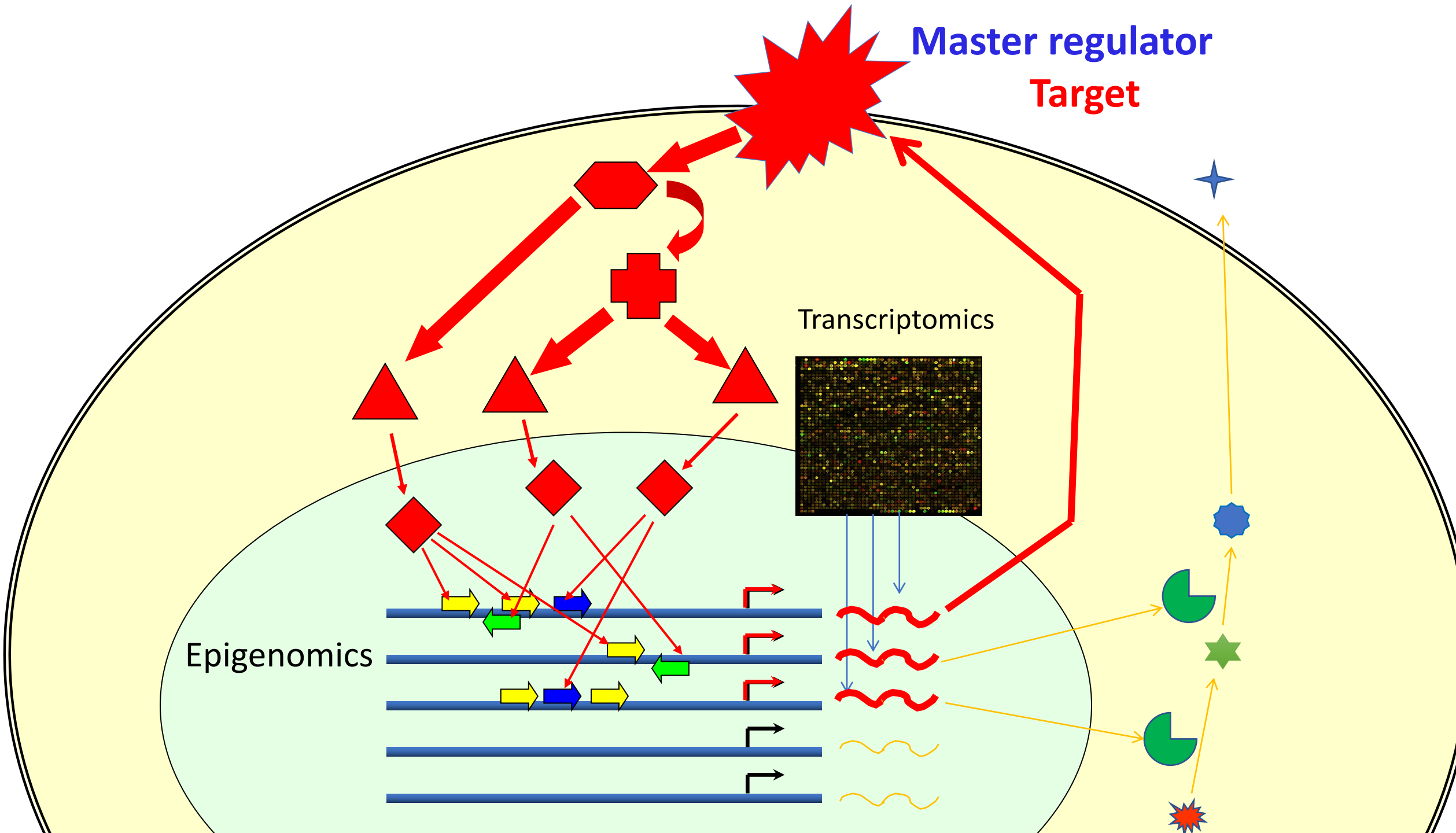
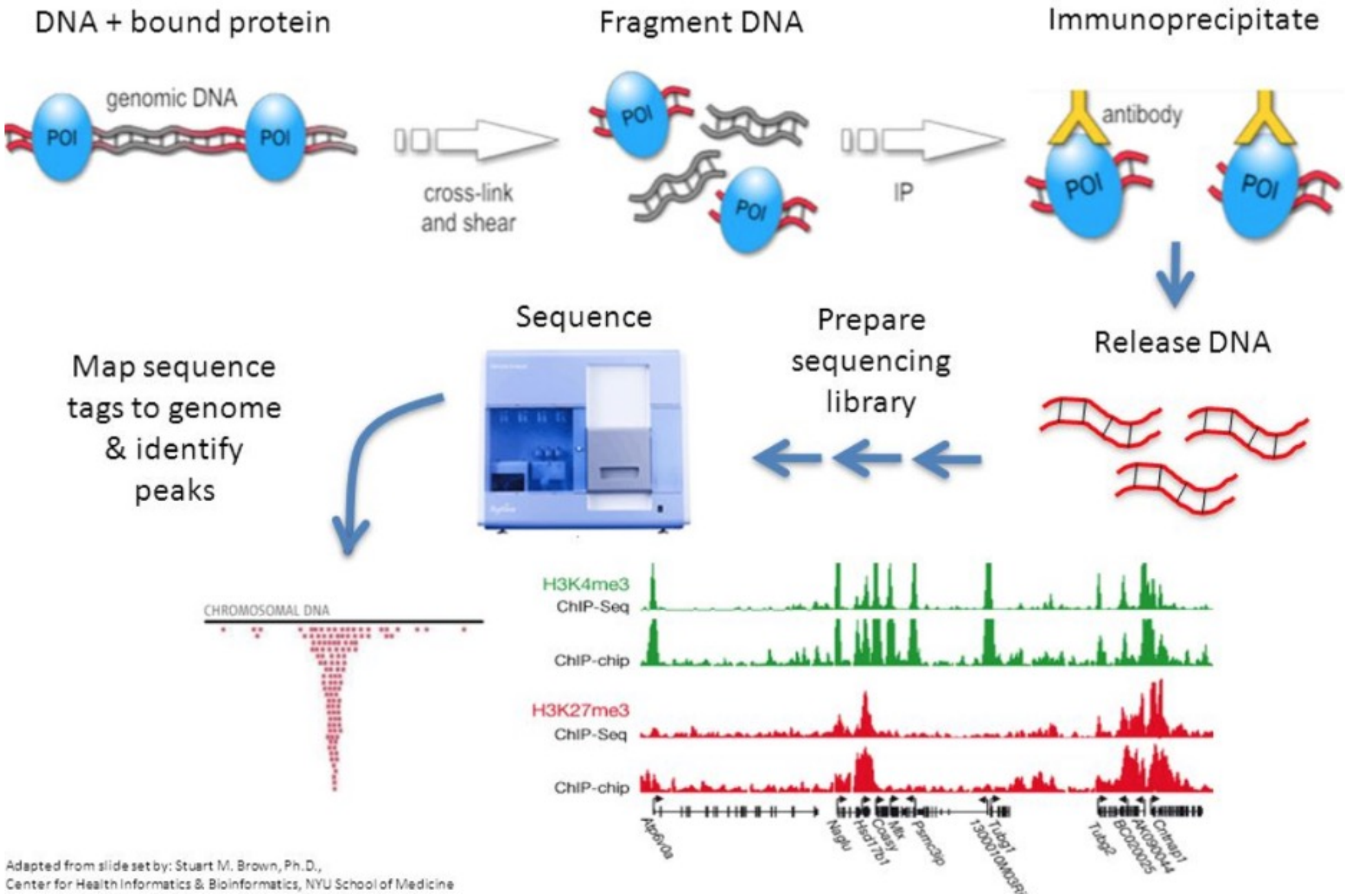
Let's listen to a symphony of epigenomics when seeking for drug targets

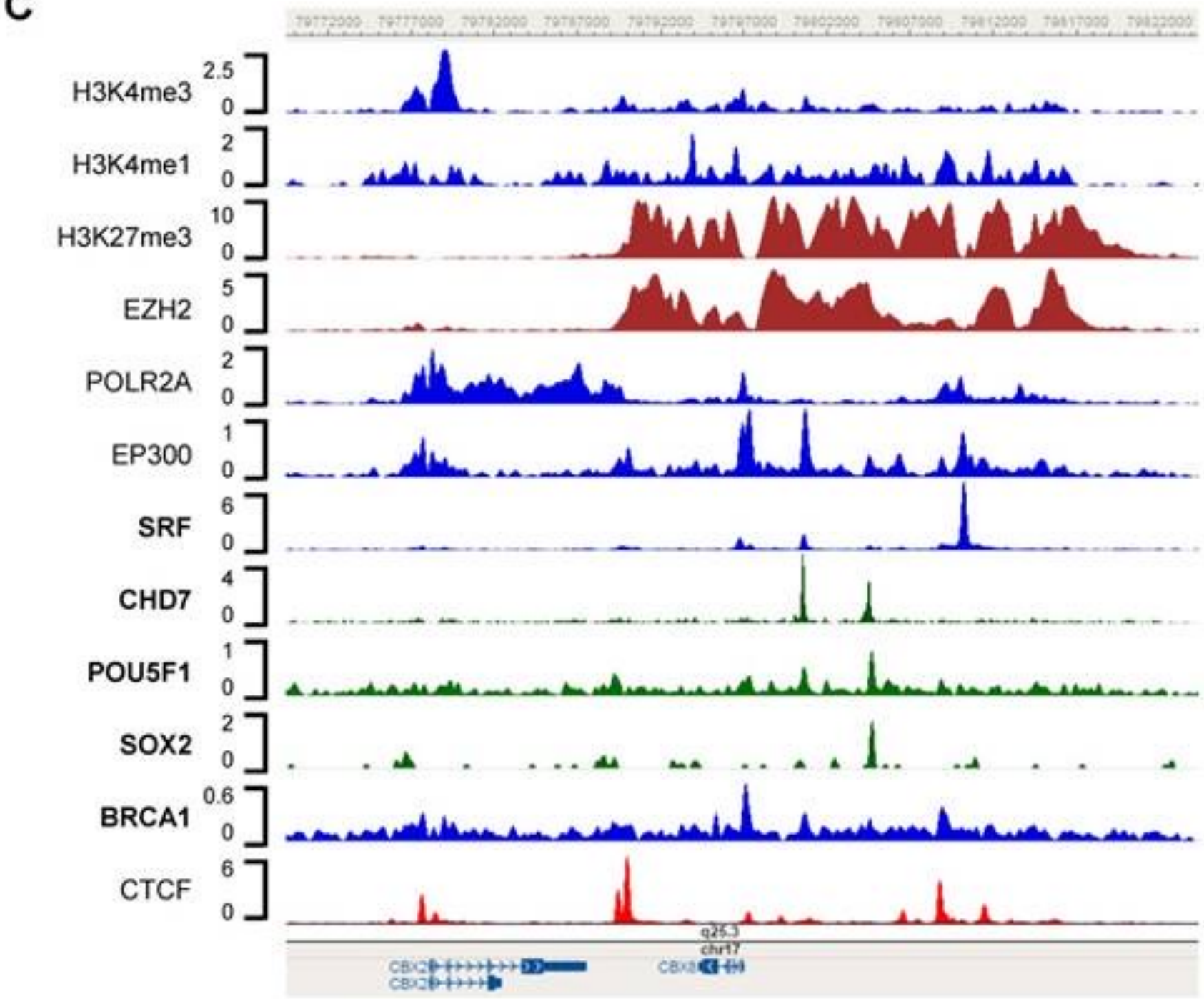
Alexander Kel

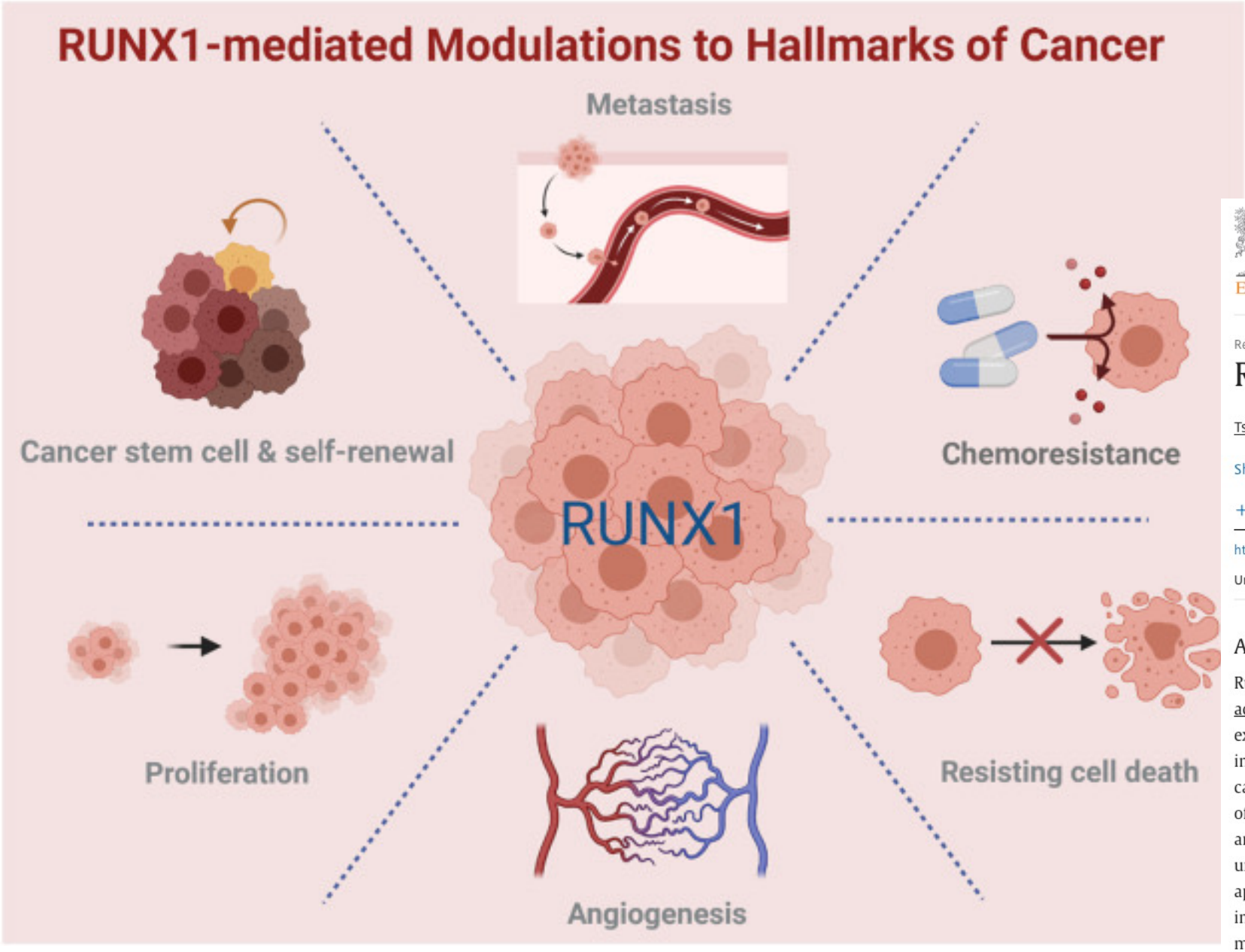

Master regulator
Target
Transcriptomics
Epigenomics



Adapted from slide set by: Stuart M. Brown, Ph.D.,
Center for Health Informatics & Bioinformatics, NYU School of Medicine

C


H3K4me3
2.5
0
H3K4me1
2
0
H3K27me3
10
0
EZH2
5
0
POLR2A
2
0
EP300
1
0
SRF
6
0
CHD7
4
0
POU5F1
1
0
SOX2
2
0
BRCA1
0.6
0
CTCF
6
0
q25.3
chr17



ELSEVIER
Biochimica et Biophysica Acta (BBA) - Reviews on Cancer
Volume 1877, Issue 3, May 2022, 188715

Review

RUNX1 and cancer

Tsung-Chieh Lin

Show more

+ Add to Mendeley Share Cite

https://doi.org/10.1016/j.bbcan.2022.188715

Get rights and content

Abstract

Runt-related transcription factor 1 (RUNX1) is frequently involved in the progression of acute leukemia. However, emerging and discoverable *RUNX1* somatic mutations, *RUNX1* expressional signatures in normal tissues and cancers, and *RUNX1*'s clinical significance in many cancer types have attracted attention for considering *RUNX1* as a biomarker for cancer. Recent discoveries have demonstrated the indirect and direct biological functions of RUNX1 in modulating cancer metastasis, proliferation, angiogenesis, cancer stemness and chemoresistance to anticancer drugs, warranting the further investigations of the underlying mechanisms to provide knowledge for developing a novel therapeutic approach. In this review article, we focused mainly on recent research developments involving oncogenic activities of RUNX1 by summarizing and integrating *RUNX1* somatic mutations, clinical trials, transcriptome data, clinical information and the discoveries related to the RUNX1-induced signaling pathway in modulating malignant phenotypes. Furthermore, a comprehensive demonstration of *RUNX1* RNA expression in a pancancer panel and specific normal cell types at single-cell level were presented, and the results suggest potential sites and cell types of RUNX1-related tumorigenesis. With this review

Series GSE129314

Status	Public
Title	Genom
Organism	Homo
Experiment type	Genom
Summary	The ov MCF10 followe cells. not id genera transd expres doxycy
Overall design	Two sa
Contributor(s)	Huang
Citation(s)	Malik N suppre *Nat Co*

