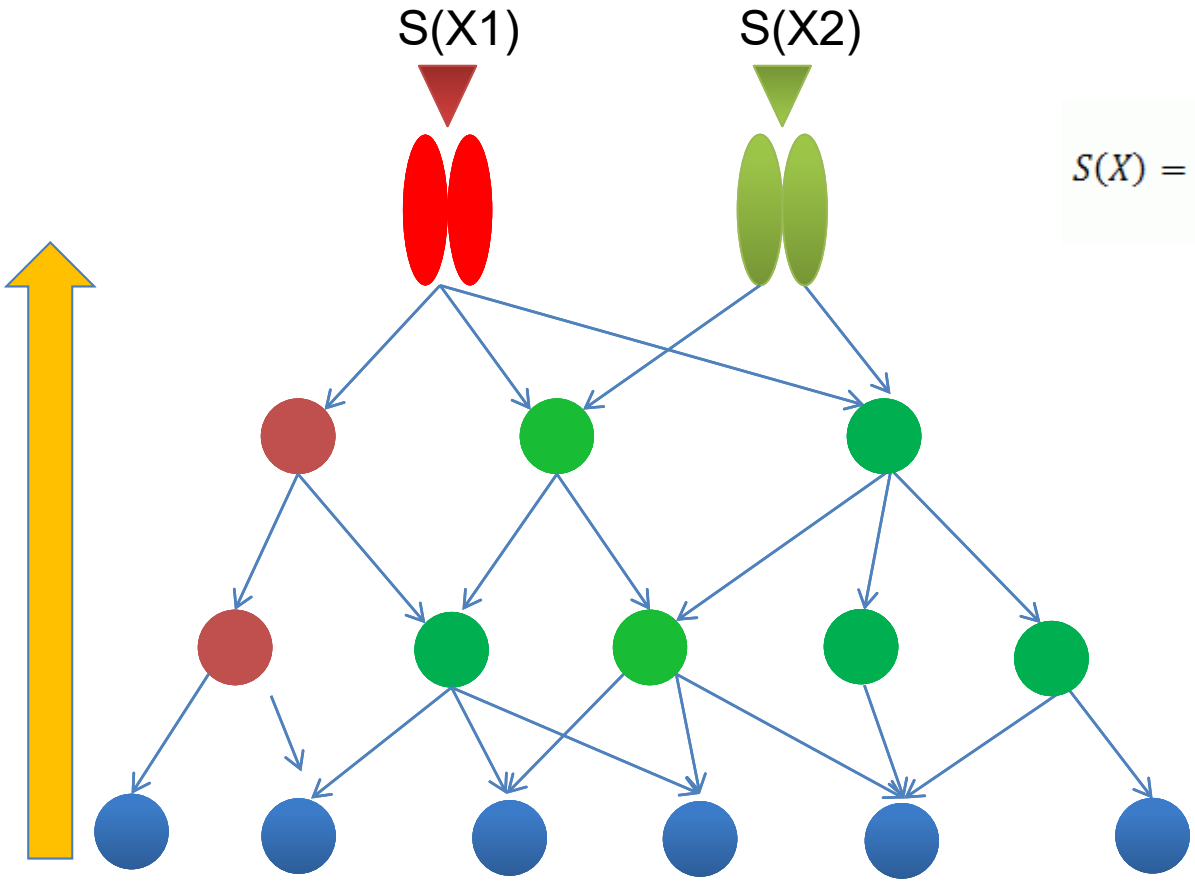
Cancer

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.

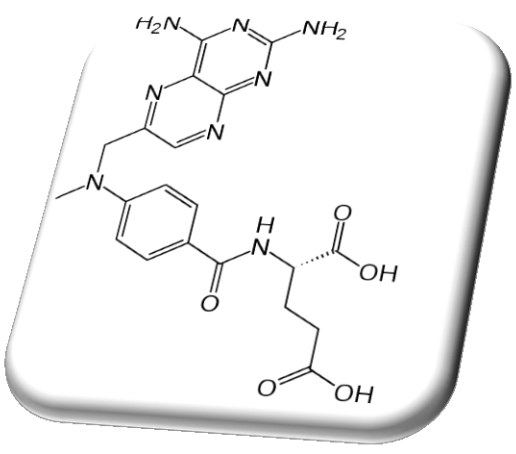
Kel, A., Voss, N., Jauregui, R., Kel-Margoulis, O. and Wingender, E.: Beyond microarrays: Find key transcription factors controlling signal transduction pathways BMC Bioinformatics 7(Suppl. 2), S13 (2006).