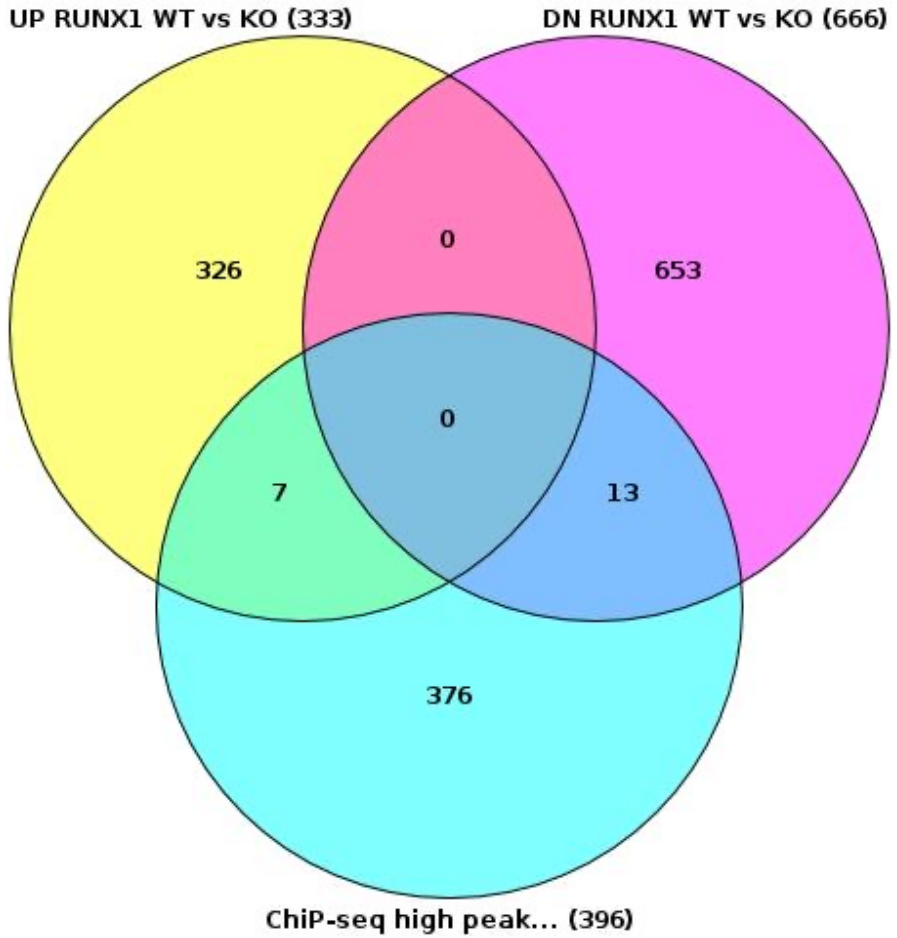
Series GSE120216

Query DataSets for GSE120216

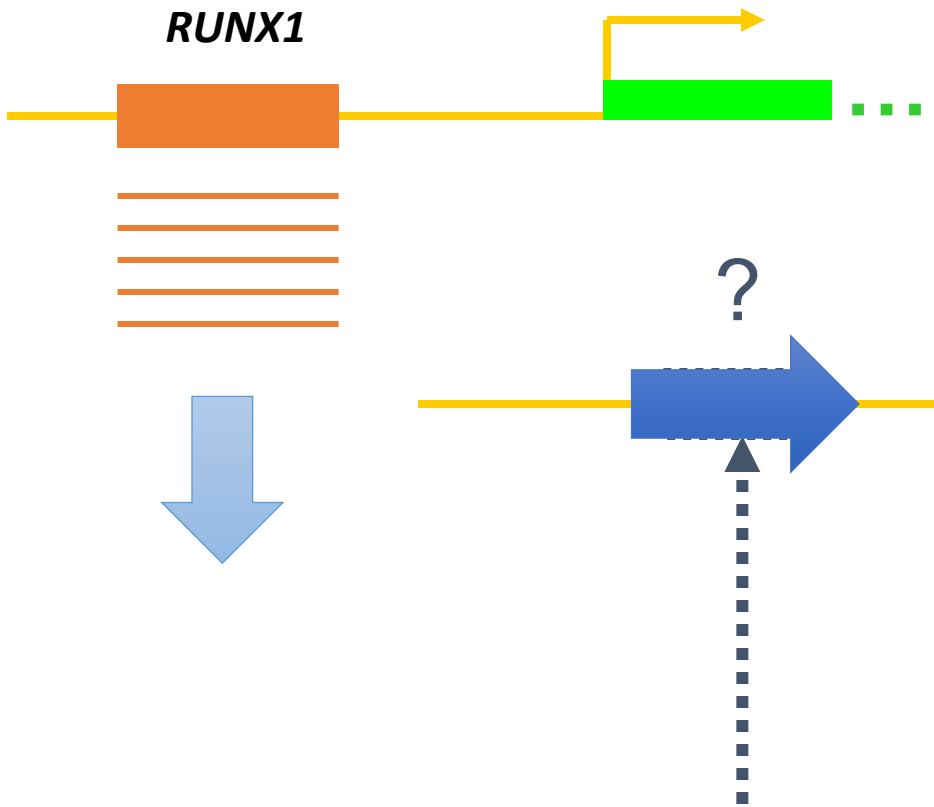
Status	Public on Apr 18, 2019
Title	Identifying transcripts that are transcriptinoally regulated by CBFB and RUNX1 using RNAseq
Organism	Homo sapiens
Experiment type	Expression profiling by high throughput sequencing
Summary	Using RNAseq to identify differentially expressed transcripts between CBFB wild type (WT) and knockout (KO) or between RUNX1 wild type (WT) and knockout (KO) MCF10A cells.
Overall design	Three repeats for CBFB KO and CBFB WT MCF10A cells. Four repeats for RUNX1 KO and RUNX1 WT MCF10A cells. One CBFB KO clone (#751) and two RUNX1 KO clones (5008 and 5010) were used.
Contributor(s)	Huang J, Malik N
Citation(s)	Malik N, Yan H, Moshkovich N, Palangat M et al. The transcription factor CBFB suppresses breast cancer through orchestrating translation and transcription. *Nat Commun* 2019 May 6;10(1):2071. PMID: 31061501

Big peaks are NOT in DEGs !

ID	Gene symbol	GSE129314_RUNX1ChIP_MCF10AOEq0.05_peaks track enrich10: Schematic	GSE129314_RUNX1ChIP_MCF10AOEq0.05_peaks track enrich10: Count
ENSG00000159216	RUNX1		13
ENSG00000162599	NFIA		3
ENSG00000286153	ENSG00000286153		3
ENSG00000069667	RORA		2
ENSG00000085733	CTTN		2
ENSG00000113719	ERGIC1		2
ENSG00000129071	MBD4		2
ENSG00000134871	COL4A2		2
ENSG00000136928	GABBR2		2
ENSG00000161217	PCYT1A		2
ENSG00000169641	LUZP1		2
ENSG00000224184	MIR3681HG		2
ENSG00000235618	FAM21EP		2
ENSG00000237686	SCIRT		2
ENSG00000249751	ECSCR		2
ENSG00000279686	ENSG00000279686		2
ENSG00000001630	CYP51A1		1
ENSG00000002746	HECW1		1
ENSG00000007402	CACNA2D2		1
ENSG00000008130	NADK		1
ENSG00000010282	HHATL		1
ENSG00000011332	DPF1		1
ENSG00000015133	CCDC88C		1
ENSG00000016402	IL20RA		1



Search for new TF binding sites with PWMs



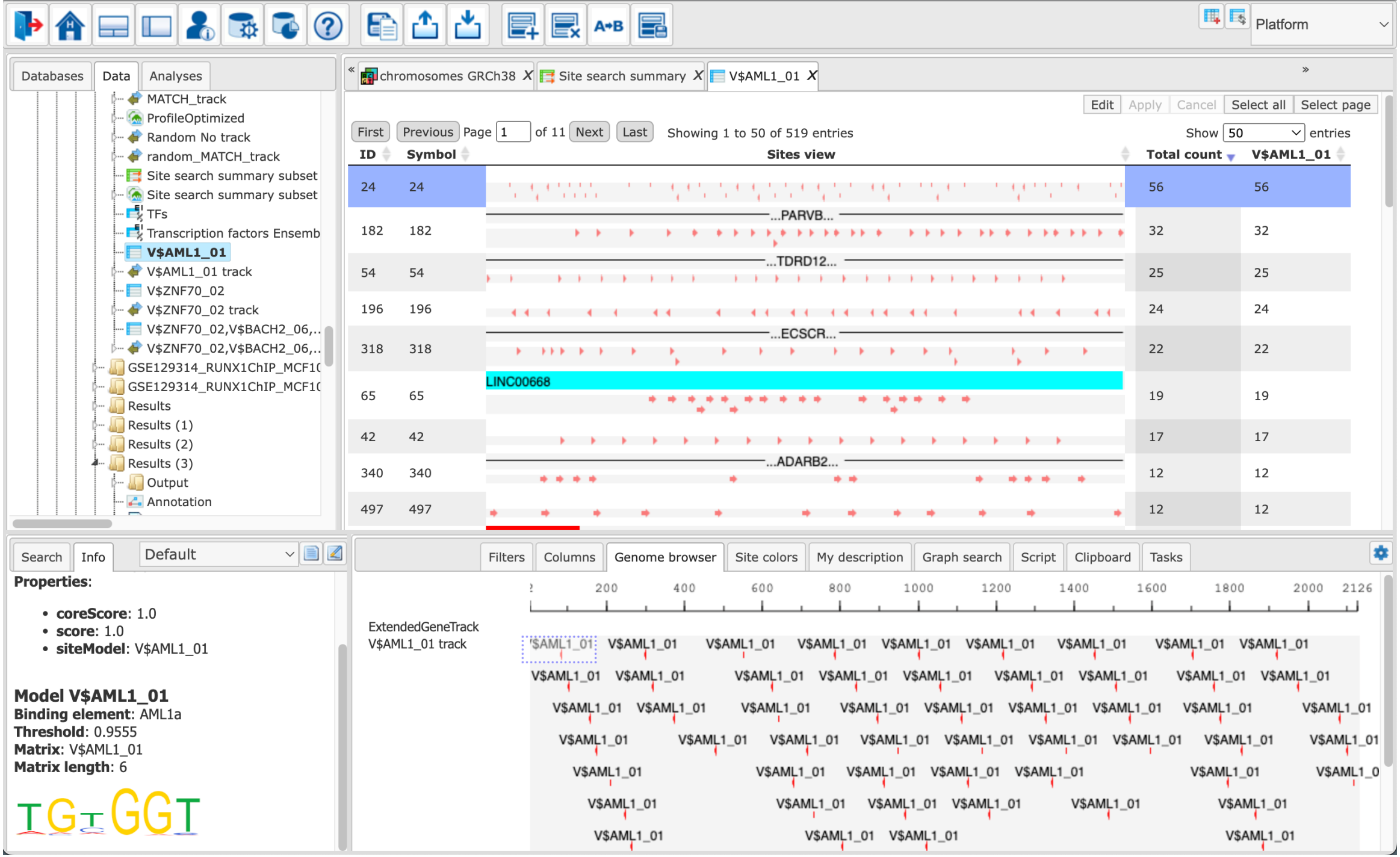
$$q = \frac{\sum_{i=1}^{l} I(i) f(b_i, i) - \sum_{i=1}^{l} I(i) f^{\min}(i)}{\sum_{i=1}^{l} I(i) f^{\max}(i)} \quad (1)$$

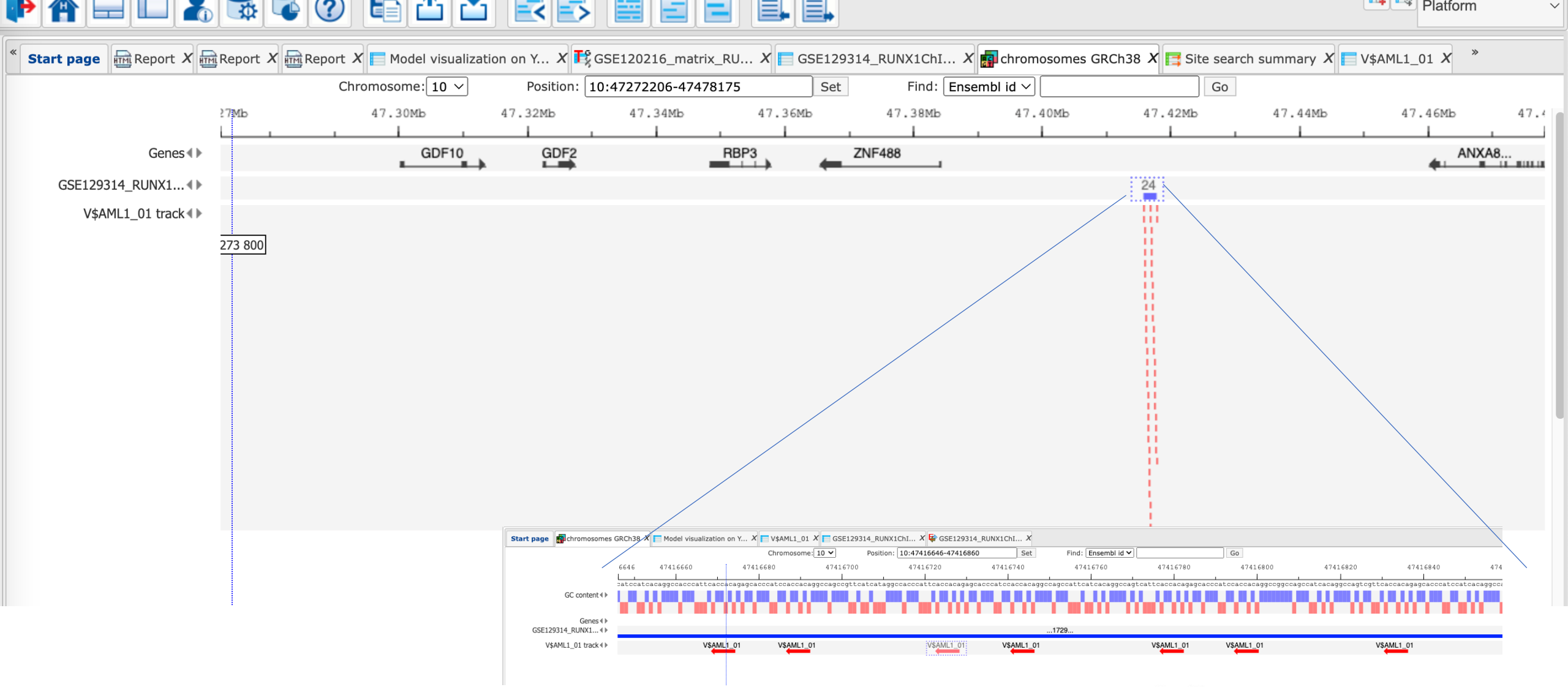
$$I(i) = \sum_{b \in \{A,T,G,C\}} f(b,i) \ln(4 f(b,i)) \quad (2)$$

Logo

A	0.172	0.092	0.114	0.122	0.068	0.008	0.034	0.002	0.002	0.018	0.032	0.176	0.114	0.098
C	0.076	0.130	0.130	0.398	0.010	0.016	0.028	0.000	0.004	0.146	0.138	0.026	0.110	0.130
G	0.316	0.522	0.292	0.176	0.016	0.924	0.026	0.998	0.974	0.112	0.160	0.192	0.388	0.380
T	0.436	0.256	0.464	0.304	0.906	0.052	0.912	0.000	0.020	0.724	0.670	0.606	0.388	0.392

TRANSFAC®


Platform
Databases
Data
Analyses
MATCH_track
ProfileOptimized
Random No track
random_MATCH_track
Site search summary subset
Site search summary subset
TFs
Transcription factors Ensemb
V$AML1_01
V$AML1_01 track
V$ZNF70_02
V$ZNF70_02 track
V$ZNF70_02,V$BACH2_06,..
V$ZNF70_02,V$BACH2_06,..
GSE129314_RUNX1ChIP_MCF10
GSE129314_RUNX1ChIP_MCF10
Results
Results (1)
Results (2)
Results (3)
Output
Annotation
chromosomes GRCh38
Site search summary
V$AML1_01
Edit
Apply
Cancel
Select all
Select page
First
Previous
Page 1 of 11
Next
Last
Showing 1 to 50 of 519 entries
Show 50 entries
ID
Symbol
Sites view
Total count
V$AML1_01
24 24 56 56
182 182 ...PARVB... 32 32
54 54 ...TDRD12... 25 25
196 196 24 24
318 318 ...ECSCR... 22 22
65 65 LINC00668 19 19
42 42 17 17
340 340 ...ADARB2... 12 12
497 497 12 12
Search
Info
Default
Properties:
coreScore: 1.0
score: 1.0
siteModel: V$AML1_01
Model V$AML1_01
Binding element: AML1a
Threshold: 0.9555
Matrix: V$AML1_01
Matrix length: 6
Filters
Columns
Genome browser
Site colors
My description
Graph search
Script
Clipboard
Tasks
200 400 600 800 1000 1200 1400 1600 1800 2000 2126
ExtendedGeneTrack
V$AML1_01 track
V$AML1_01



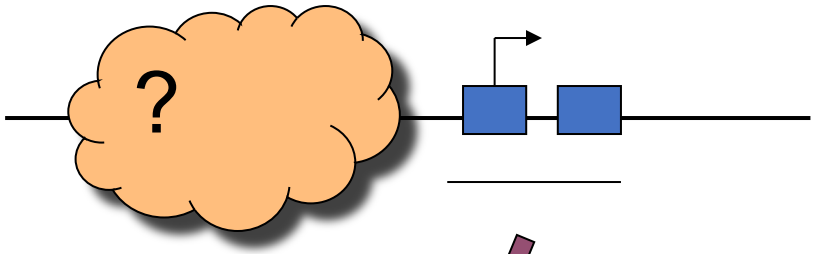
- **coreScore**: 1.0
- **score**: 1.0
- **siteModel**: V$AML1_01

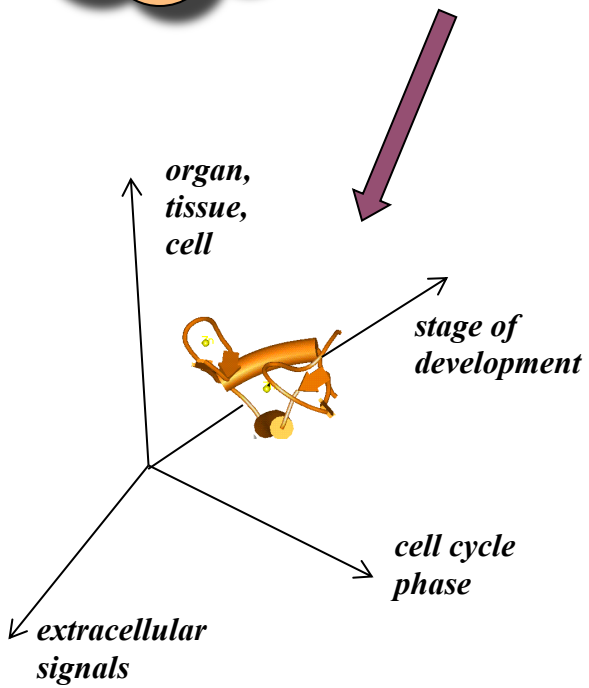
Model V$AML1_01
Binding element: AML1a
Threshold: 0.9555
Matrix: V$AML1_01
Matrix length: 6

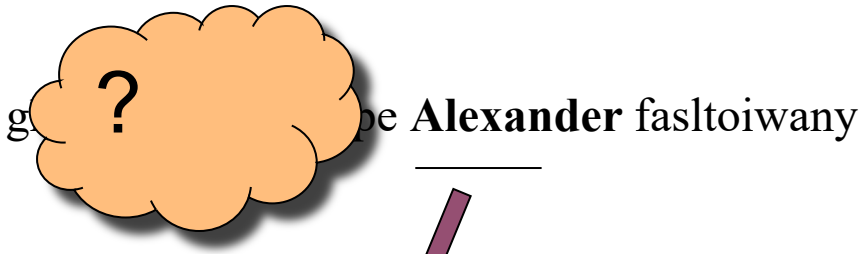
Non-coding DNA

we need to know “regulatory code”

(Epi-Genetic code)







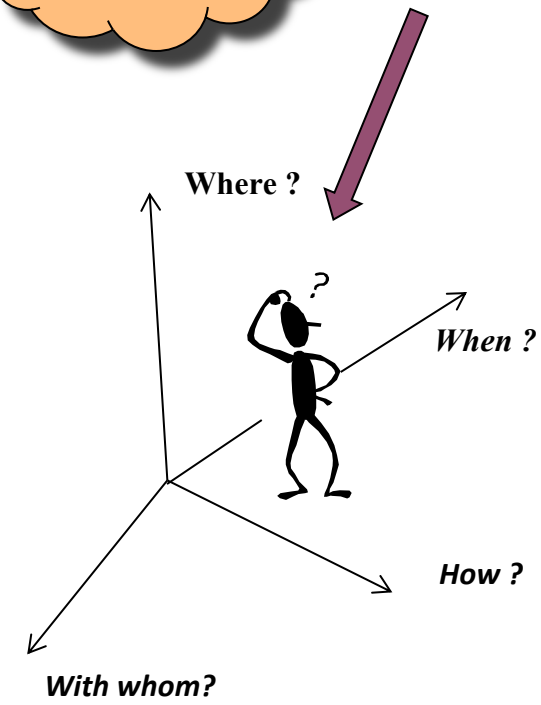


Таблица 17-2

Порядок символов в кодонах, полученный путем использования принципа связности, состава кодонов (см. табл. 17-1) и списка мутационных замен аминокислот в белках [Ратнер, 1966]

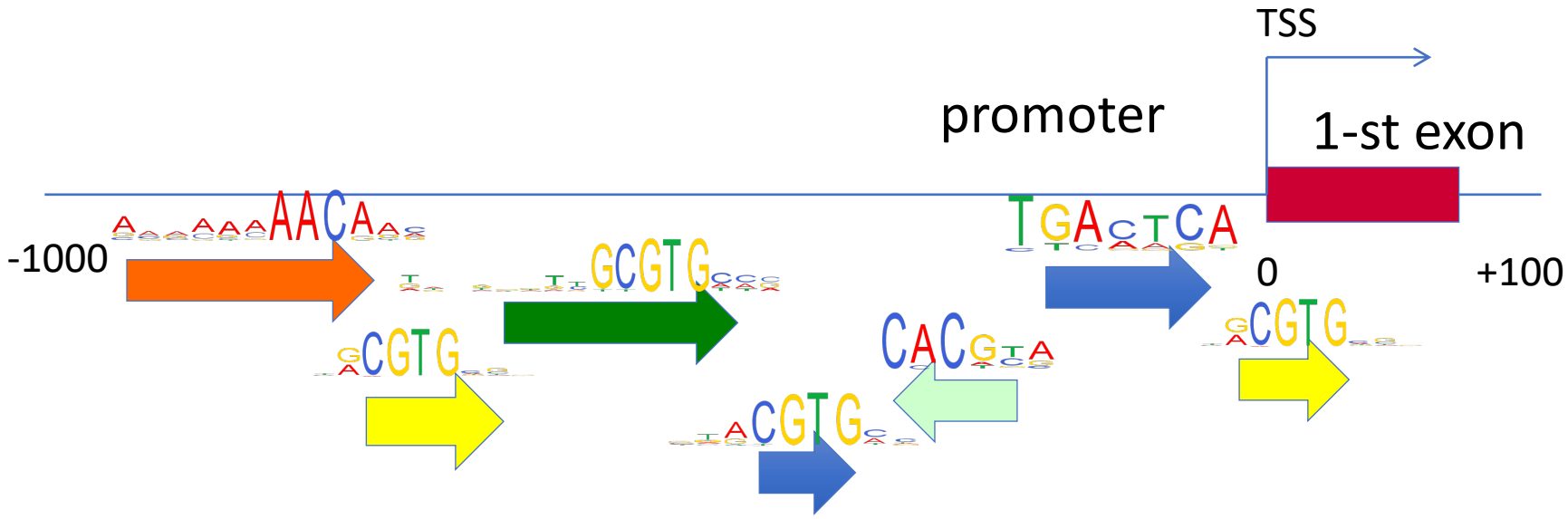
1↓2→	U	C	A	G	↓3
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC (Tyr)	UGC Ser	C
	UUA Leu	UCA Thr	UAA Lys	UGA ?	A
	UUG Leu	UCG [Ser]	UAG ?	UGG (Try)	G
C	CUU Leu	CCU Pro	CAU (His)	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Gln	CCA Pro	CAA Gln	CGA Arg	A
	CUG [Leu]	CCG ?	CAG ?	CGG ?	G
A	AUU Ile	ACU Asn	AAU Asn	AGU ?	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG [Thr]	AAG ?	AGG ?	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC [Val]	GCC Ala	GAC Asp	GGC Gly	C
	GUA Glu	GCA Ala	GAA Glu	GGA Gly	A
	GUG [Val]	GCG ?	GAG ?	GGG ?	G

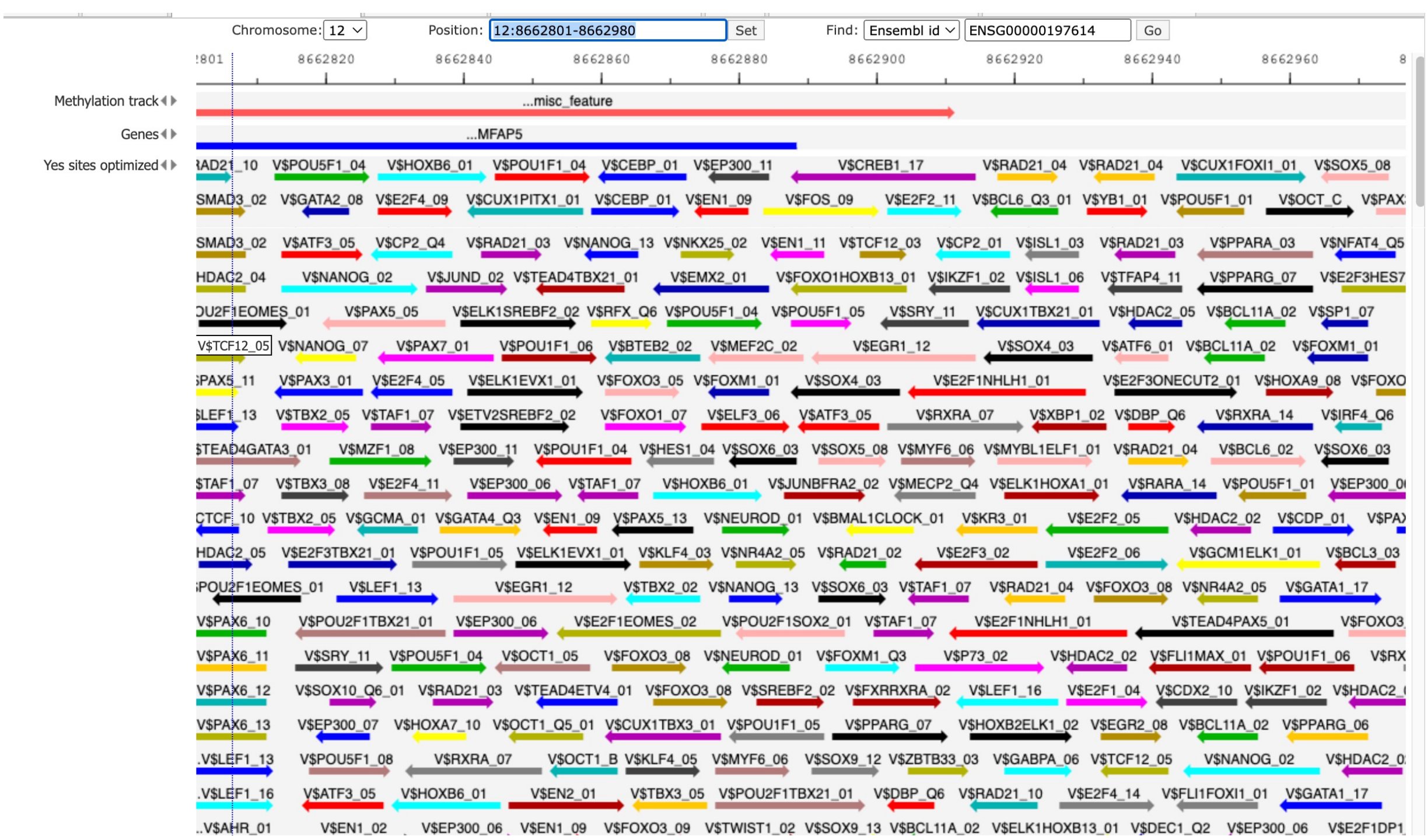
Примечание. Подчеркнуты кодоны, которые в дальнейшем оказались ошибочными. В круглых скобках — кодоны, в которых порядок символов установлен неоднозначно, в квадратных — кодоны, добавленные в ходе процедуры вывода для объяснения мутаций и связности. Из 64 кодонов 47 совпадают с кодом Ниренберга, 6 не совпадают, а 11 (см. табл. 17-1) не были определены по составу (из них 9 с *G* в третьих позициях кодонов).

Fundamental principle of genetic code:

Codons are not overlapping

Regulatory „instruction“ in promoter of an Up-regulated gene

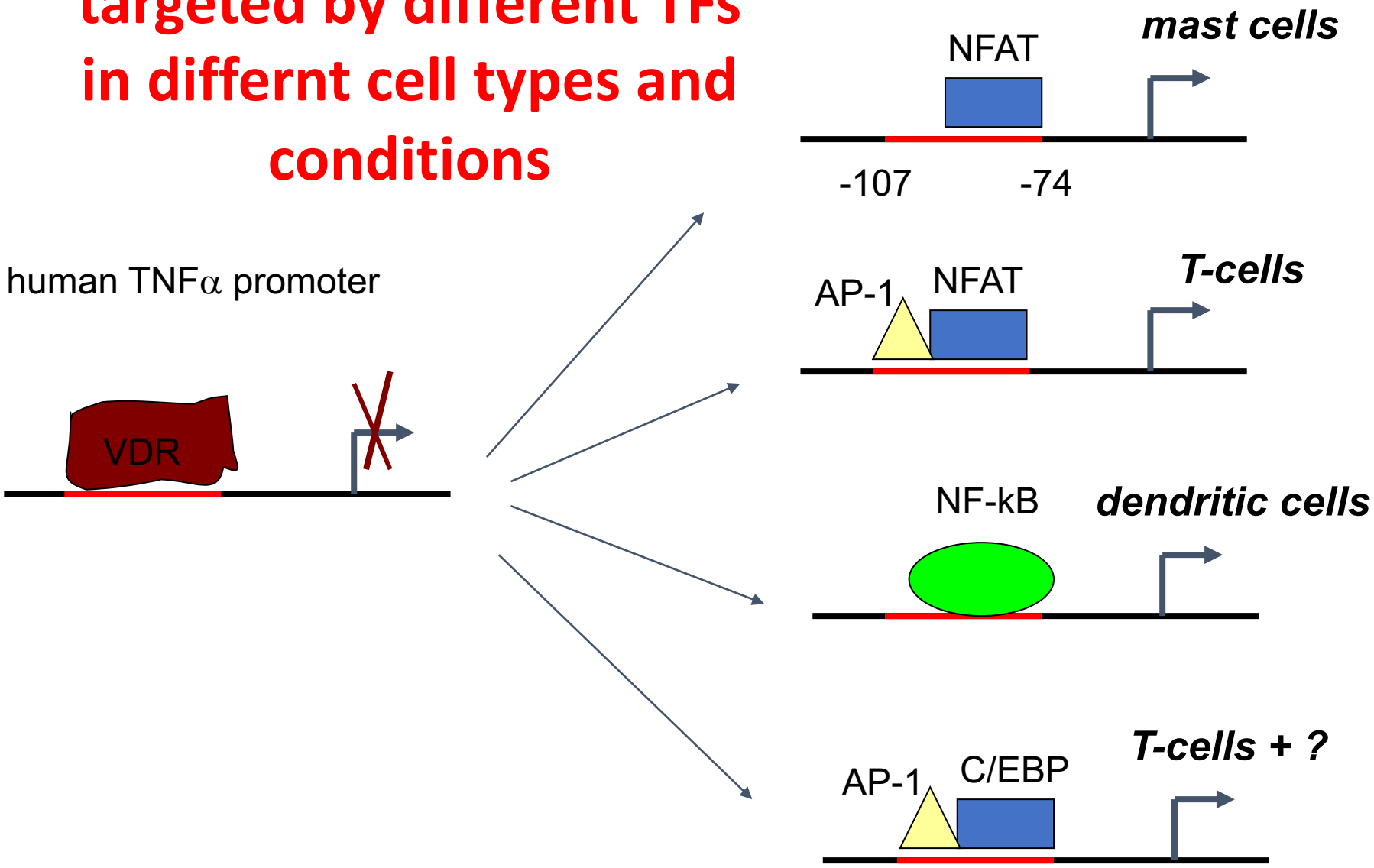



Chromosome: 12
Position: 12:8662801-8662980
Set
Find: Ensembl id
ENSG00000197614
Go
8662820
8662840
8662860
8662880
8662900
8662920
8662940
8662960
Methylation track
...misc_feature
Genes
...MFAP5
Yes sites optimized