RESISTANCE

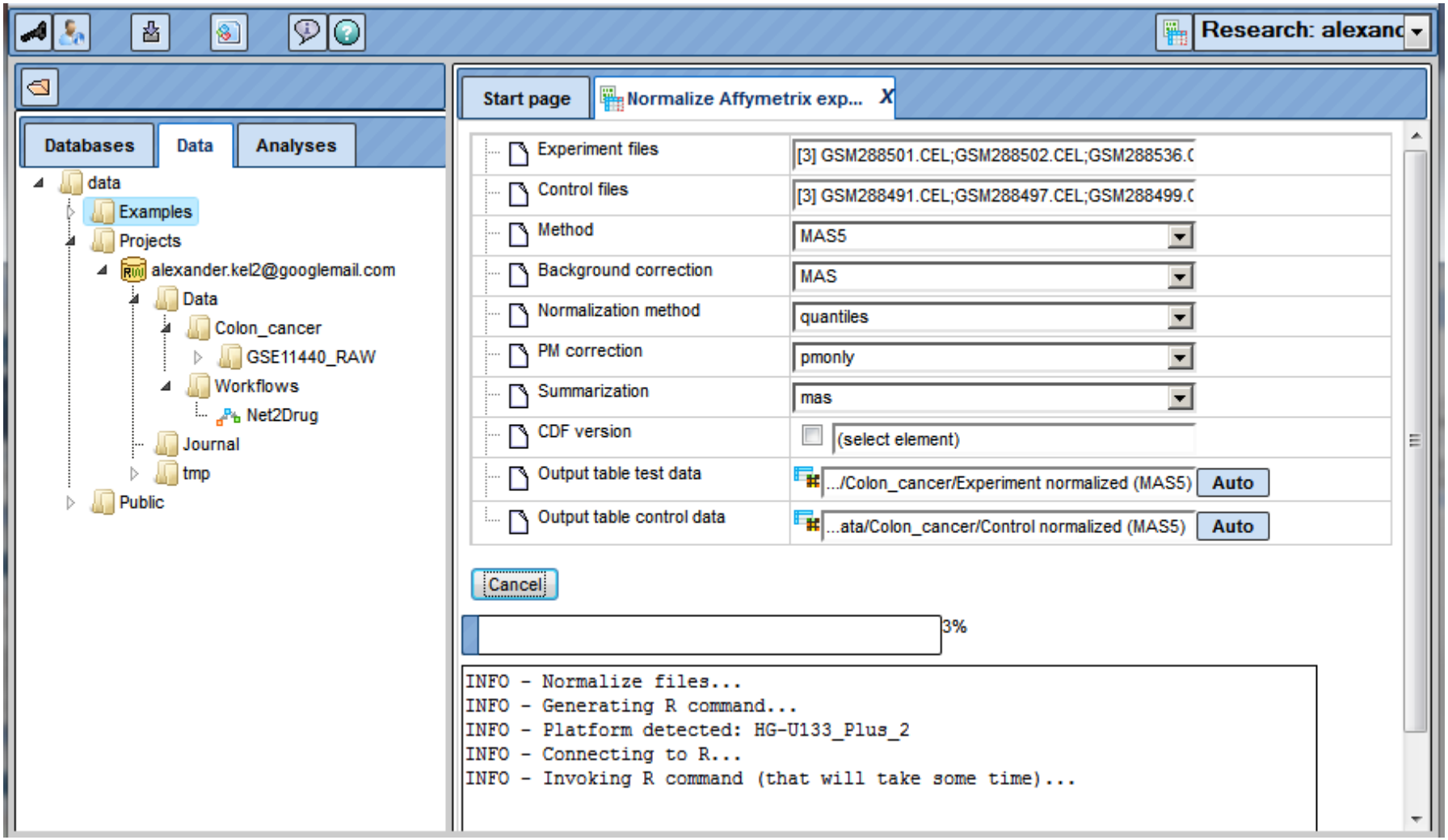
Series GSE11440

Query DataSets for GSE11440

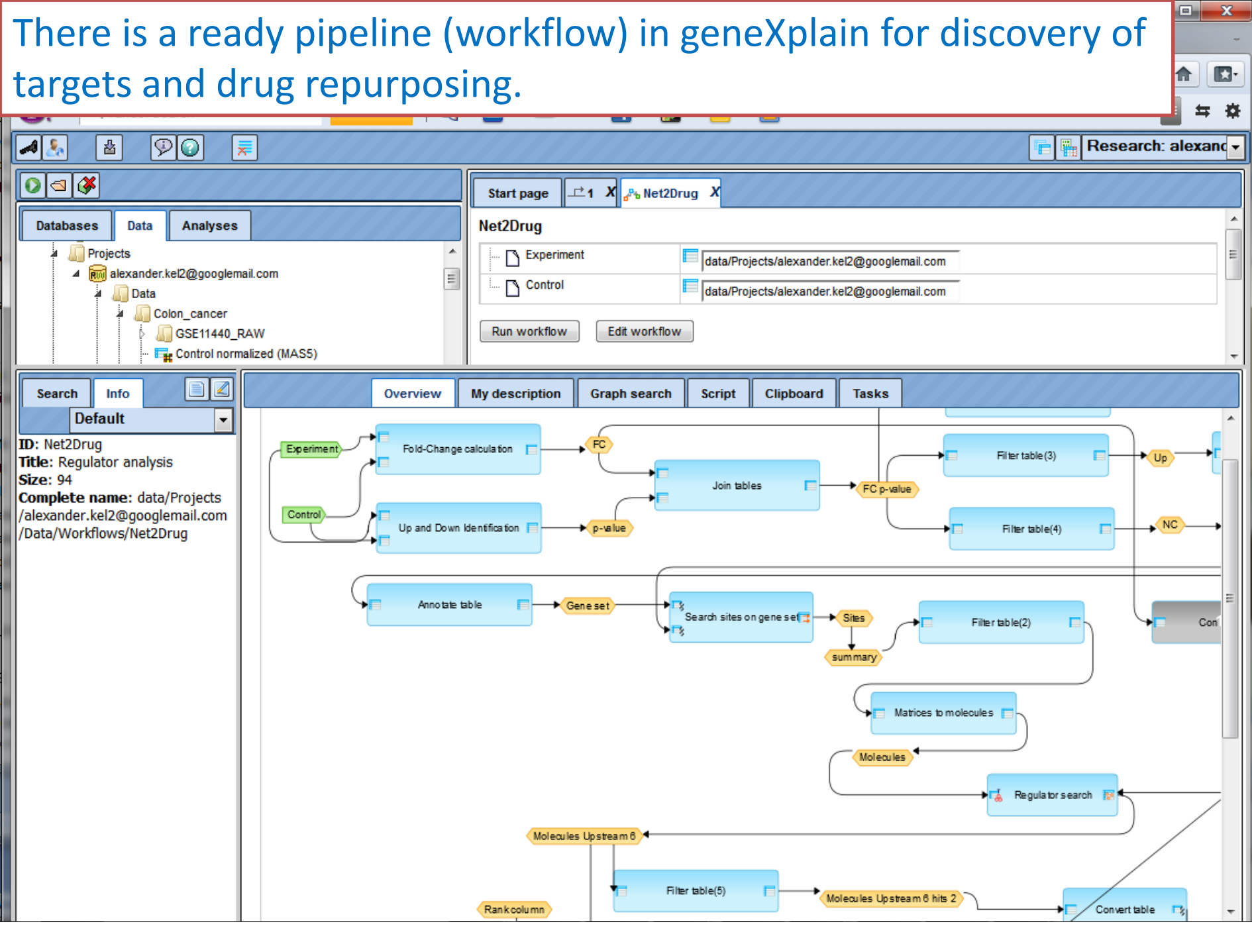
Status	Public on Sep 08, 2008
Title	Role of Caveolin 1, E-Cadherin, Enolase 2 and PKCa on resistance to methotrexate in human HT29 colon cancer cells
Organism	Homo sapiens
Experiment type	Expression profiling by array
Summary	A summary of the work associated to these microarrays is the following: Methotrexate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

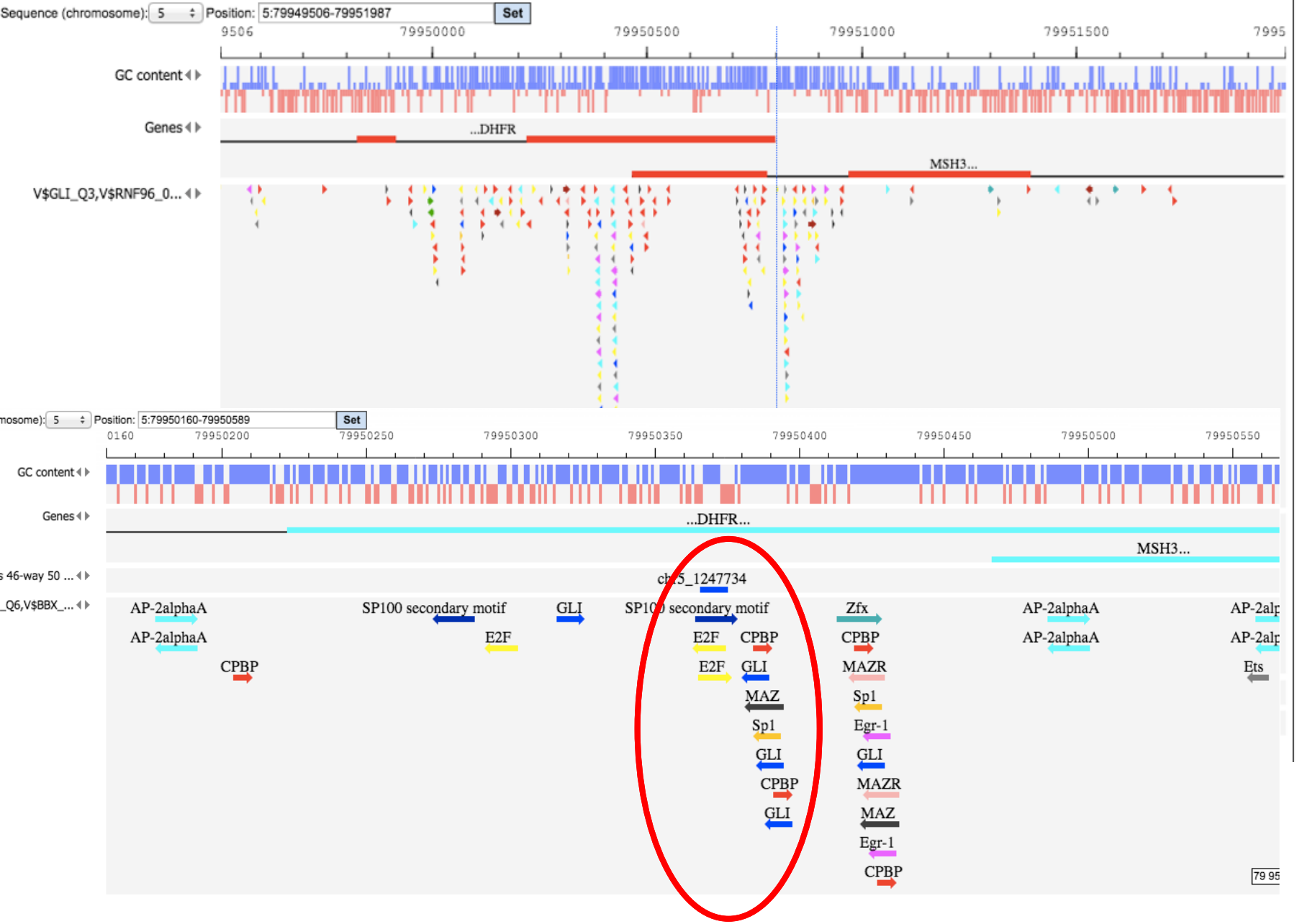


We took data on 3 MTX resistant patients versus 3 MTX sensitive and loaded them into geneXplain platform.