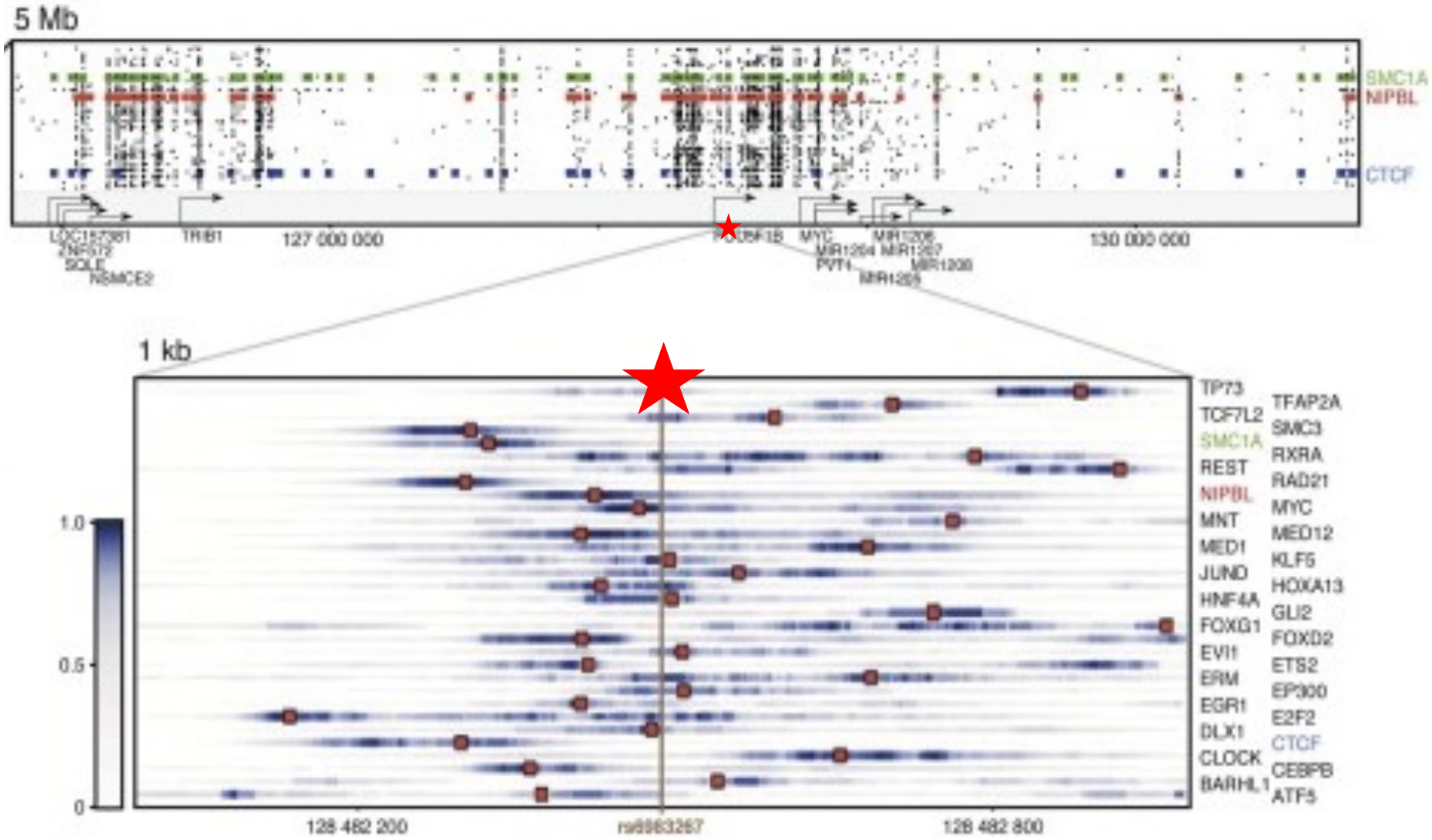
Promoter is a parking place

Parking in Italy

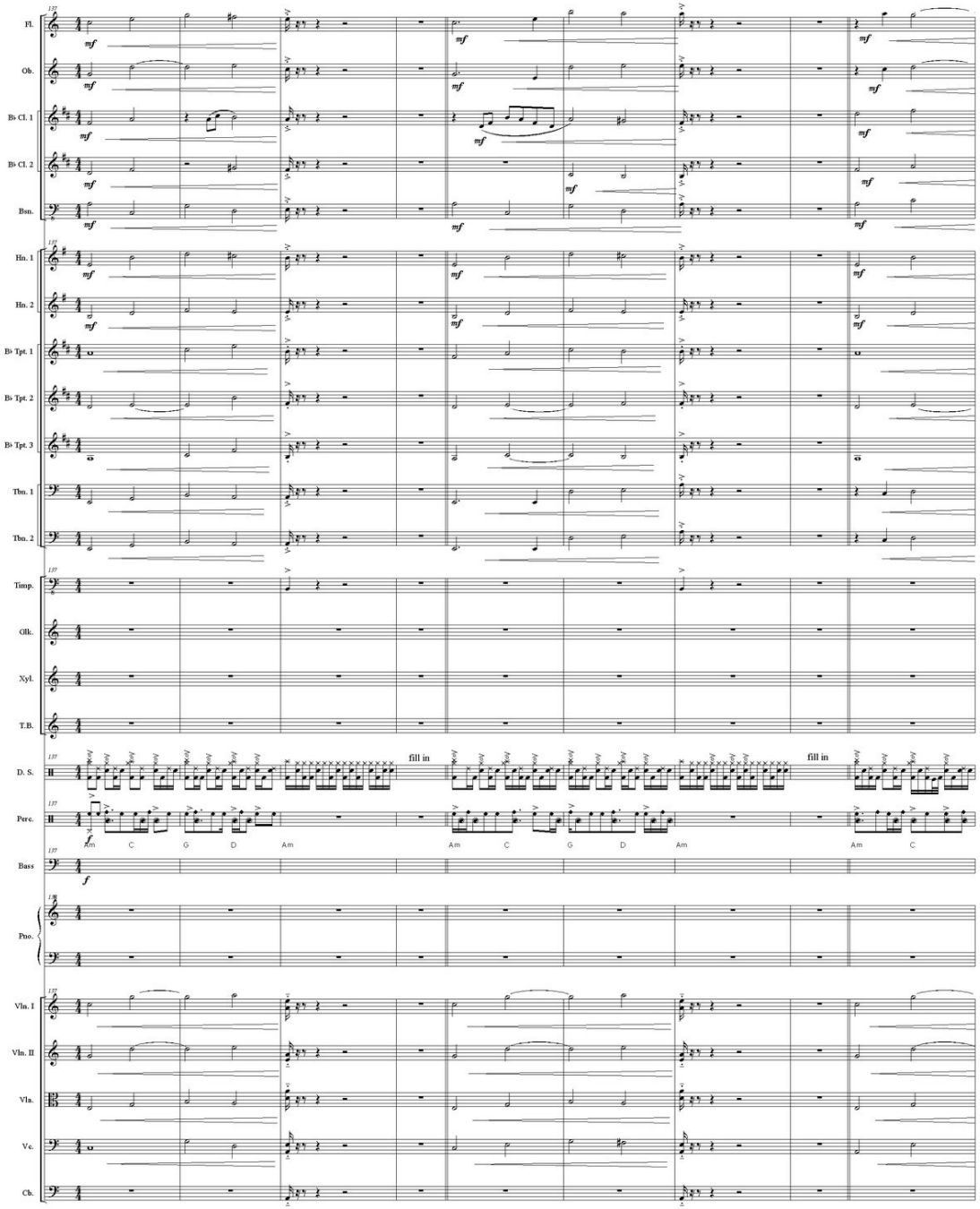
The same place in promoter is targeted by different TFs in differnt cell types and conditions



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



A score of a Symphony



Charme

Genetic code

George Gamow

Paradigm shift

Forget about biochemistry!
Nucleotides are just

Letters

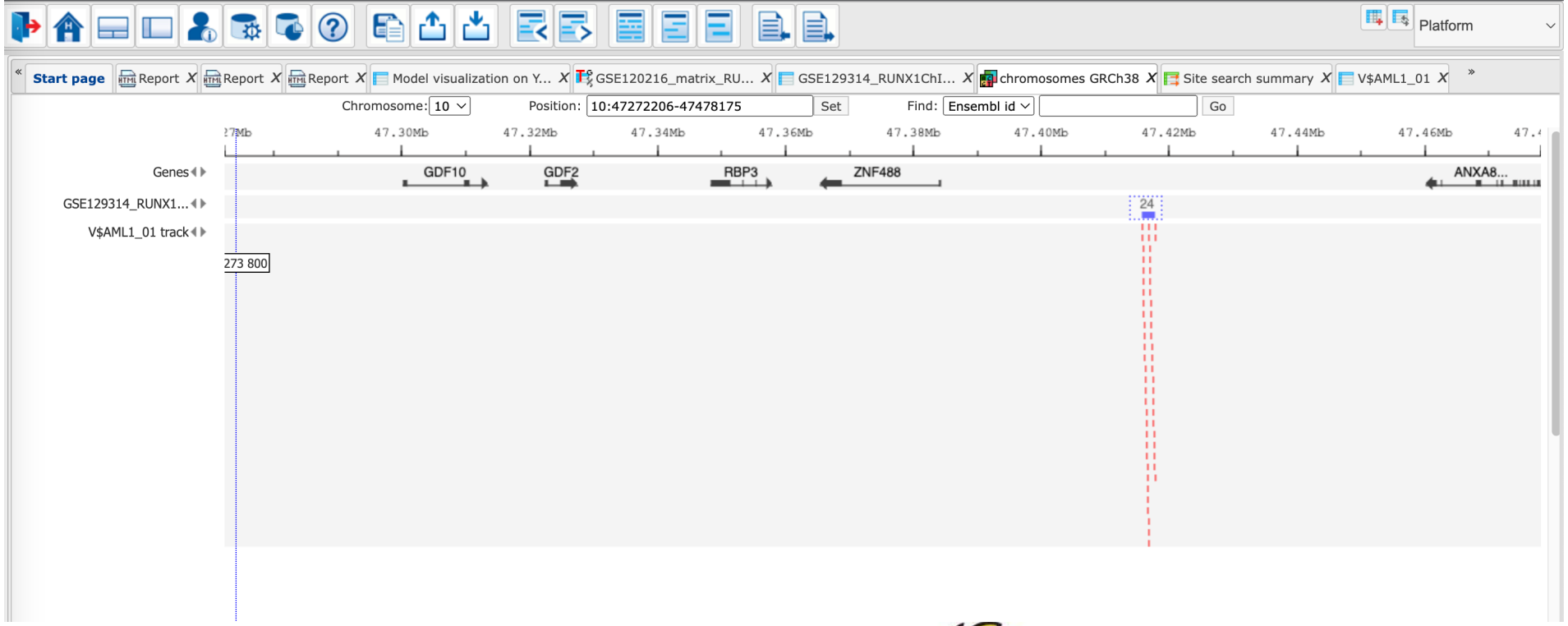
of an unknown language

Epi-Genetic code:

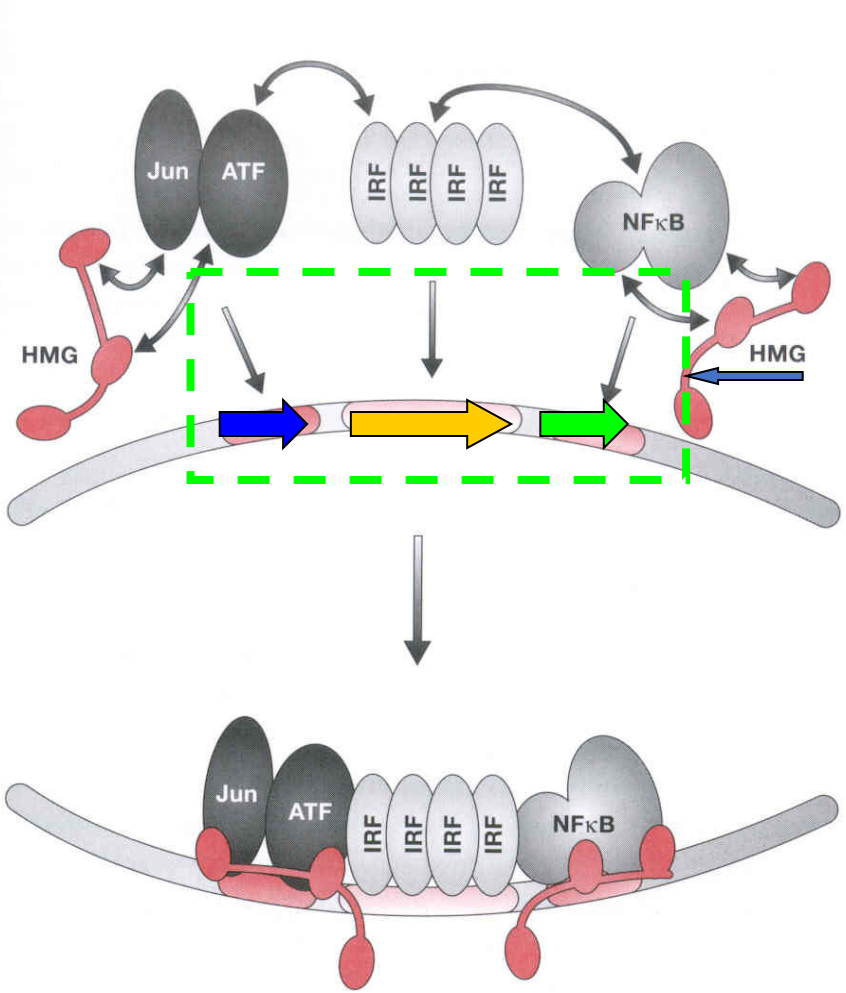
Forget about letters!

It is not text!

It is music!


Platform
Start page
Report
Report
Report
Model visualization on Y...
GSE120216_matrix_RU...
GSE129314_RUNX1ChI...
chromosomes GRCh38
Site search summary
V$AML1_01
Chromosome: 10
Position: 10:47272206-47478175
Set
Find: Ensembl id
Go
47.30Mb
47.32Mb
47.34Mb
47.36Mb
47.38Mb
47.40Mb
47.42Mb
47.44Mb
47.46Mb
Genes
GDF10
GDF2
RBP3
ZNF488
ANXA8...
GSE129314_RUNX1...
24
V$AML1_01 track
273 800

Composite Modules (CM)



Composite Module (CM)

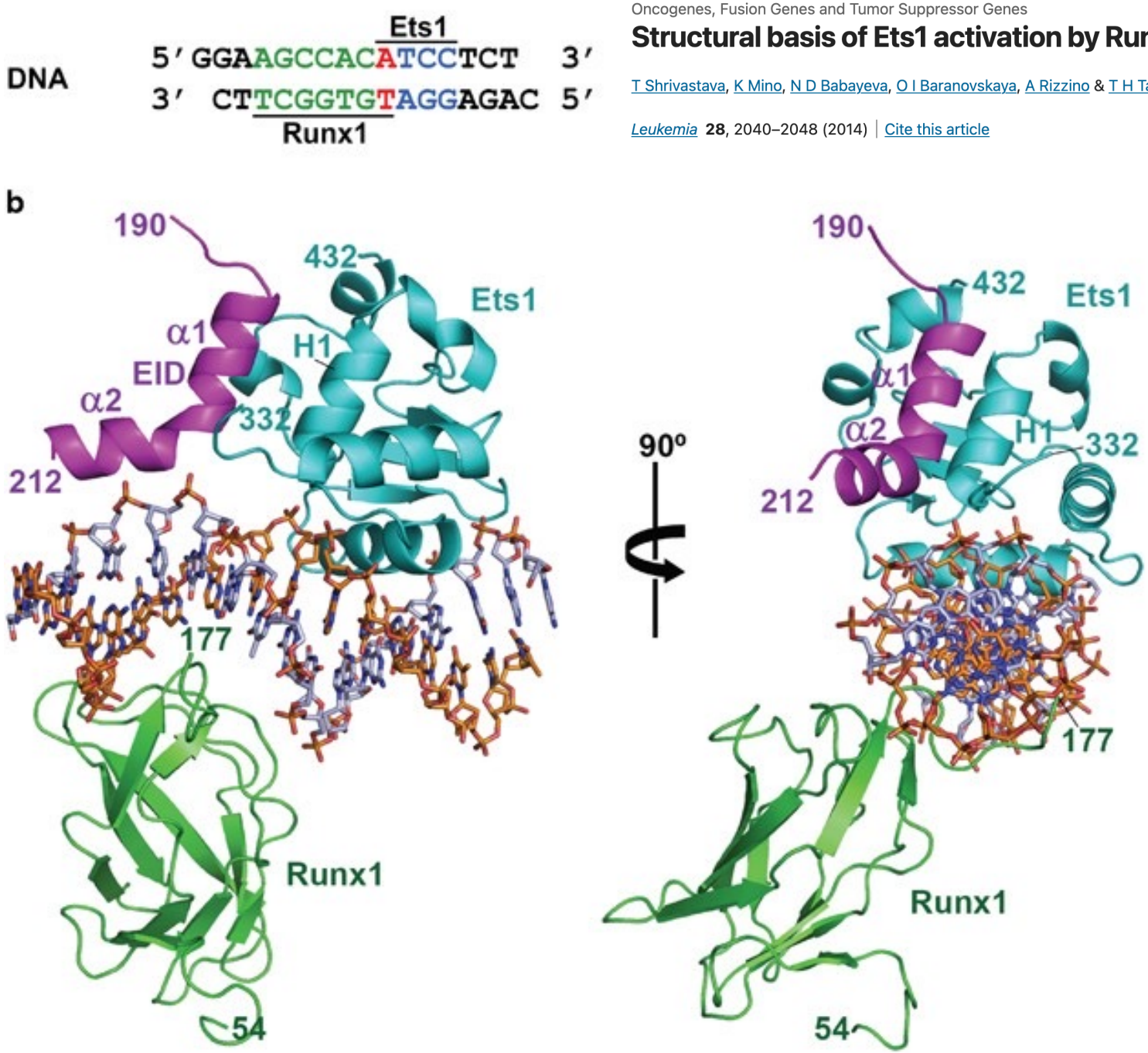
FIGURE 3.3. The human interferon-β enhanceosome. HMG represents HMGI/Y, a ubiquitous protein that binds cooperatively with the three activators. HMGI/Y both bends the DNA and contacts the activators. Each of the transcription factors shown is a member of a family of related activators. (Mark Ptashne,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