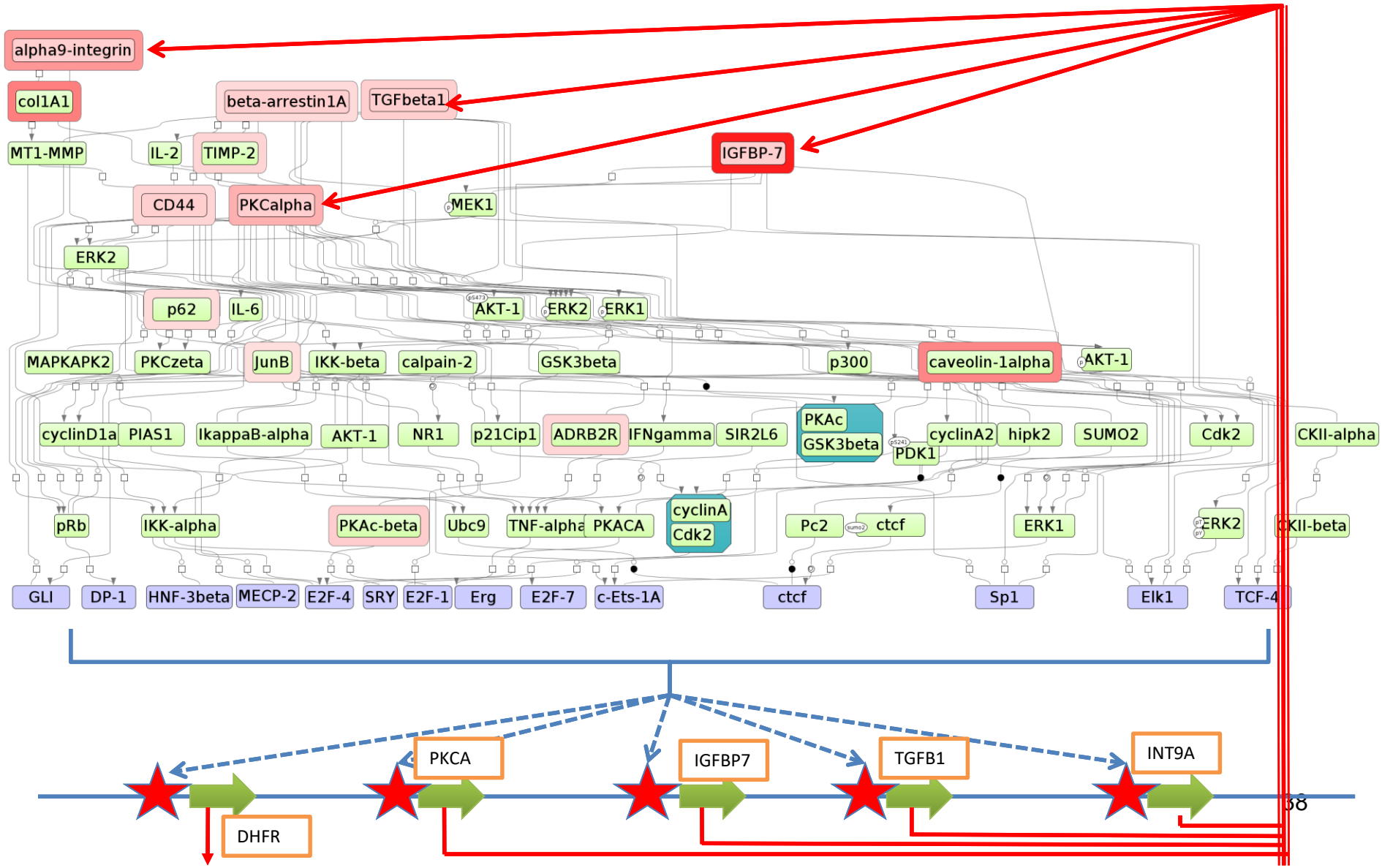


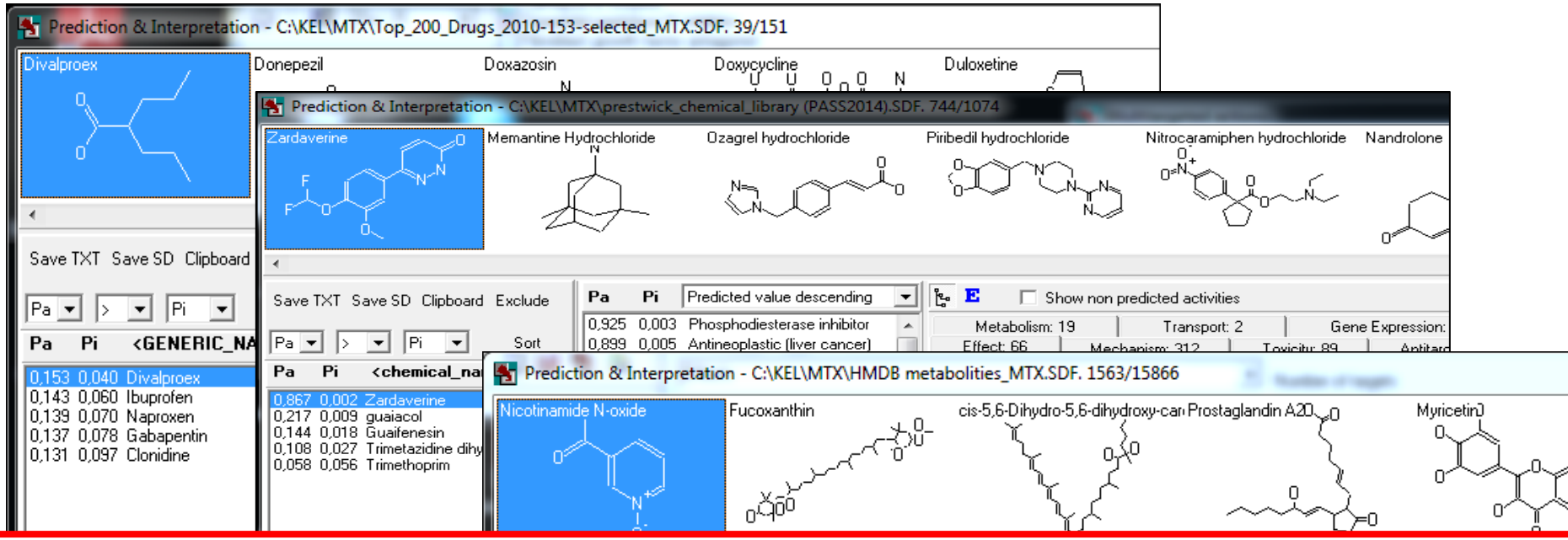
There is a ready pipeline (workflow) in geneXplain for discovery of targets and drug repurposing.




Sequence (chromosome): 5
Position: 5:79949506-79951987
Set
9506
79950000
79950500
79951000
79951500
7995
GC content
Genes
...DHFR
MSH3...
VGLI_Q3,VRNF96_0...
mosome): 5
Position: 5:79950160-79950589
Set
0160
79950200
79950250
79950300
79950350
79950400
79950450
79950500
79950550
GC content
Genes
...DHFR...
MSH3...
s 46-way 50 ...
Q6,V$BBX...
ch5_1247734
AP-2alphaA
AP-2alphaA
CPBP
SP100 secondary motif
E2F
GLI
SP100 secondary motif
E2F
CPBP
E2F
GLI
MAZ
Sp1
GLI
CPBP
GLI
Zfx
CPBP
MAZR
Sp1
Egr-1
GLI
MAZR
MAZ
Egr-1
CPBP
AP-2alphaA
AP-2alphaA
AP-2alp
AP-2alp
Ets
79 95

Resistance to MTX





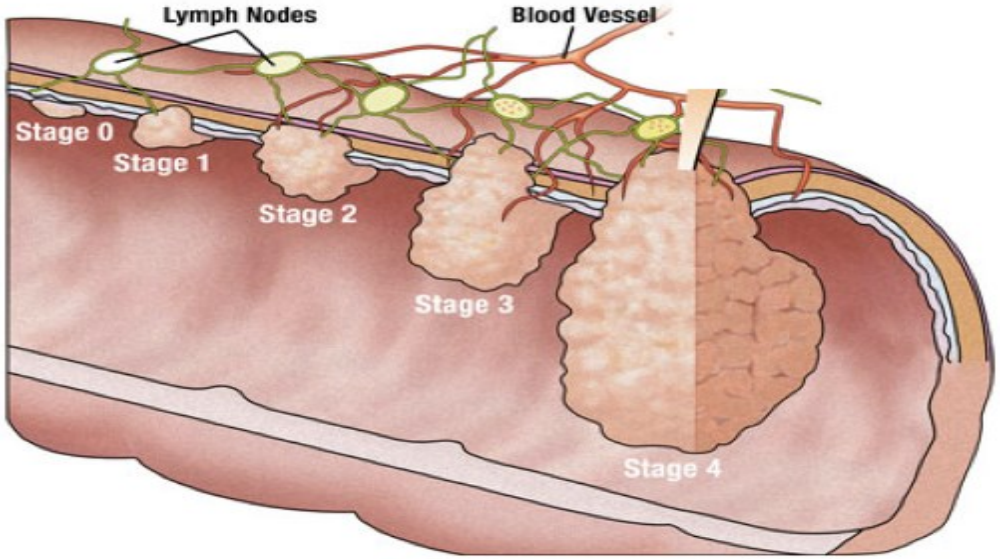
1) Divalproex, which is also known as valproic acid, is an old drug primarily used to treat epilepsy and bipolar disorder and to prevent migraine headaches. Recently a number of clinical trials were performed with this drug and they confirmed its efficacy for treatment of **Acute Myeloid Leukaemia, Cervical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


EARLY
DETECTION

Colorectal cancer



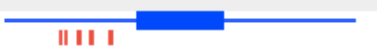
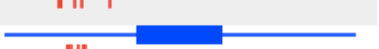
SYSCOL

SYSTEMS BIOLOGY OF COLORECTAL CANCER

SYSCOL

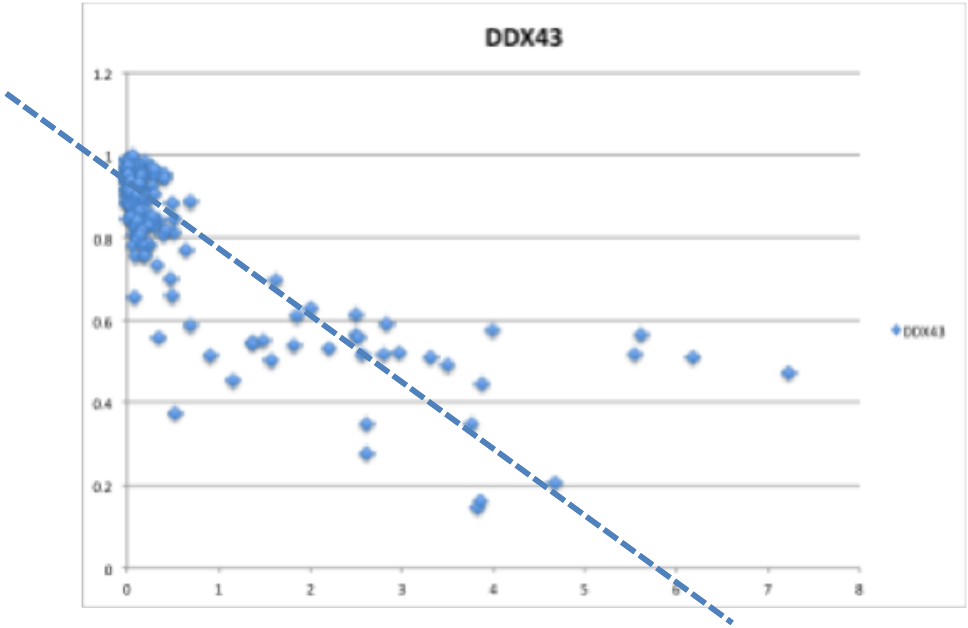
Systems Biology of Colorectal Cancer

SEVENTH FRAMEWORK PROGRAMME

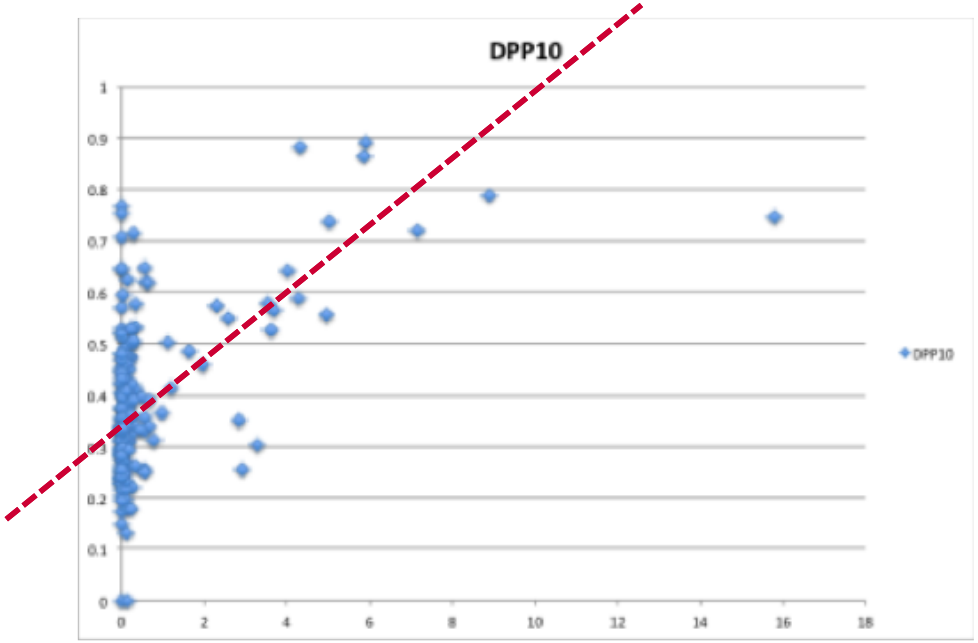
Gene symbol	Methylation vs Expression metastasis fc04 corr03: Count	Methylation vs Expression metastasis fc04 corr03: Count in exons	Methylation vs Expression metastasis fc04 corr03: Count in introns	Methylation vs Expression metastasis fc04 corr03: Count in 5'	Methylation vs Expression metastasis fc04 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