Oncogenes, Fusion Genes and Tumor Suppressor Genes

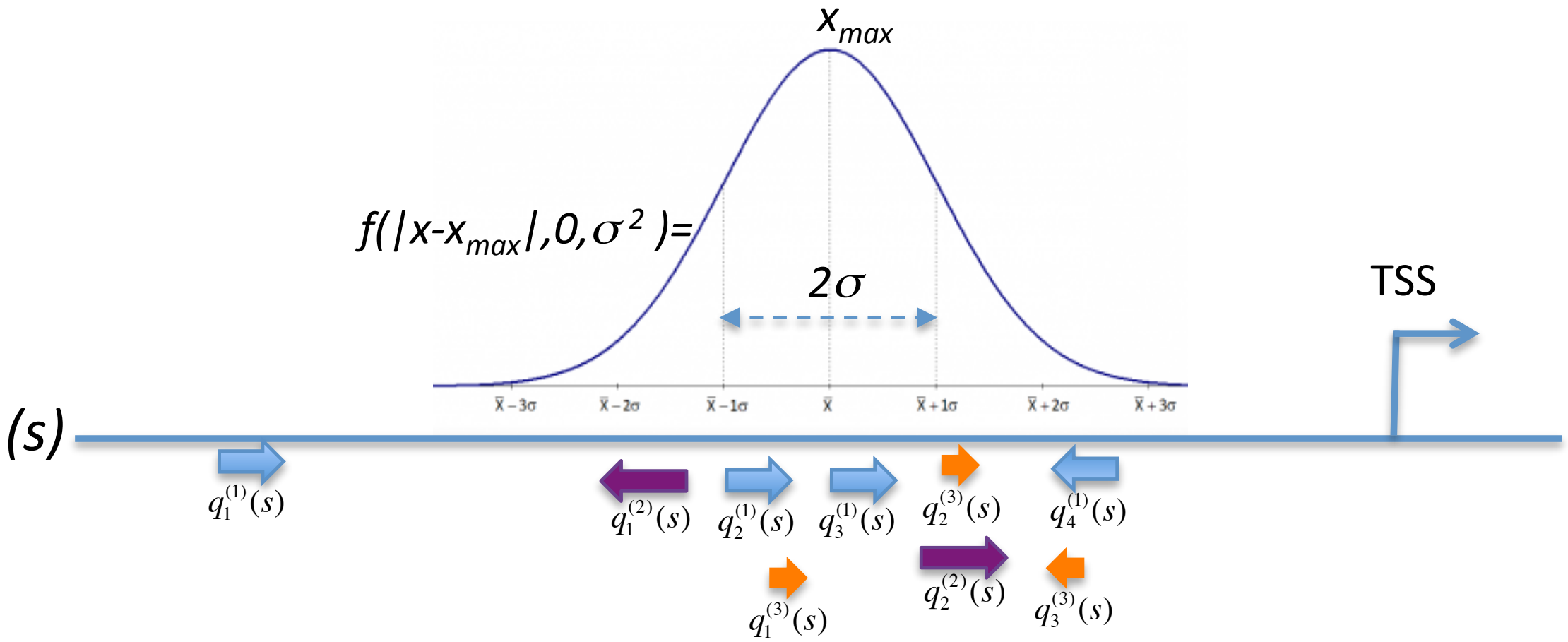
Structural basis of Ets1 activation by Runx1

T Shrivastava, K Mino, N D Babayeva, O I Baranovskaya, A Rizzino & T H Tahirov

Leukemia **28**, 2040–2048 (2014) | Cite this article



Composite model



Module 1:

V$STAT_01	V$GR_Q6_02	V$HAND1E47_01	V$ZNF462_01	V$TFE_Q6	V$RUNX2_05
0.89; N=2	0.92; N=2	0.92; N=2	0.93; N=3	0.96; N=3	0.92; N=2

Module width: 141

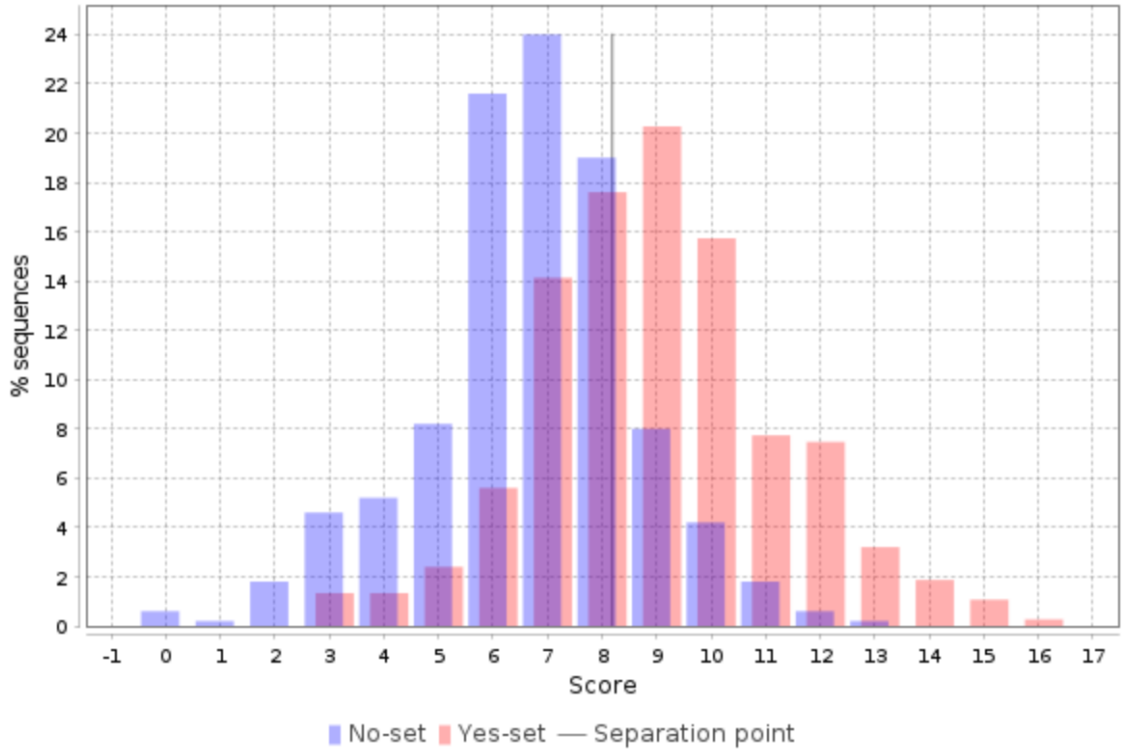
Module 2:

V$TFAP2A_11	V$OSF2_Q6	V$NR3C2_02	V$MZF1_Q5	V$STAT5A_02	V$NFATC3_06	V$AML_Q6
0.92; N=2	0.90; N=3	0.99; N=2	0.99; N=3	0.80; N=3	0.62; N=3	0.94; N=3

Module width: 167

Model score (-p*log10(pval)): 21.18
Wilcoxon p-value (pval): 1.94e-46
Penalty (p): 0.463
Average yes-set score: 8.96
Average no-set score: 6.71
AUC: 0.78
Separation point: 8.19
False-positive: 19.00%
False-negative: 37.60%
The AUC of the model achieves value significantly higher than expected for a random set of regulatory regions
Z-score = 3.59

UP-regulated genes



Module 1:

V$RUNX2_05	V$CTCF_08	V$RUNX3_06	V$PAX2_Q2	V$TEAD4TCF3_01	V$HAND1E47_01	V$GR_Q6_01
0.91; N=3	0.80; N=3	0.96; N=2	0.86; N=2	0.72; N=3	0.92; N=2	0.95; N=2

Module width: 140

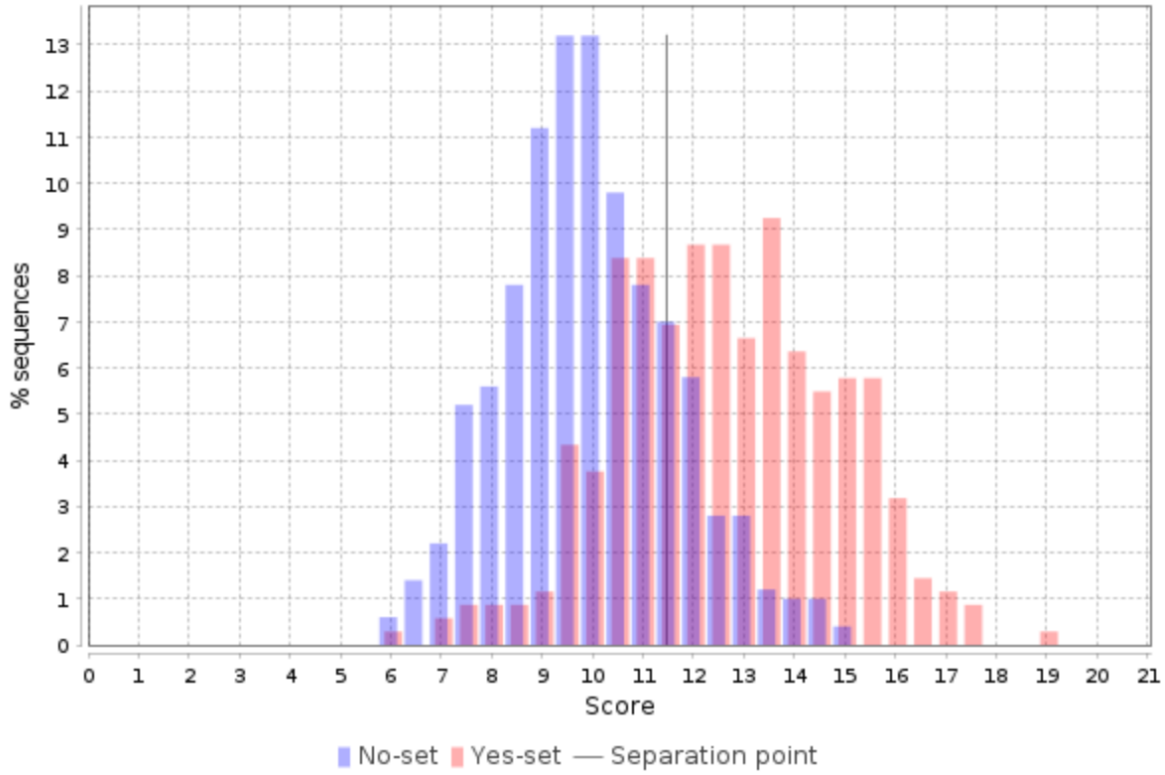
Module 2:

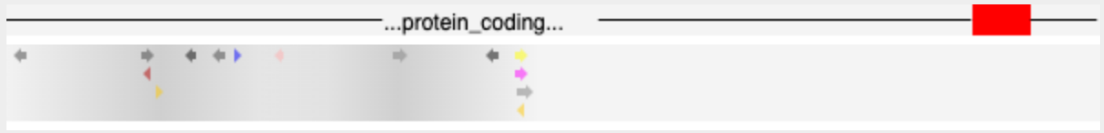
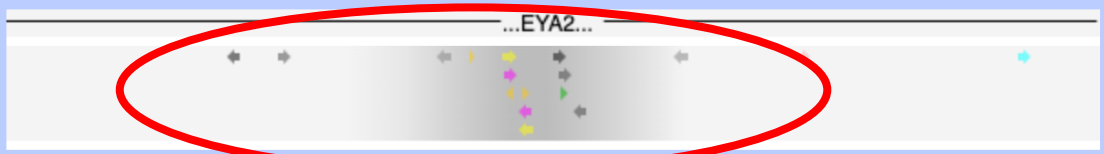
V$TFAP2A_09	V$SREBP_Q3	V$JUN_03	V$ZNF462_01	V$RUNX2_05	V$PAX6_Q2
0.90; N=3	0.83; N=1	0.73; N=3	0.93; N=2	0.91; N=3	0.68; N=3

Module width: 176

Model score (-p*log10(pval)): 27.62
Wilcoxon p-value (pval): 2.35e-60
Penalty (p): 0.463
Average yes-set score: 12.60
Average no-set score: 9.97
AUC: 0.83
Separation point: 11.47
False-positive: 17.60%
False-negative: 31.79%
The AUC of the model achieves value significantly higher than expected for a random set of regulatory regions
Z-score = 5.41

DN-regulated genes



ID	Ensembl IDs	Names	Model	Sites	Score	logFC	P.Value
30	ENSG00000047936	ROS1		V$NFATC3_06, V$MZF1_Q5, V$OSF2_Q6, V$GR_Q6_02, VTFE_Q6, VZNF462_01, V$HAND1E47_01, V$AML_Q6, V$RUNX2_05 (less)	8.28345	7.15284282543761	3.9480148660214E-8
130	ENSG00000064655	EYA2		V$ZNF462_01, V$TFE_Q6, V$TFAP2A_11, V$GR_Q6_02, V$OSF2_Q6, V$AML_Q6, V$RUNX2_05, V$NFATC3_06, V$MZF1_Q5 (less)	9.99289	5.94765294762649	5.04417038797566E-8
131	ENSG00000064655	EYA2		V$ZNF462_01, V$AML_Q6, V$OSF2_Q6, V$NFATC3_06, V$STAT_01, V$RUNX2_05, VGR_Q6_02, VTFE_Q6, V$HAND1E47_01 (less)	10.8041	5.94765294762649	5.04417038797566E-8
132	ENSG00000064655	EYA2		VGR_Q6_02, VNFATC3_06, V$HAND1E47_01, V$OSF2_Q6, VAML_Q6, VRUNX2_05, V$STAT_01, V$ZNF462_01, V$TFAP2A_11 (less)	11.56019	5.94765294762649	5.04417038797566E-8
				V$HAND1E47_01, V$ZNF462_01,			

Chromosome: 20 Position: 20:46070730-47980614 Set Find: Ensembl id ENSG00000064655 Go

)7Mb 46.20Mb 46.40Mb 46.60Mb 46.80Mb 47.00Mb 47.20Mb 47.40Mb 47.60Mb

Genes

COA5 lncRNA SLC35C2 MKRN7P SLC13A3 processed_pseudogene processed_pseudogene lncRNA LINC01754 NCOA3

IPL13P2 CDH22 lncRNA lncRNA TP53RK-DT RN7SKP33 GAPDHP54 MIR3616 lncRNA RNU6-497P lncRNA

CD40 ELMO2 ZNF840P lncRNA lncRNA processed_pseudogene ZMYND8 RPL35AP processed_pseudogene

ZNF663P OCSTAMP TP53RK EYA2 lncRNA RNU6-563P

ZNF334 SLC2A10

GSE129314_RUNX1...

01 7702 7703 7706 7707 7708 7710

7704 7709 7711

7705

yes track

M1

46 538 000


Chromosome: 20
Position: 20:46891490-47010858
Set
Find: Ensembl id
ENSG00000064655
Go
Genes
GSE129314_RUNX1...
yes track
EYA2...
lncRNA
1765
847
5297
M1
M2
M2
46 894 400
...EYA2...
5297
M2 M1
V$GR_Q6_02
V$HAND1E47_01
V$RUNX2_05
V$STAT_01
V$HAND1E47_01
V$ZNF462_01
V$NFATC3_06
V$OSF2_Q6
V$RUNX2_05
V$OSF2_Q6
V$NFATC3_06
V$AML_Q6
V$NFATC3_06
V$OSF2_Q6
V$AML_Q6
V$GR_Q6_02
score: 0.969094
sequenceName: 132
siteModel: V$RUNX2_05
subModel: M1
weight: 0.957369
Model V$RUNX2_05
Binding element: AML3
Threshold: 0.9
Matrix: V$RUNX2_05
Matrix length: 14
TGTGG

Introduction

Description

Eyes absent homolog 2, a transcription coactivator that acts in TGF-beta signaling, epithelial to mesenchymal transition, apoptosis, and somitogenesis, may play a role in eye development; SNPs correlate with non-small-cell lung carcinoma

Gene symbol

EYA2

Synonyms

EAB1; EYA2; MGC10614; RP5-890O15.2; Eyes absent homolog 2; eyes absent homolog 2 (Drosophila); EYA transcriptional coactivator and phosphatase 2

Gene Ontology what is this?

Molecular function

G-protein alpha-subunit binding [E], protein binding [E], transcription coactivator activity [E] details

Biological process

eye development [S], mesodermal cell fate specification [E], positive regulation of epithelial to mesenchymal transition [E], somitogenesis [P]... details

Cellular component

cytoplasm [E], cytosol [Y], nucleus [E] details

Hierachy of orthologous relationships for this locus

+ View orthologous relationships

Biomarker Associations what is this?