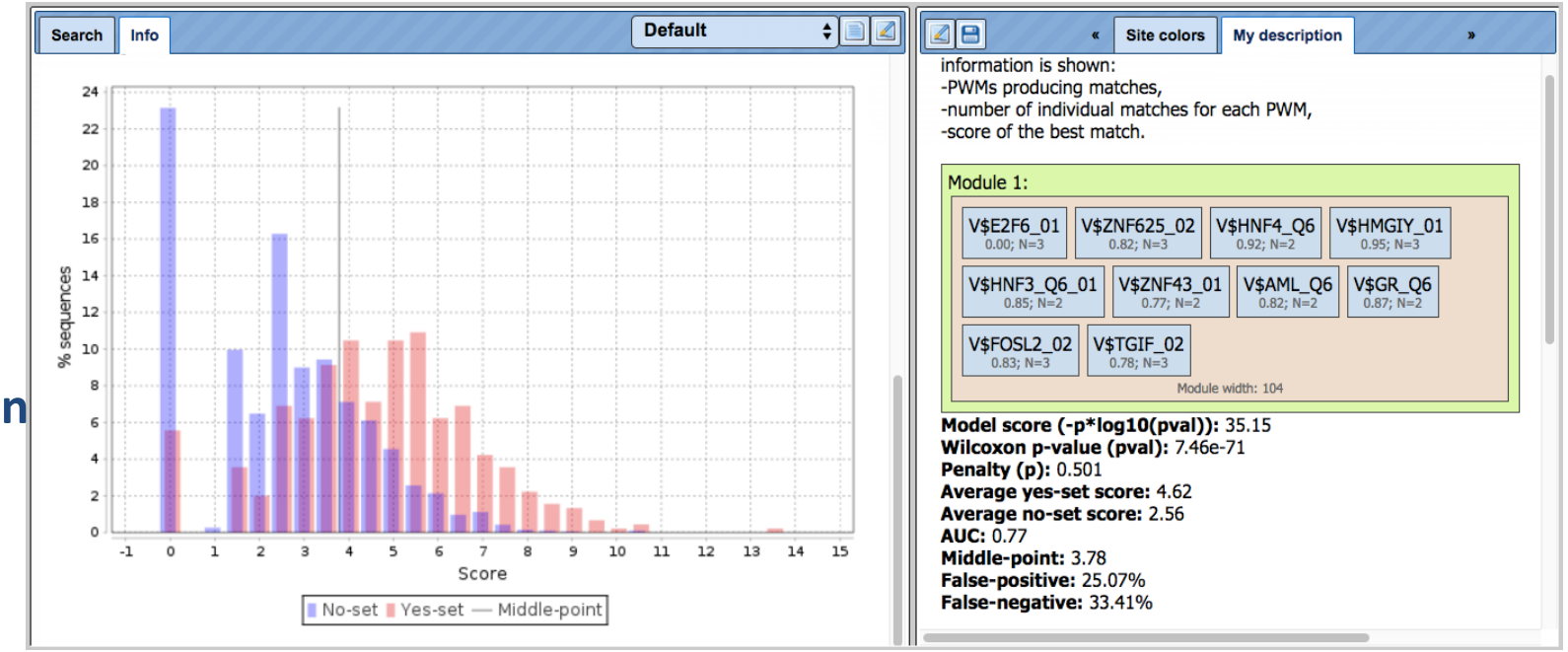


Positive correlation

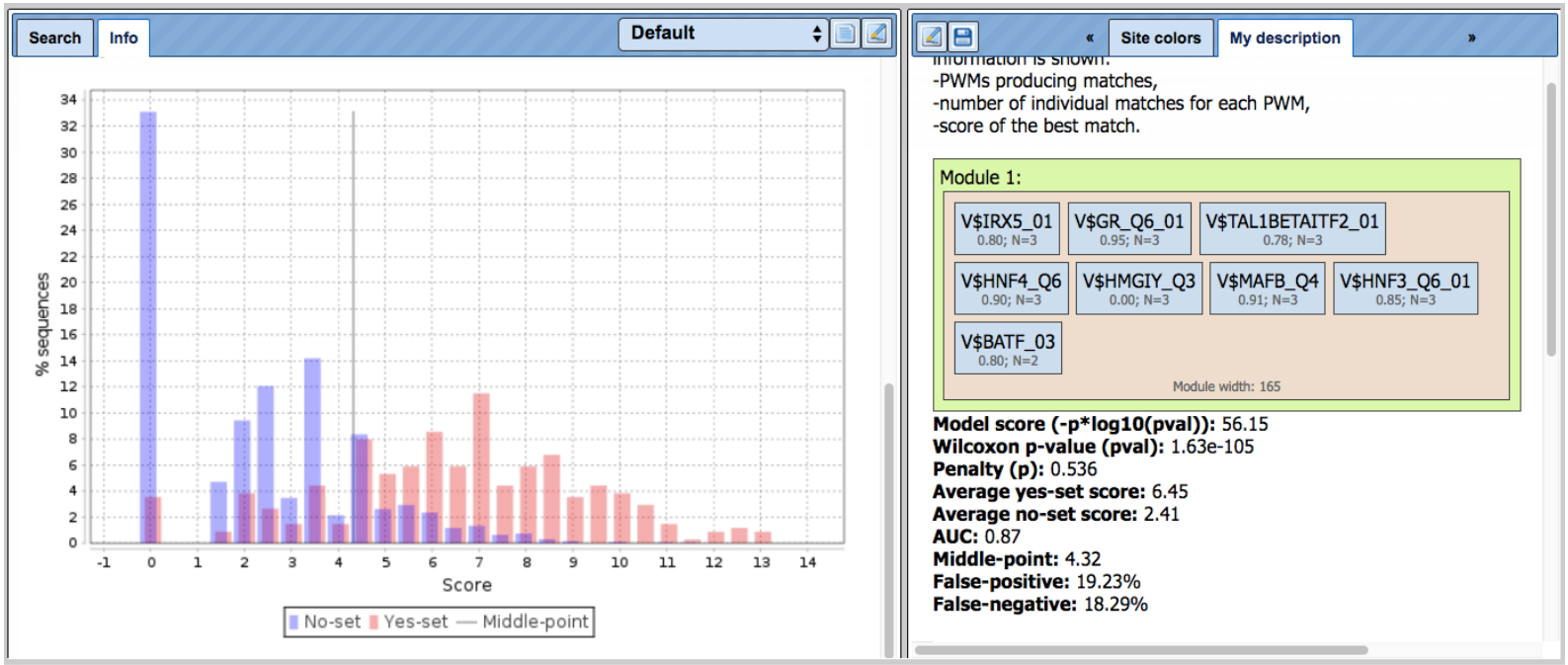


Meth. vs Expr.

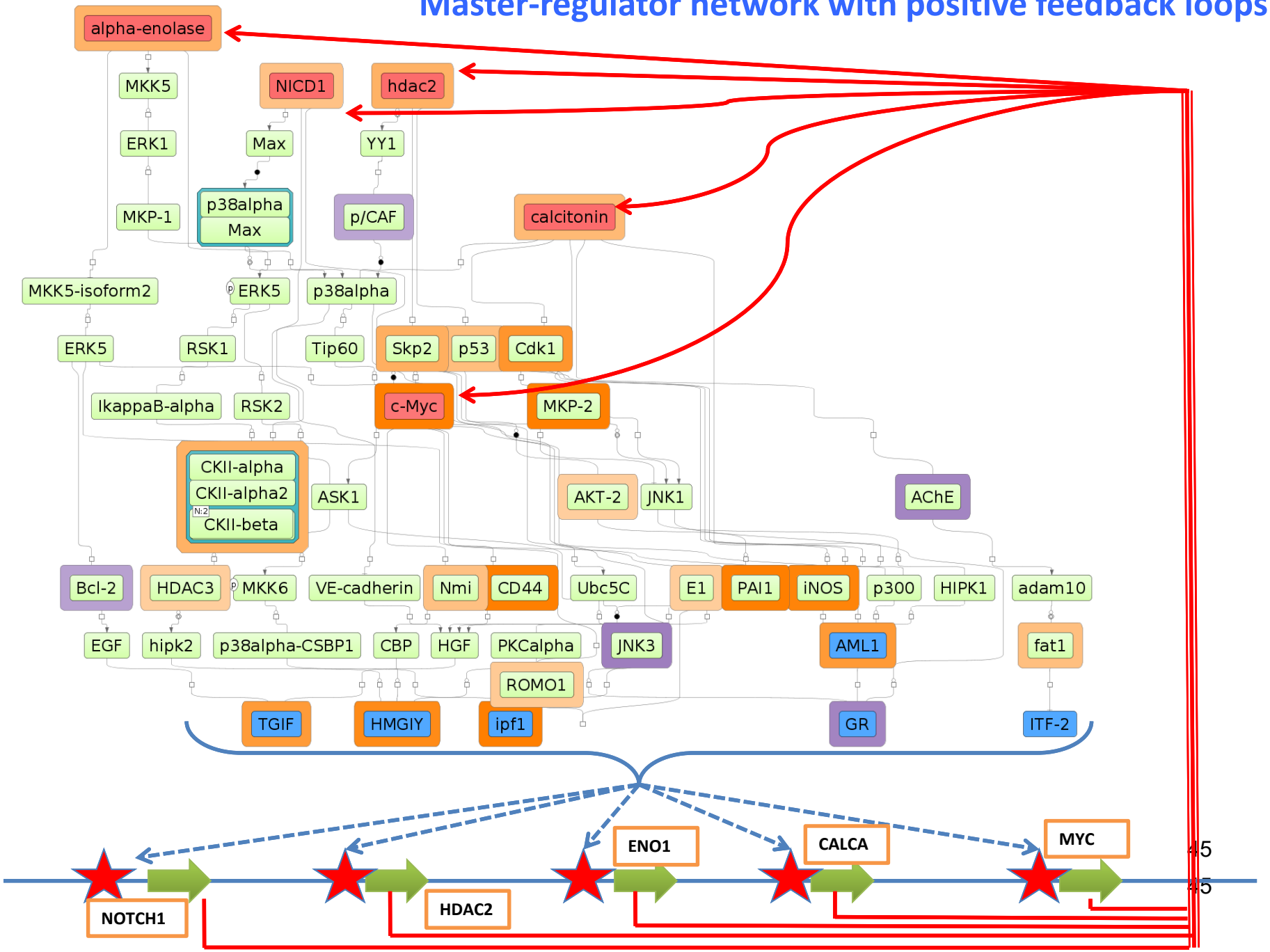
Negative correlation

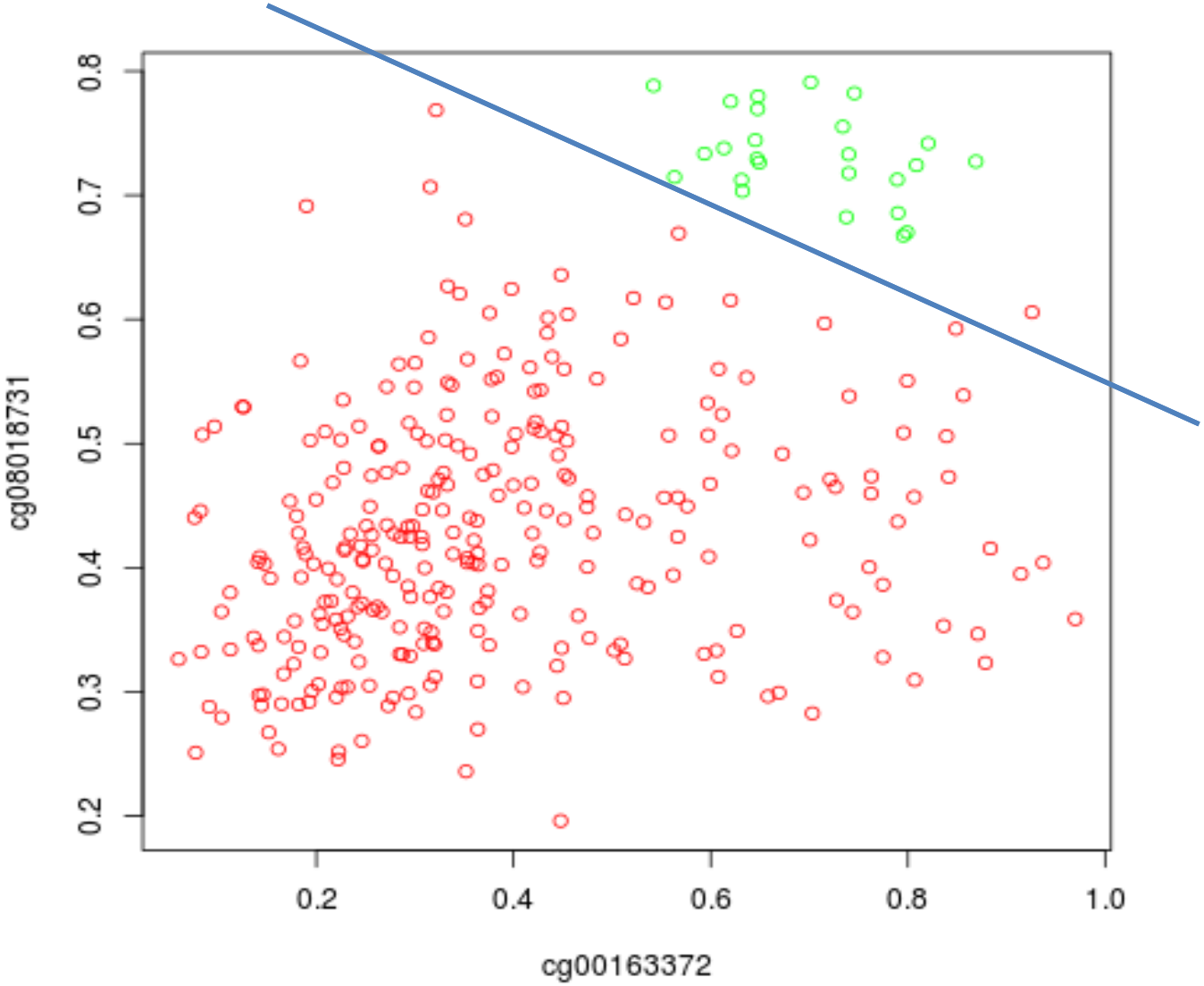


Positive correlation



Master-regulator network with positive feedback loops





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

6 CpG methylation biomarkers for yearly detection of CRC

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

93% success rate on independent cohort (from Siberia)

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 OverRignt	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 OverRignt	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	

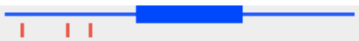
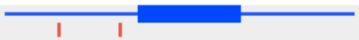
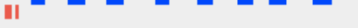
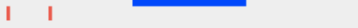
65 HIGH impact mutations in proteins.

Edit

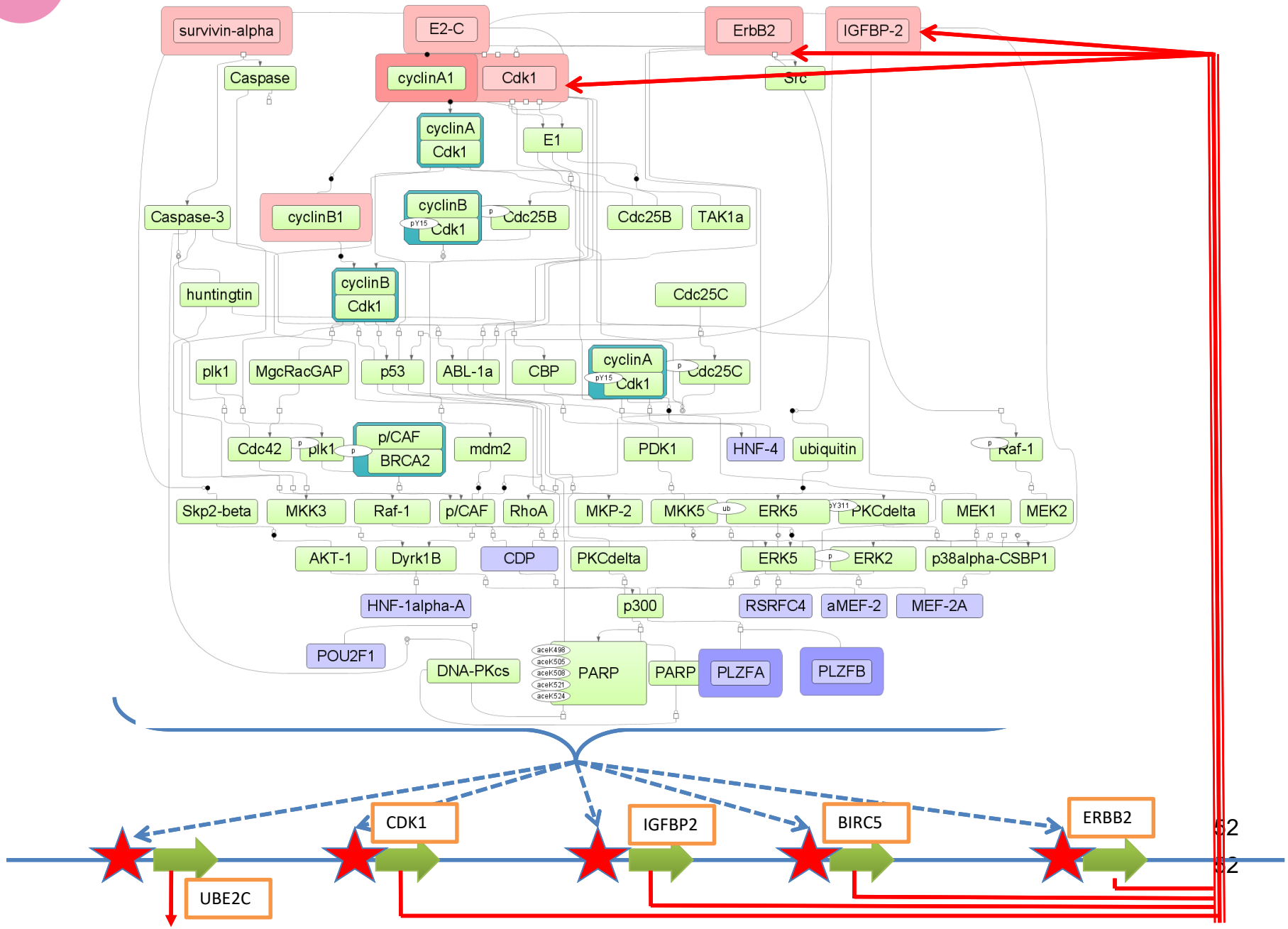
First Previous Page 1 of 2 Next Last Showing 1 to 50 of 65 entries

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
ENSG00000031691	centromere protein Q	CENPQ	6:49492214-49492214	G	stop_gained	HIGH
ENSG00000055609	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
ENSG00000121053	eosinophil	EBX	17:58196140-58196140	A	stop_gained	HIGH

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

ID	LogFC Endom	LogFC OverLinks	LogFC OverRignt	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 !

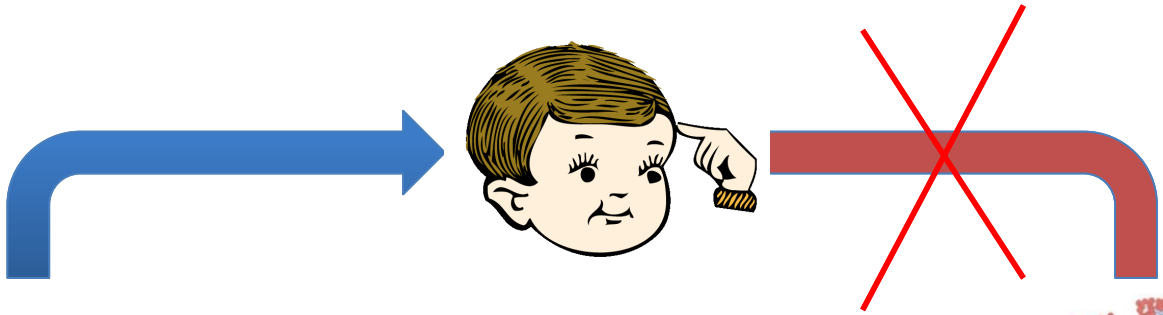
Sites do not exist !

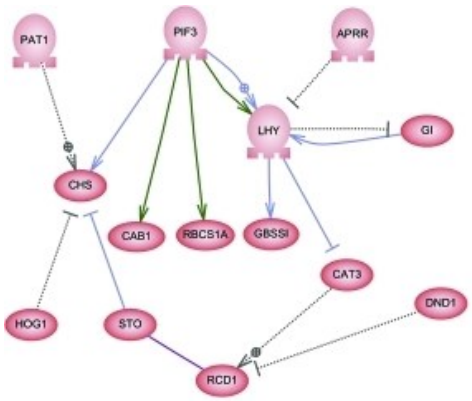
Pathways do not exist !

Things are very complex….So what?

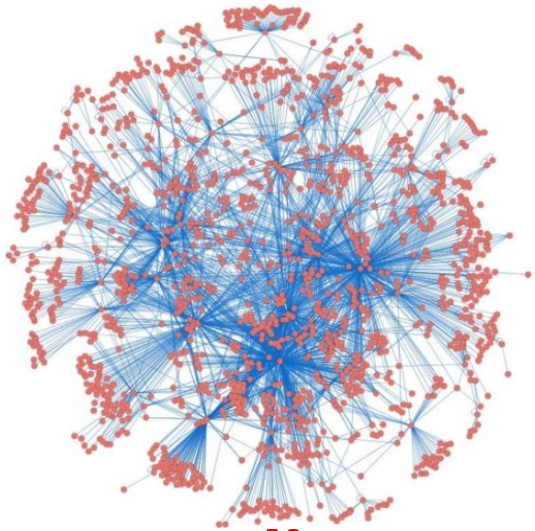
Simplified model
fits into our brain!