Diseases associated with EYA2 (9 entries)

Show All entries Search:

		Type of Association				Indication		
Disease details-all	Significance	Causal 2 associations	Correlative 13 associations	Preventative	Negative	Disease Mechanism 2 associations	Prognosis 1 associations	Therapeutic Target
Ovarian Neoplasms	6 associations	1 associations	5 associations			1 associations	1 associations	
Carcinoma, Non-Small-Cell Lung	2 associations		2 associations					
Breast Neoplasms	1 associations		1 associations					
Rectal Neoplasms	1 associations		1 associations					
Colorectal Neoplasms	1 associations		1 associations					
Nerve Sheath Neoplasms	1 associations	1 associations				1 associations		
Lung Neoplasms	1 associations		1 associations					
Papillomavirus Infections	1 associations		1 associations					
Thyroid Neoplasms	1 associations		1 associations					

Showing 1 to 9 of 9 entries First Previous 1 Next Last

Evidence	Description
E	Experimentally determined
S	Predicted by sequence similarity
P	Predicted by analysis other than sequence similarity
K, BioKnowledge Transfer	Predicted by BioKnowledge Transfer by BIOBASE
see	Indicates that a reference mentions and cites a finding, but does not demonstrate it directly
hpa	Data imported from the Human Protein Atlas version 20 available at www.proteinatlas.org
hpa evidence codes:	W - uncertain, X - approved, Y - supported, Z - enhanced

Biomarkers associated with Correlative associations for Breast Neoplasms

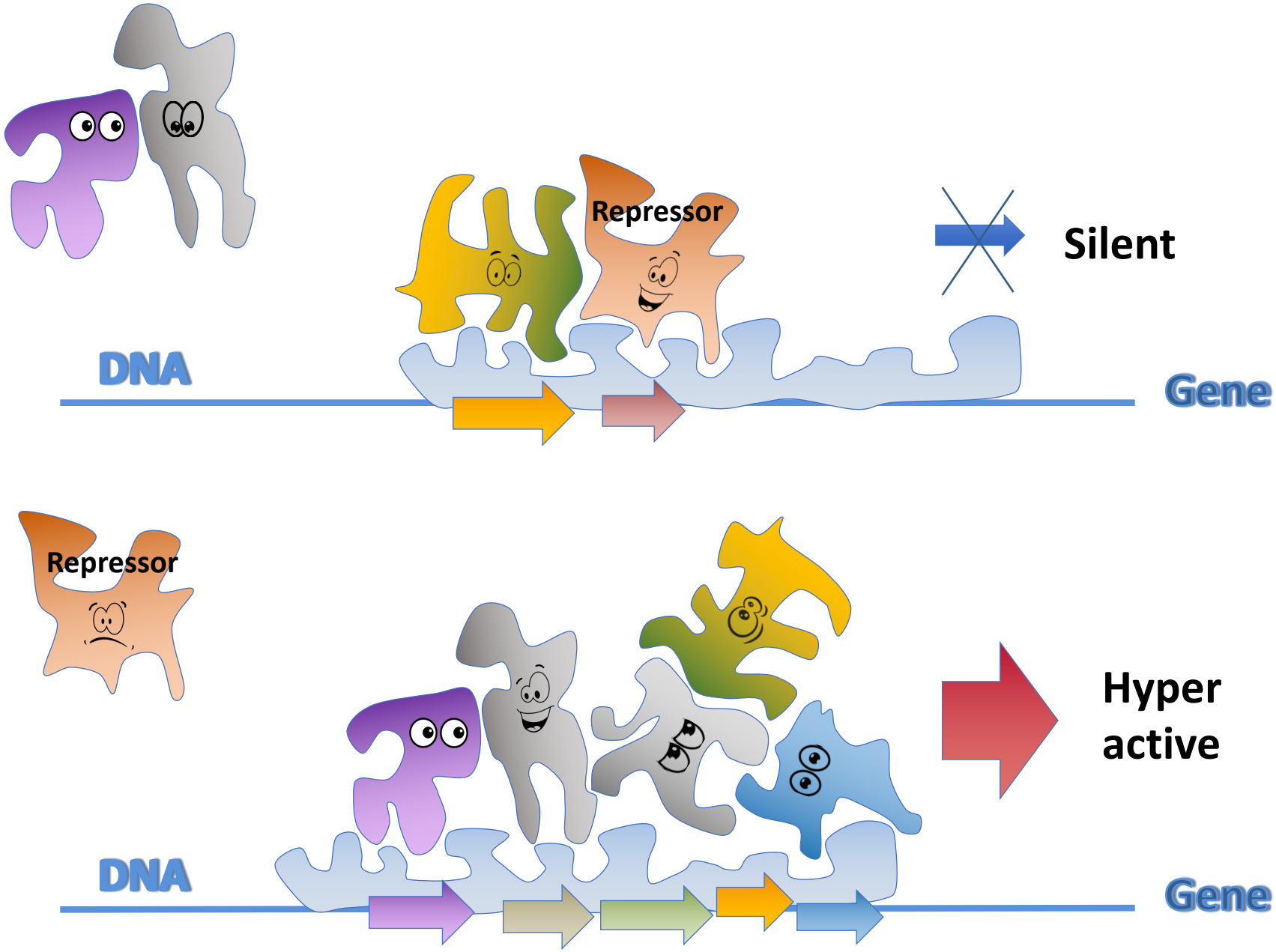
Show 5 entries Search:

Gene/Protein	Molecule	Alteration	Relationship	Sub-type / Linked disorders	Biological Process	Drug	Disease	Cell/Tissue Specificity	Indication	Annotation & Reference
EAB1	DNA	promoter hypermethylation	correlates with				Breast Neoplasms			hypermethylation of the EYA2 promoter correlates with breast neoplasms 1, E

Showing 1 to 1 of 1 entries First Previous 1 Next Last

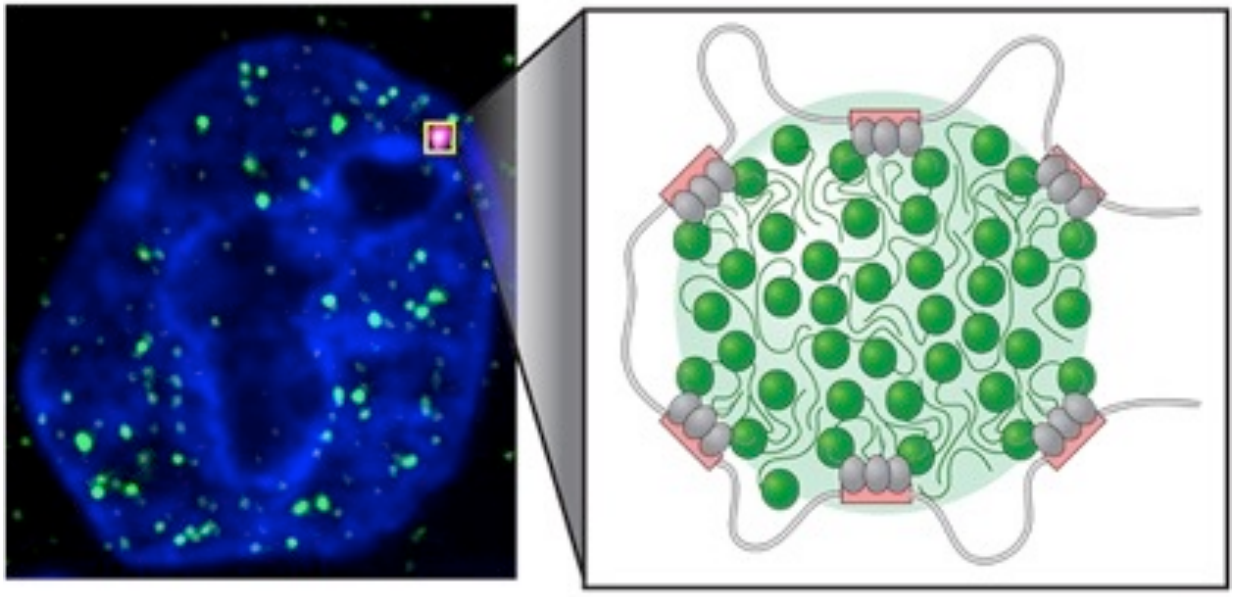
PM000620159	Runx1(h)	20:46988315..46988700	VCaP + DHT 24h	V$AML1_Q4_01	-1520	1.000	0.989	GCACCACAC	GEO Data 23193258 25228652 29126285
PM000620158	Runx2(h)	20:46979042..46979226	iMSC3	V$AML3_Q3	-130	1.000	0.894	GGGGTGGTGA	ArrayExpress Data 25361974 29126285
PM000620158	Runx1(h)	20:46979450..46979628	CD34+ cells, adult	V$AML1_Q4_01	268	1.000	0.994	AAACCACAA	GEO Data 23193258 26766440 29126285
PM000915834	Runx2(h)	20:46895075..46895653	iMSC3	V$AML3_Q3	269	0.870	0.864	CCACCCCAGG	ArrayExpress Data 25361974 29126285
PM000620159	Runx1(h)	20:46990274..46991206	VCaP + DHT 24h	V$AML1_Q4_01	379	0.820	0.842	CCACCACGG	GEO Data 23193258 25228652 29126285
PM000620159	AR(h)	20:46990444..46990895	VCaP + DHT 24h + SHRUNX1	V$AR_04	650	1.000	0.704	AGAACAATTCCAGCT	GEO Data 23193258 25228652 29126285
PM000915834	Runx1(h)	20:46895504..46895866	VCaP + DHT 24h	V$AML1_Q4_01	793	0.771	0.806	CTCTGGTTC	GEO Data 23193258 25228652

It’s a fuzzy puzzle!

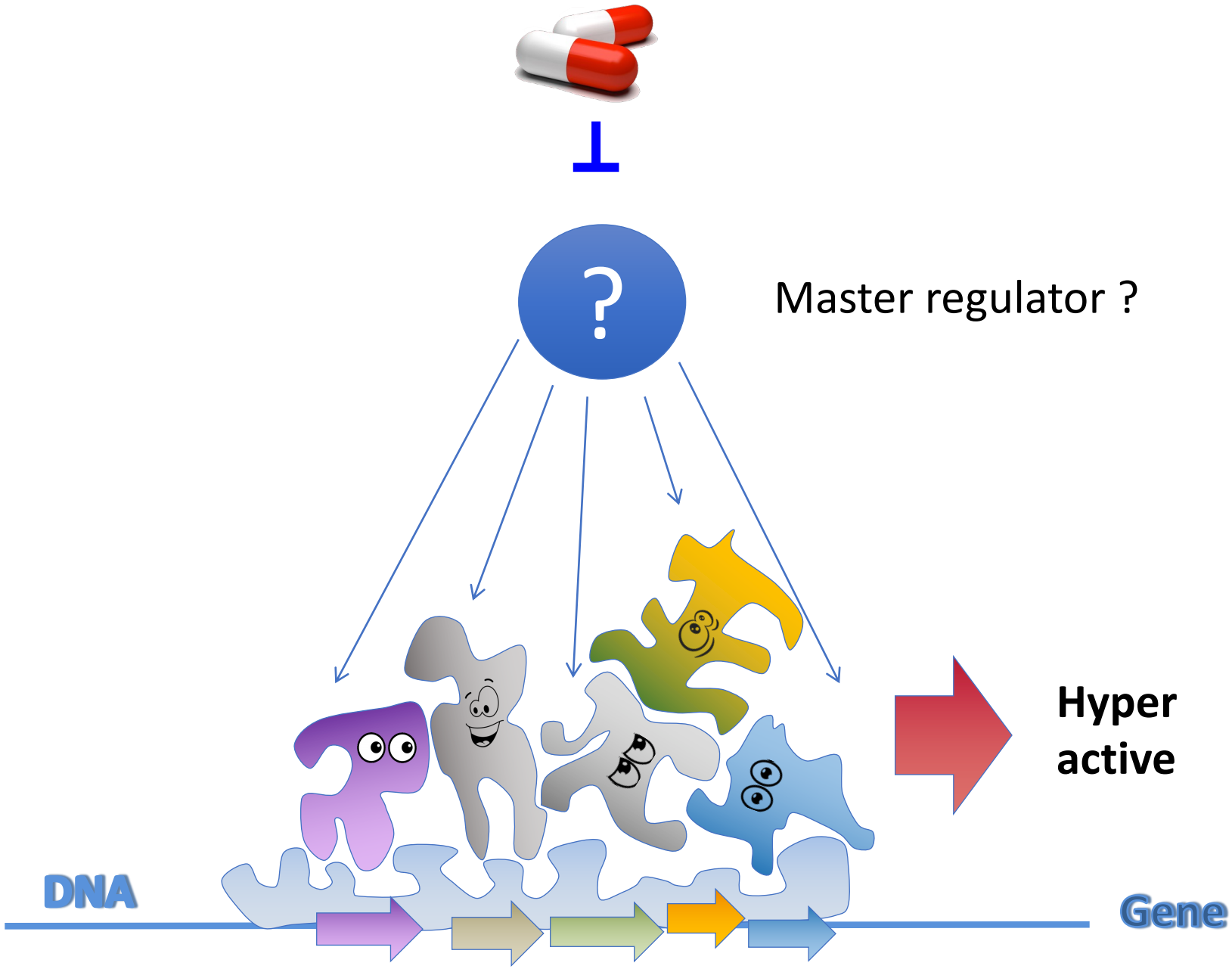


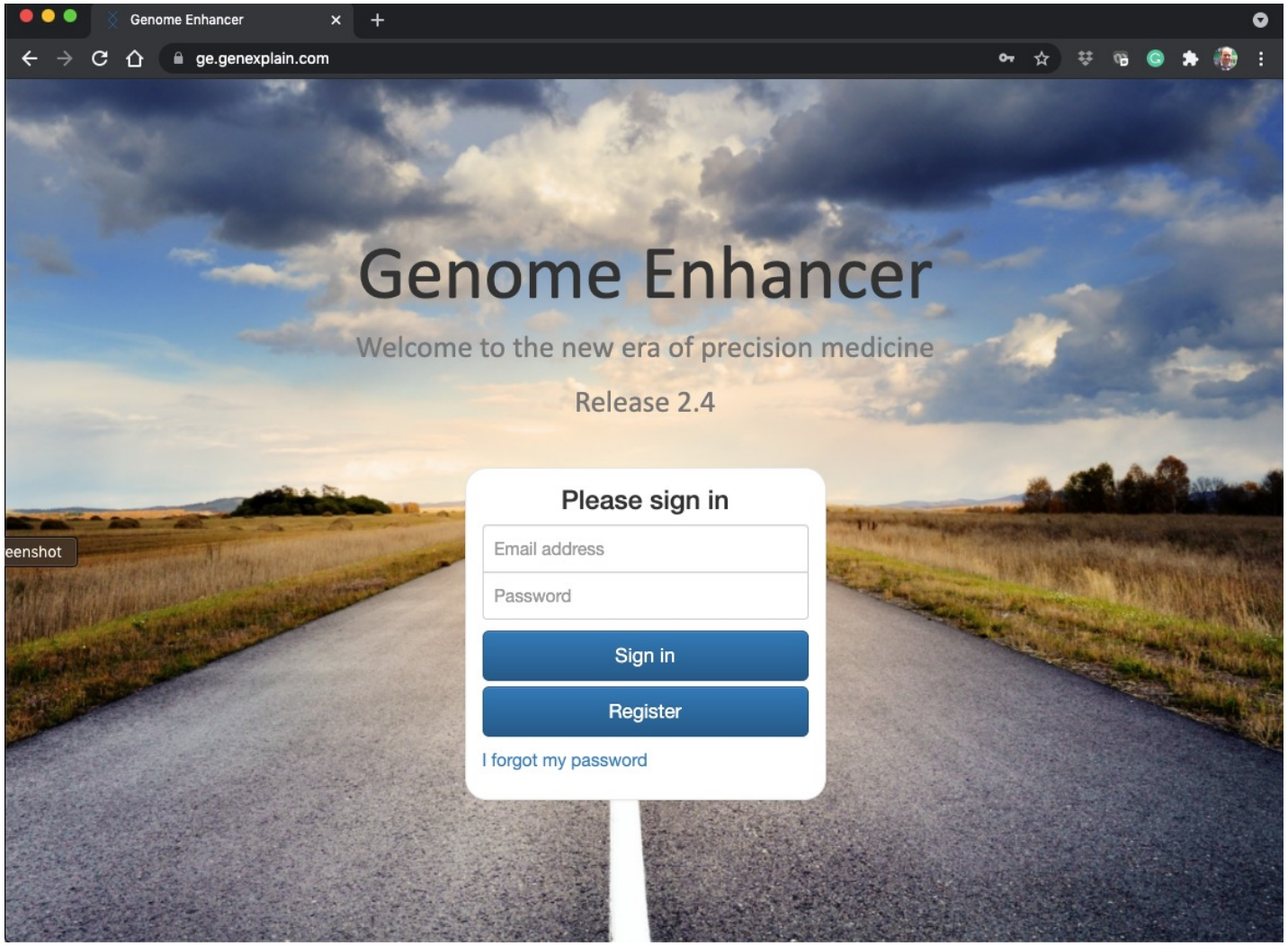
Phase-separated condensates

TFs have unstable and dynamic protein structure that promotes formation of such condensates.



Richard Young and collegues at Massachusetts Institute of Technology (MIT).


?
Master regulator ?
Hyper
active
DNA
Gene

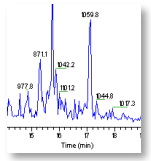

Genome Enhancer
ge.genexplain.com
Genome Enhancer
Welcome to the new era of precision medicine
Release 2.4
Please sign in
Email address
Password
Sign in
Register
I forgot my password

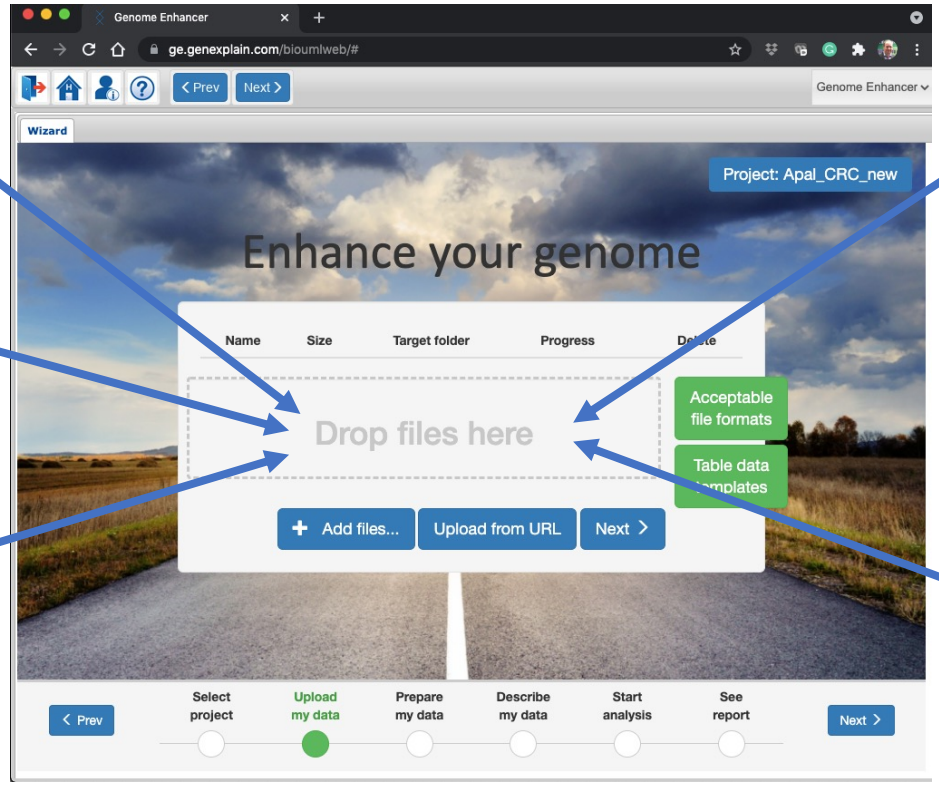
Multi-omics data input

Transcriptomics

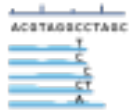
Epigenomics

Proteomics

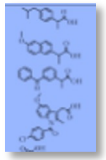




Genomics



Metabolomics



Seque

ITGI
targ
cont
tran
expi

Alexan
geneX
alexan
Data rec

Genome

Abstra

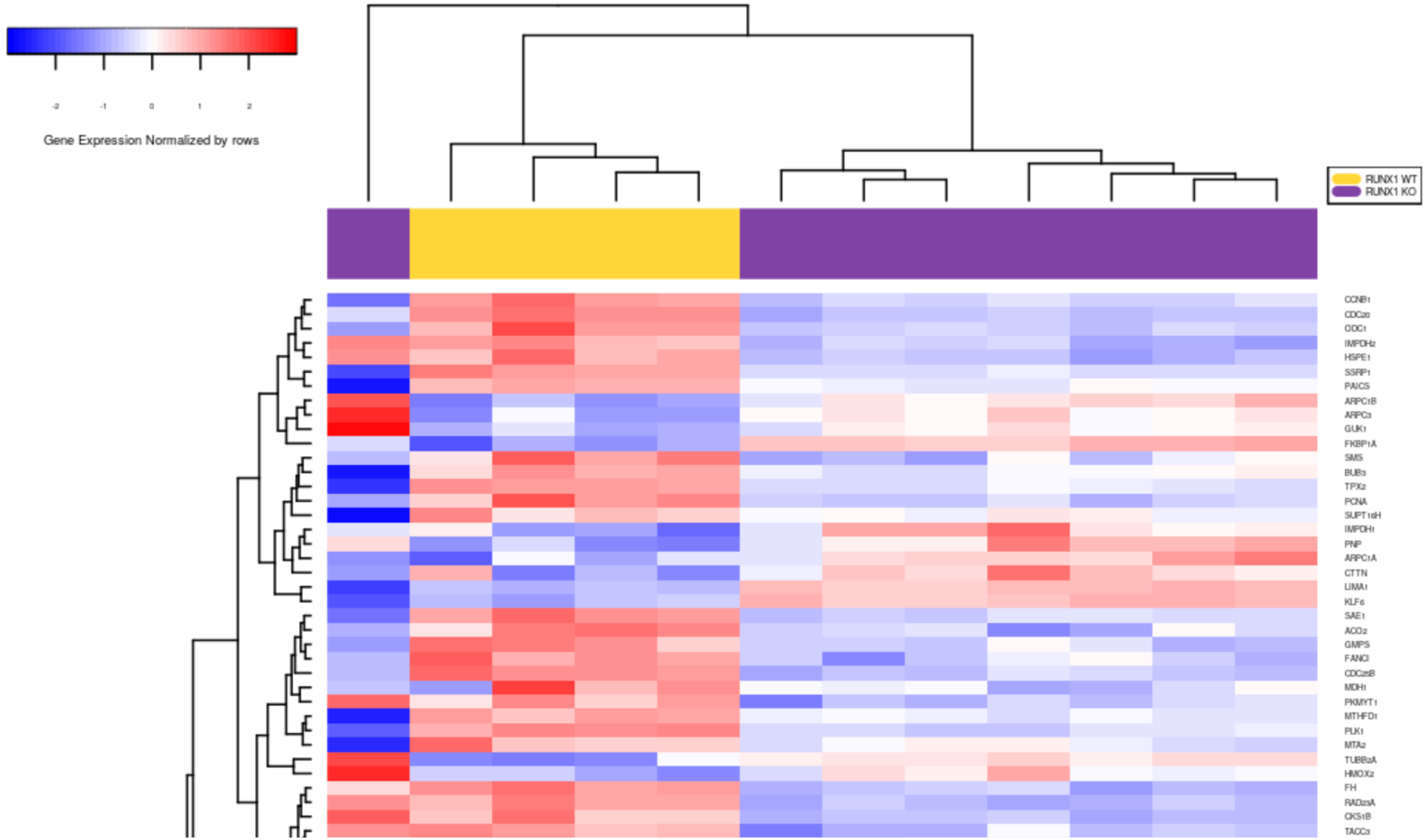
In the p
that cor
Neoplas
that go
transcrip

activities

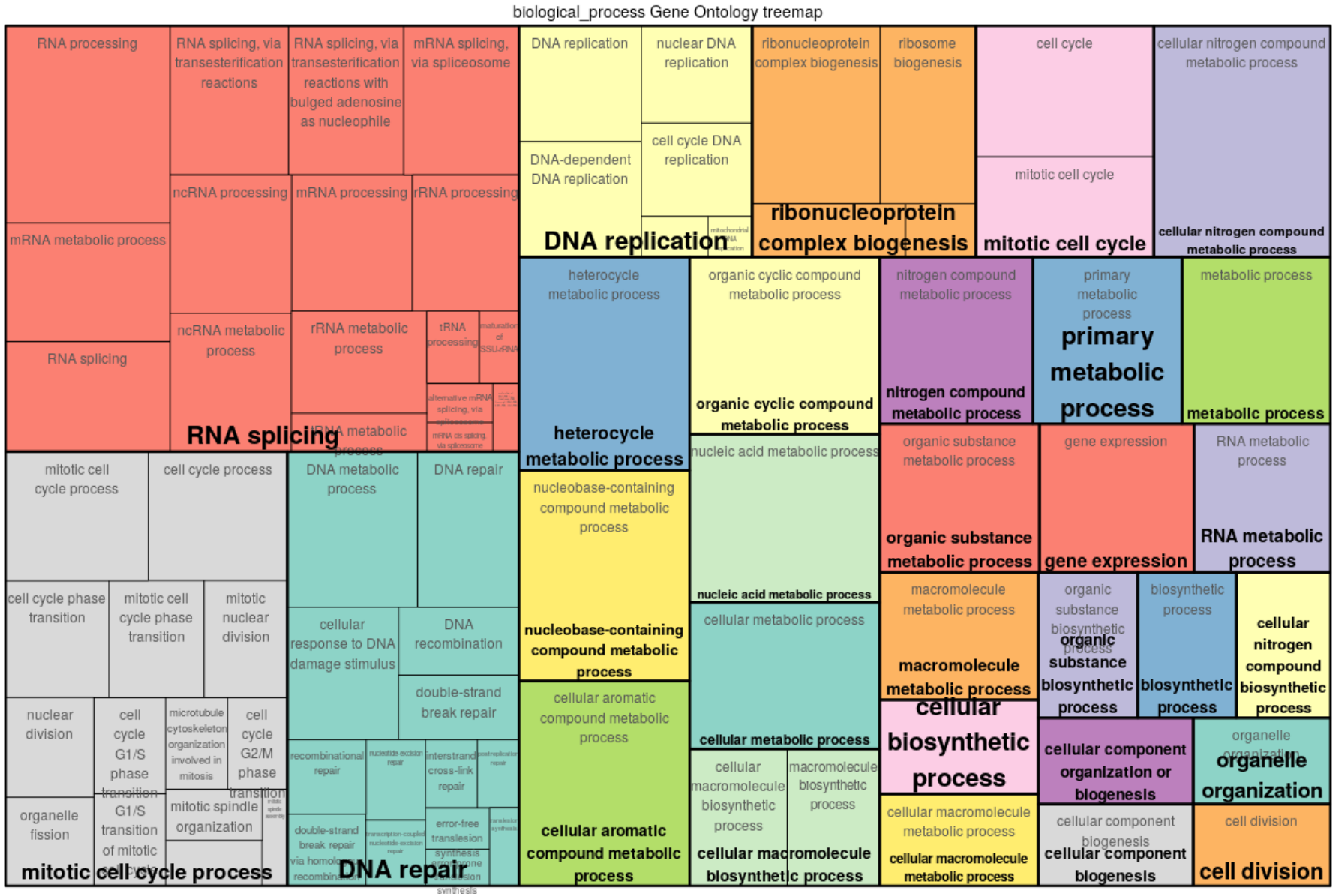
3. Results

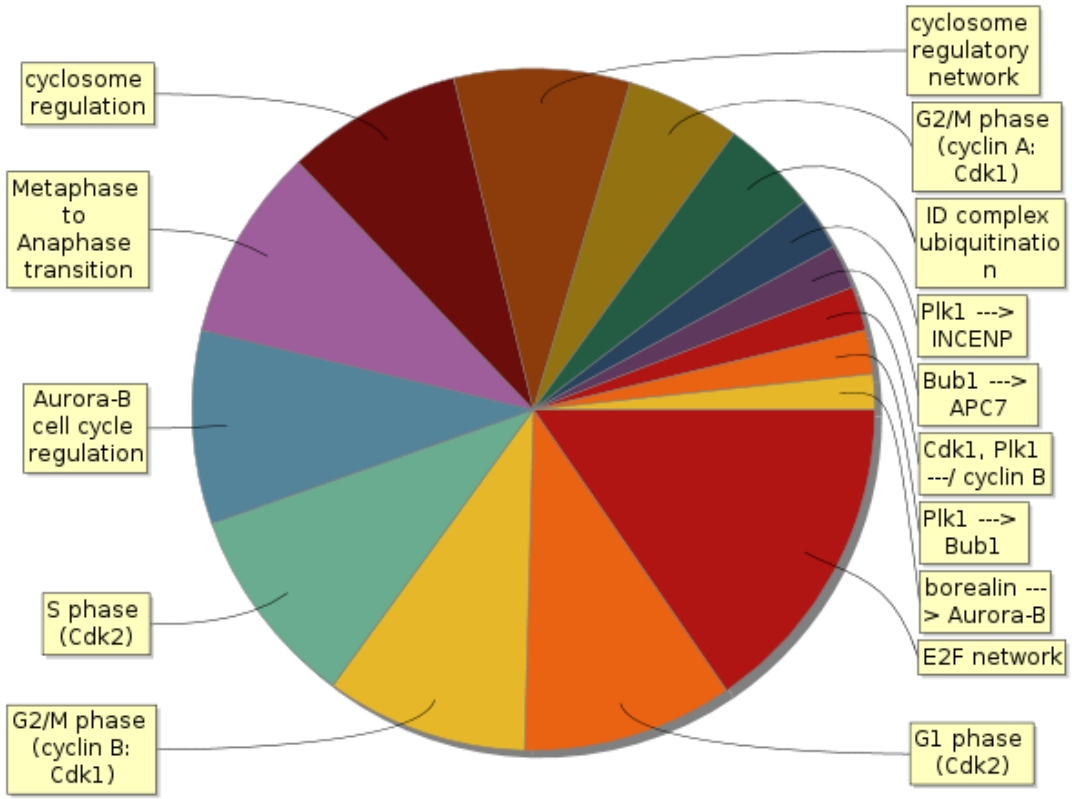
Heatmap of differentially expressed genes in RUNX1 WT vs. RUNX1 KO

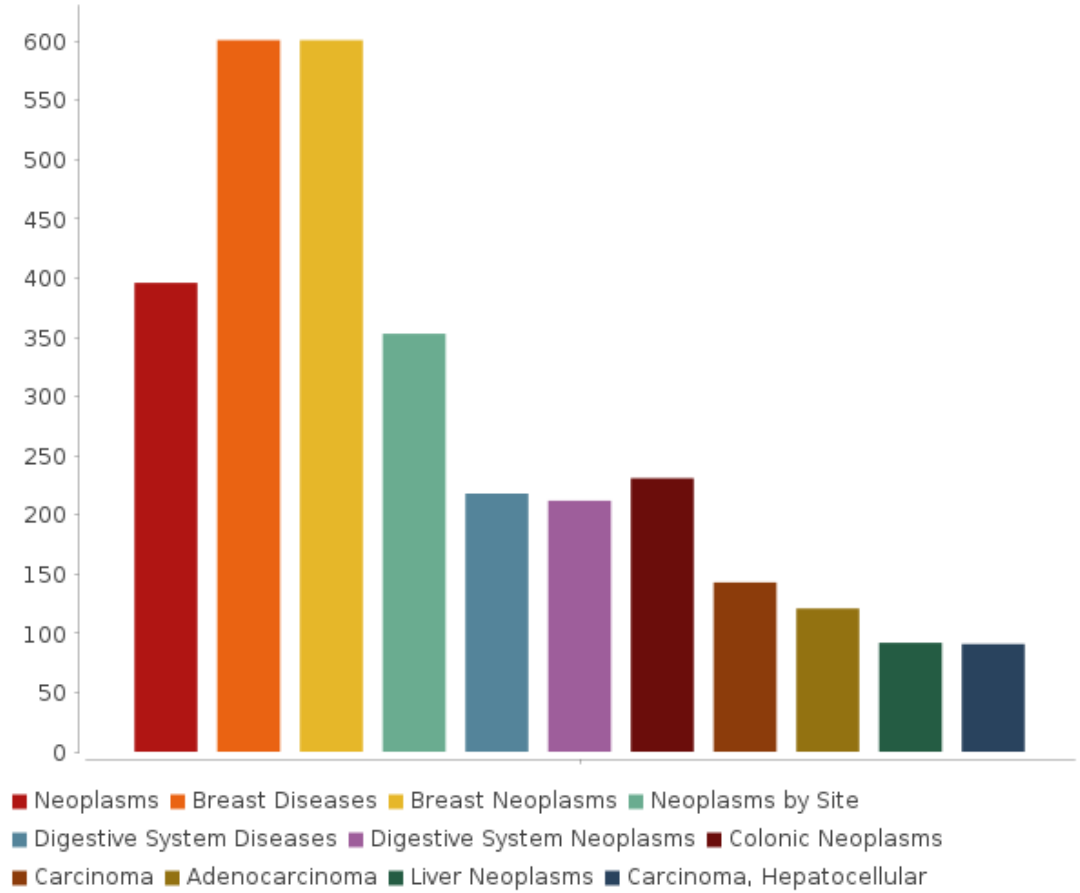
A heatmap of all differentially expressed genes playing a potential regulatory role in the system (enriched in TRANSPATH® pathways) is presented in Figure 2.