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



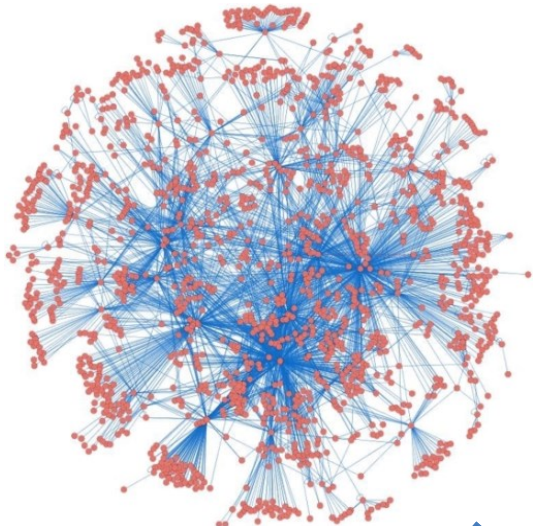


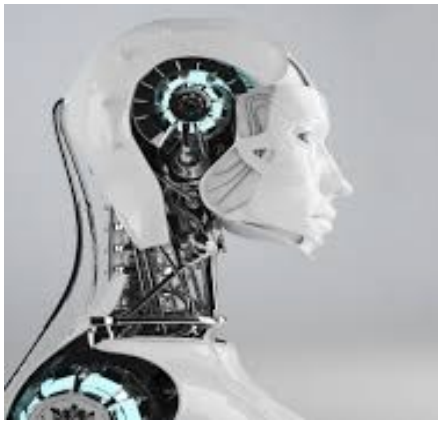
Model



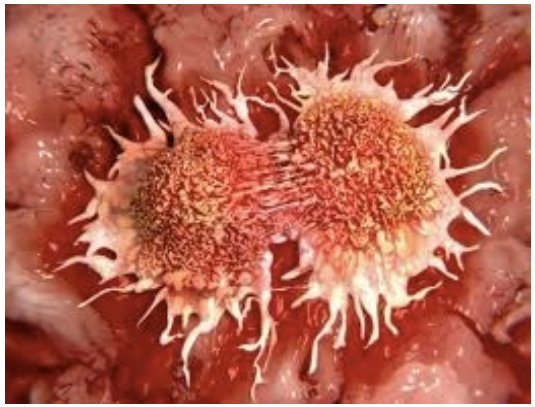
Reality

How to treat cancer?





Cancer



CADE METZ BUSINESS 01.27.16 01:00 PM

IN A HUGE BREAK GOOGLE'S AI BEA PLAYER AT THE G

News & Analysis

Microsoft, Google Beat Humans at Image Recognition

Deep learning algorithms compete at ImageNet challenge

DeepMind

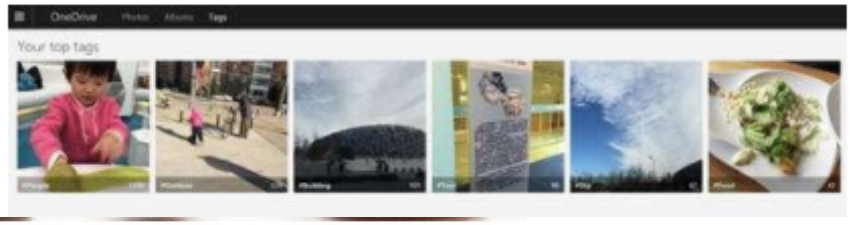
MUST READ NEW DOCUMENTS REVEAL FBI PAID GEEK SQUAD REPAIR STAFF AS INFORMANTS

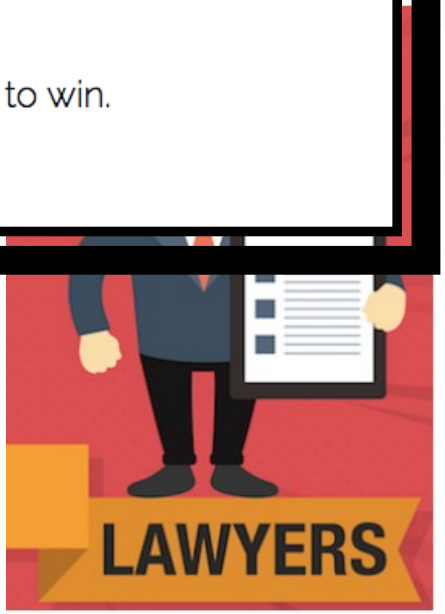
Gaming AI beats human top scores by cheating

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

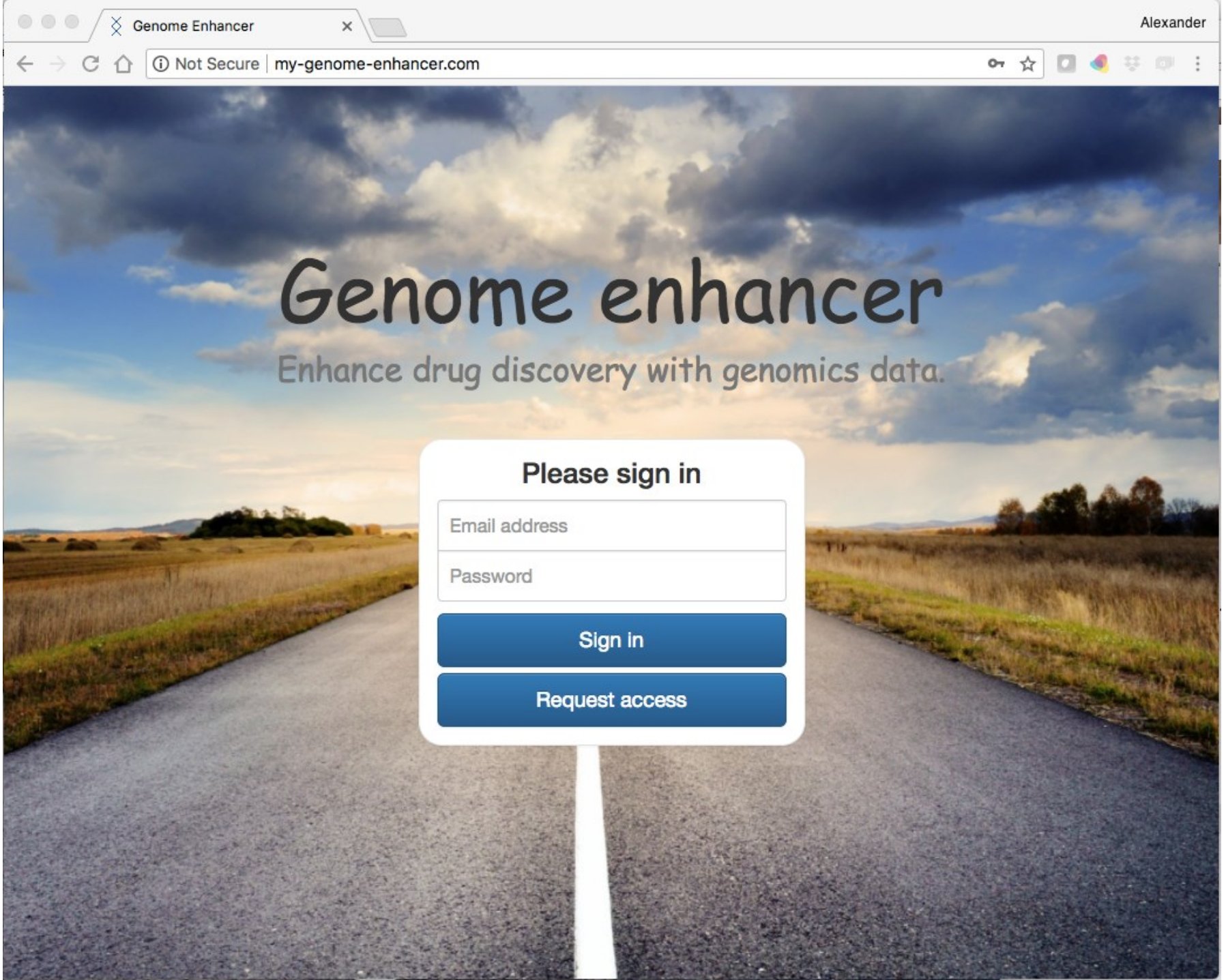
By Charlie Osborne for Between the Lines | March 2, 2018 -- 11:12 GMT (11:12 GMT) | Topic: Innovation

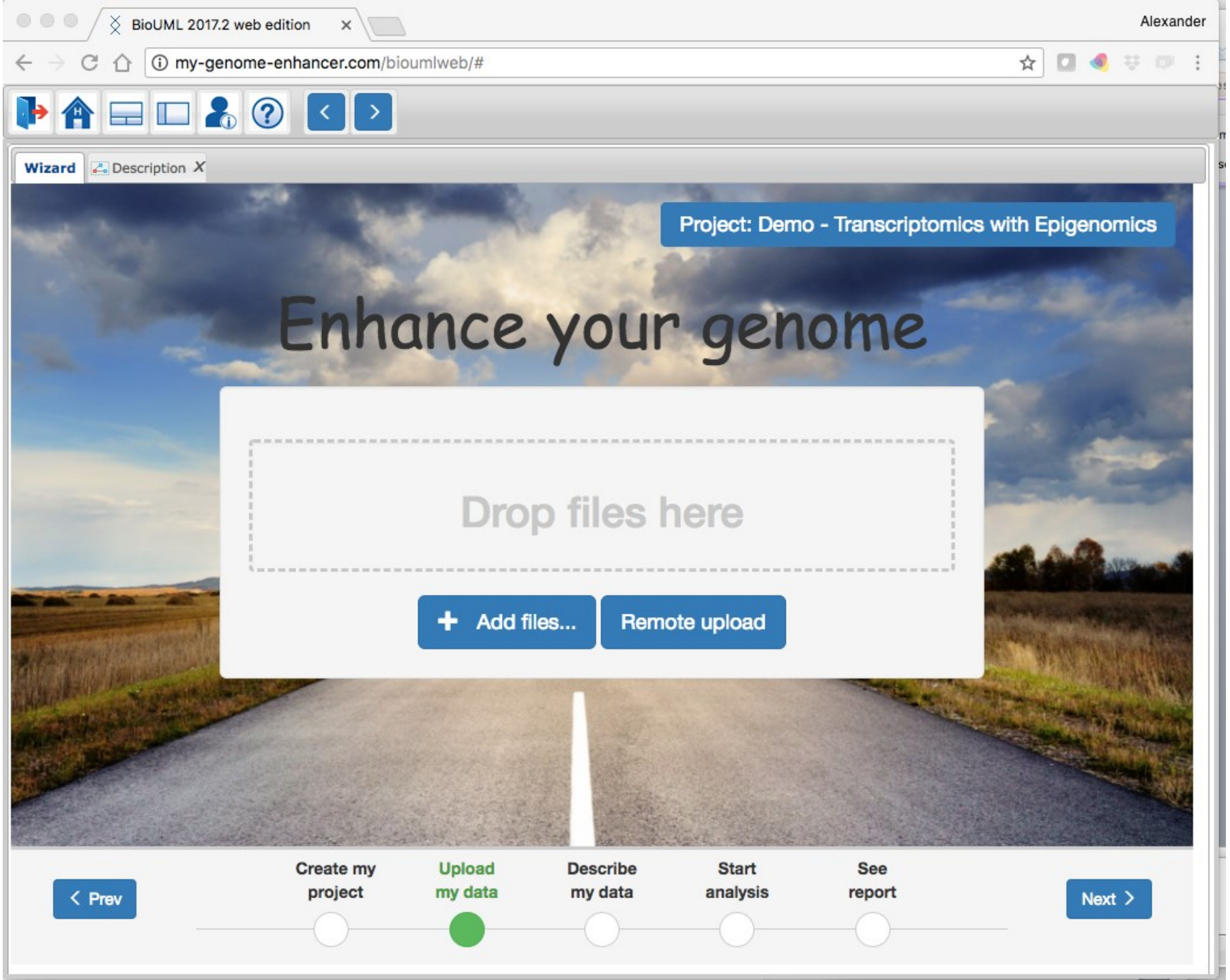
human benchmark of 5.1% errors with a 4.94% error grabbing neural network, Google announced it had one-upped Microsoft by 0.04%.





▲ 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
Alexander
Not Secure
my-genome-enhancer.com
Genome enhancer
Enhance drug discovery with genomics data.
Please sign in
Email address
Password
Sign in
Request access


BioUML 2017.2 web edition
Alexander
my-genome-enhancer.com/bioumlweb/#
Wizard
Description
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

Sequence and Pathway analysis

Identification of master-regulators in gene regula transduction pathways.

Alexander
geneXplai
Data received on

Data

For this study the following

Table 1. Experimental datasets u

File name
Normalized (RMA) (1)

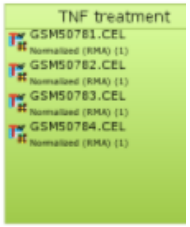
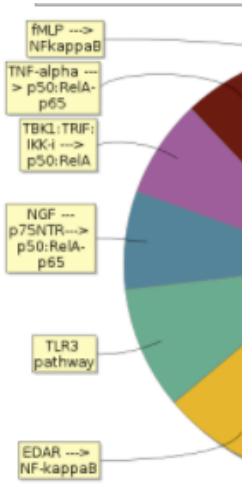


Figure 1. Annotation diagram of
data that are compared in our ani



Summary

In this repo
automatizec
The goal of
targets for
factors (TFs
search for s
The identifi
checkup, th
pathology.

Here we ap
transcripton
involved in
CXCL8, RP
us to better
an approac
with true po

Introducti

Multiple "-o
genetic and
pathological
mechanism
comparison
networks c
regulatory n
pathology p
potential m
pathological
pathological
reconstruct
such approa
actually do

Results

We have run the workflow t

Identification of target

In the first step of the ana
applied the Limma tool (R
expression in the following
on the basis of 2 of the fol
(corrected to the multiple te
(Supplementary table 1) a
LogFC>0 for up-regulated a
genes according to their p-v
most downregulated genes

Table 2. Identified **upregulated** g
See full table →

ID	Gene symbol
ENSG00000162692	VCAM1
ENSG00000007908	SELE
ENSG00000090339	ICAM1
ENSG00000100906	NFKBIA
ENSG00000110848	CD69
ENSG00000104312	RIPK2

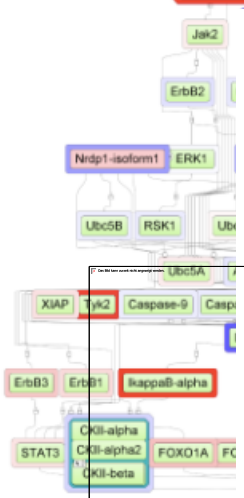
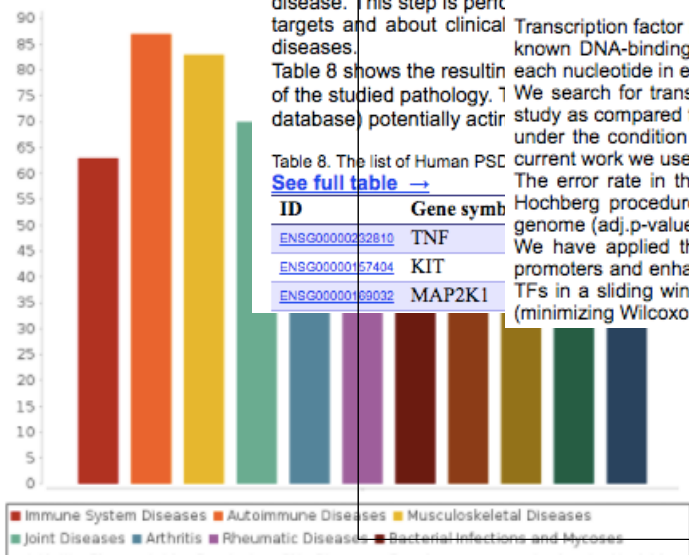


Figure 10. Diagram of intracellu
See full diagram →

Figure 3. Enriched TRANSPATH Pat
Full classification →

HumanPSD(TM) disease

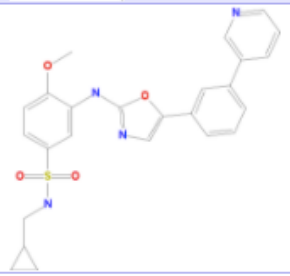


Conclusion

At the last step of the ar
activation) of the identifie
disease. This step is perfo
targets and about clinical
diseases.
Table 8 shows the resultin
of the studied pathology. T
database) potentially actin

Table 8. The list of Human PSD
See full table →

ID	Gene symb
ENSG00000232810	TNF
ENSG00000157404	KIT
ENSG00000169032	MAP2K1

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-PYRIDINYL)PHENYL]-1,3-OXAZOL-2-YL}AMINO)BENZENESULFO...		NEK7, MAPK6, TRAF6, CDK1, CCNA1, DUSP4, BARD1...	4.19	0.16

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).

Robot-scientist

Is this our future?