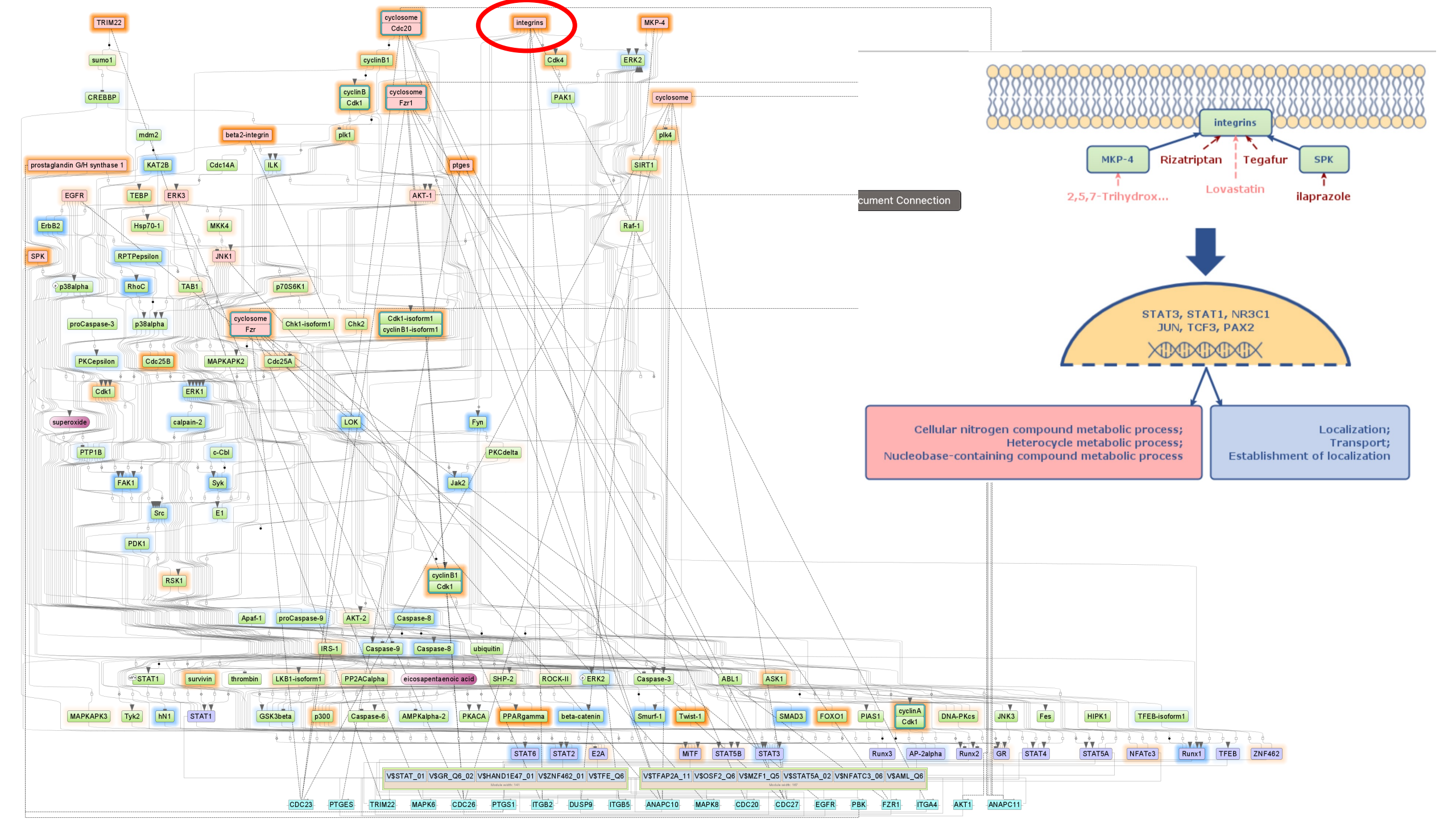


UP


biological_process Gene Ontology treemap
RNA processing
RNA splicing, via transesterification reactions
RNA splicing, via transesterification reactions with bulged adenosine as nucleophile
mRNA splicing, via spliceosome
mRNA metabolic process
ncRNA processing
mRNA processing
rRNA processing
RNA splicing
ncRNA metabolic process
rRNA metabolic process
tRNA processing
maturation of SSU-rRNA
alternative mRNA splicing, via spliceosome
mRNA cis splicing, via spliceosome
RNA splicing
DNA replication
nuclear DNA replication
DNA-dependent DNA replication
cell cycle DNA replication
ribonucleoprotein complex biogenesis
ribosome biogenesis
ribonucleoprotein complex biogenesis
cell cycle
mitotic cell cycle
mitotic cell cycle
cellular nitrogen compound metabolic process
cellular nitrogen compound metabolic process
heterocycle metabolic process
heterocycle metabolic process
organic cyclic compound metabolic process
organic cyclic compound metabolic process
nitrogen compound metabolic process
nitrogen compound metabolic process
primary metabolic process
primary metabolic process
metabolic process
metabolic process
mitotic cell cycle process
cell cycle process
cell cycle phase transition
mitotic cell cycle phase transition
mitotic nuclear division
nuclear division
cell cycle G1/S phase transition
microtubule cytoskeleton organization involved in mitosis
cell cycle G2/M phase transition
organelle fission
G1/S transition of mitotic cell cycle
mitotic spindle organization
mitotic cell cycle process
DNA metabolic process
DNA repair
cellular response to DNA damage stimulus
DNA recombination
double-strand break repair
recombinational repair
nucleotide-excision repair
interstrand cross-link repair
double-strand break repair via homologous recombination
transcription-coupled nucleotide-excision repair
error-free translesion synthesis
DNA repair
nucleobase-containing compound metabolic process
nucleobase-containing compound metabolic process
nucleic acid metabolic process
nucleic acid metabolic process
cellular metabolic process
cellular metabolic process
cellular aromatic compound metabolic process
cellular aromatic compound metabolic process
cellular macromolecule biosynthetic process
macromolecule biosynthetic process
cellular macromolecule biosynthetic process
organic substance metabolic process
organic substance metabolic process
gene expression
gene expression
RNA metabolic process
RNA metabolic process
macromolecule metabolic process
macromolecule metabolic process
cellular biosynthetic process
cellular macromolecule metabolic process
cellular macromolecule metabolic process
organic substance biosynthetic process
organic substance biosynthetic process
biosynthetic process
biosynthetic process
cellular nitrogen compound biosynthetic process
cellular component organization or biogenesis
organelle organization
organelle organization
cellular component biogenesis
cellular component biogenesis
cell division
cell division


cyclosome regulation
Metaphase to Anaphase transition
Aurora-B cell cycle regulation
S phase (Cdk2)
G2/M phase (cyclin B: Cdk1)
cyclosome regulatory network
G2/M phase (cyclin A: Cdk1)
ID complex ubiquitinatio n
Plk1 ---> INCENP
Bub1 ---> APC7
Cdk1, Plk1 ---/ cyclin B
Plk1 ---> Bub1
borealin ---> Aurora-B
E2F network
G1 phase (Cdk2)


600
550
500
450
400
350
300
250
200
150
100
50
0
Neoplasms
Breast Diseases
Breast Neoplasms
Neoplasms by Site
Digestive System Diseases
Digestive System Neoplasms
Colonic Neoplasms
Carcinoma
Adenocarcinoma
Liver Neoplasms
Carcinoma, Hepatocellular


integrins
MKP-4
Rizatriptan
Tegafur
SPK
2,5,7-Trihydrox...
Lovastatin
ilaprazole
STAT3, STAT1, NR3C1
JUN, TCF3, PAX2
Cellular nitrogen compound metabolic process;
Heterocycle metabolic process;
Nucleobase-containing compound metabolic process
Localization;
Transport;
Establishment of localization
cument Connection

Drugs approved in clinical trials for Oncology

Table 12. Clinically approved (FDA, ENA, etc.) drugs for the studied pathology (most promising and clinically approved treatment candidates selected for the identified drug targets on the basis of literature curation in HumanPSD™ database)

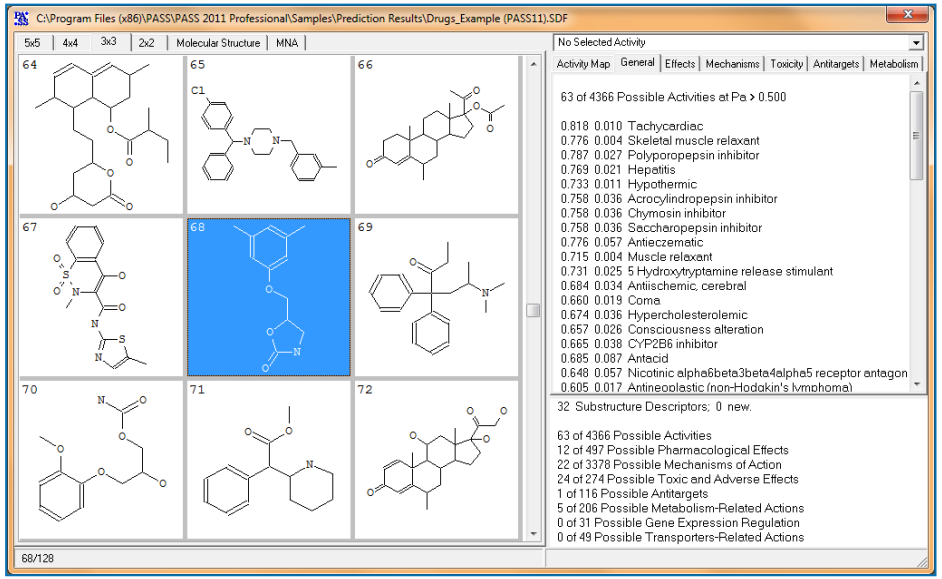
See full table →

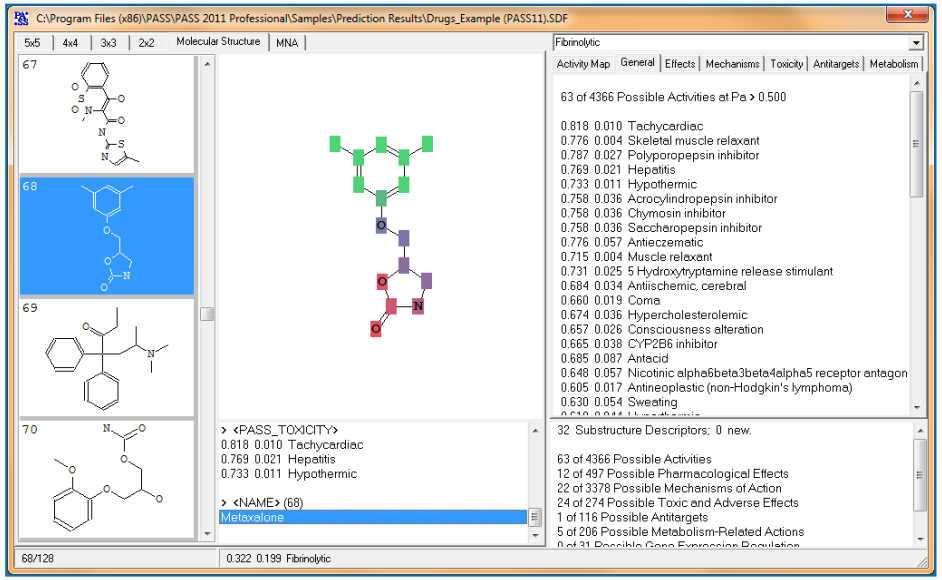
Name	Target names	Drug score	Disease activity score	Disease trial phase	Approved
Lapatinib	BMPR1A, NEK2, MAP2K3, NEK6, LIMK1, PRKD1, PIP5K1A... (more)	97	6	Phase 3: Breast Neoplasms, Breast Diseases, Liver Neoplasms, Neoplasm Metastasis, Neoplasms	Breast Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, FDA)
Paclitaxel	MAPK8, BCL2, E2F1, BIRC5, CDK1, CDK2, BRCA1... (more)	86	8	Phase 3: Breast Neoplasms, Adenocarcinoma, Anus Diseases, Anus Neoplasms, Arterial Occlusive Disease... (more)	Breast Neoplasms (FDA, FDA)
trastuzumab deruxtecan	PARP1, AKT1, CHEK1	80	6	Phase 3: Breast Neoplasms, Neoplasms	Breast Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, DailyMed, FDA)
neratinib	EGFR, AKT3, AKT1, RB1, AKT2, CDKN1B	80	3	Phase 2: Breast Neoplasms, Carcinoma, Non-Small-Cell Lung, Ependymoma, Fibroma, Glioblastoma, Hemang... (more)	Breast Neoplasms (ClinicalTrials, ClinicalTrials, ClinicalTrials, FDA, PUBMED)
Everolimus	AKT3, BCL2, AKT1, RB1, MAPKAP1,	78	7	Phase 4: Breast Neoplasms, Acute Coronary Syndrome, Angina Pectoris, Angina	Breast Neoplasms (FDA, PUBMED)

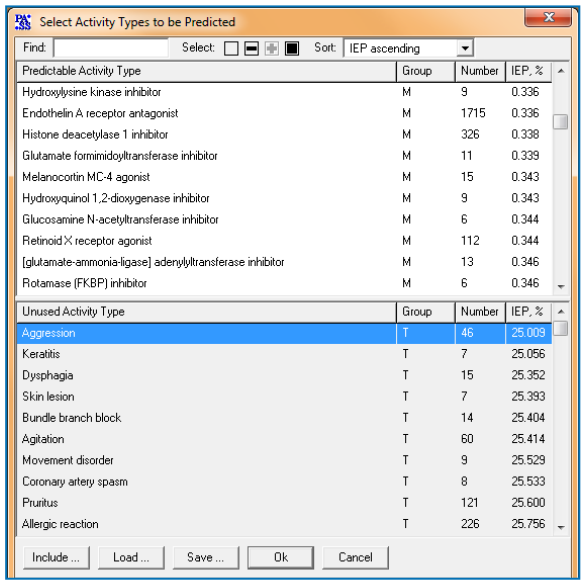
PASS

The acronym PASS stands for Prediction of Activity Spectra for Substances. PASS performs an instant prediction and computational evaluation of biological activity spectra for organic chemical compounds.

PASS results can be further interpreted via the PharmaExpert tool and combined with the structure-activity relationship models built in GUSAR.







Target activity score

$$T\text{-}score(s) = \frac{|T|}{|T| + w(|AT| - |T|))} \sum_{m \in M(s)} \left(pa(m) \sum_{g \in G(m)} IAP(g) optWeight(g) \right),$$

M(s) is the set of activity-mechanisms for the given structure *s*

G(m) is the set of targets (converted to genes)

IAP(g) is the invariant accuracy of prediction

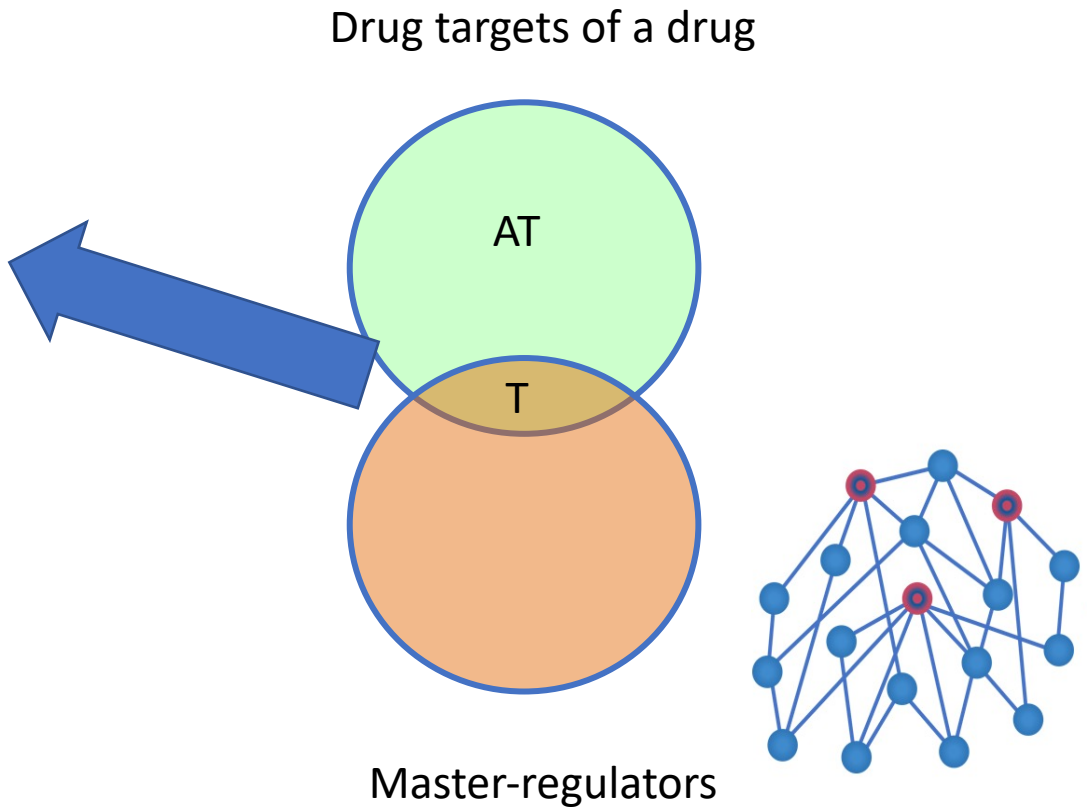


Table 15. Prospective drugs, predicted by PASS software to be active against the identified drug targets with predicted activity against the studied disease(s) (drug candidates predicted with the cheminformatics tool PASS)
See full table →

Name	Target names	Drug score	Target activity score
Disulfiram	SOD1, MAPK8, LAP3, MAPK12, MAPK9, HSPD1, MAPK6	81	0.13

Table 16. Prospective drugs, predicted by PASS software to be active against the identified drug targets, though without cheminformatically predicted activity against the studied disease(s) (drug candidates predicted with the cheminformatics tool PASS)
See full table →

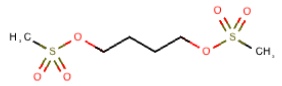
Name	Target names	Drug score	Target activity score
2,5,7-Trihydroxynaphthoquinone	MAPK8, DUSP23, MAPK9, SENP6, EPM2A, MAPK6, PTEN...(more)	97	0.92
Busulfan	DUSP23, EPM2A, PTPN2, PTPRS, EYA2, PTEN, CDC25C...(more)	96	1.74
Pentabromophenol	LAP3, DUSP23, EPM2A, PTPN2, PTPRS, EYA2, HDAC3...(more)	96	2
7-[4-(Dimethylamino)Phenyl]-N-Hydroxy-4,6-Dimethyl-7-Oxo-2,4-Heptadienamide	HDAC4, HDAC2, HDAC3	96	1.62
Iodophenyl	RPS6KA3, ROCK2, CSNK1E, RPS6KA1, AURKB, VRK1, ARAF...(more)	95	5.52

As the result of drug search we propose the following drugs as most promising candidates for treating the pathology under study: Lapatinib, seliciclib, Disulfiram and 2,5,7-Trihydroxynaphthoquinone. These drugs were selected for acting on the following targets: BIRC5, RPS6KA2, MAPK12 and DUSP9, which were predicted to be active in the molecular mechanism of the studied pathology.

The selected drugs are top ranked drug candidates from each of the four categories of drugs: (1) FDA approved drugs or used in clinical trials drugs for the studied pathology; (2) repurposing drugs

Busulfan

☆ Star 0

Summary	**Busulfan** is an alkylating agent used to treat chronic myelogenous leukemia.
Brand Names	*Busulfex, Myleran*
Generic Name	Busulfan
DrugBank Accession Number	DB01008
Background	Busulfan is a bifunctional alkylating agent, having a selective immunosuppressive effect on bone marrow. It is not a structural analog of the nitrogen mustards. It has been used in the palliative treatment of chronic myeloid leukemia (myeloid leukemia, chronic), but although symptomatic relief is provided, no permanent remission is brought about. According to the Fourth Annual Report on Carcinogens (NTP 85-002, 1985), busulfan is listed as a known carcinogen.
Type	Small Molecule
Groups	Approved, Investigational
Structure	
Weight	Average: 246.302 Monoisotopic: 246.02317956
Chemical Formula	$C_6H_{14}O_6S_2$

Provided by ClinicalTrials.gov and the registries and data partners contributing to the OpenTrials project.

Show 5 entries Search: breast

Disease	Phase	Study Title	Status	Start Date	End Date	Process Date
Neoplasms	Phase 3	Combination Chemotherapy and Peripheral Stem Cell Transplantation in Treating Patients With Stage II or Stage IIIA Breast Cancer View study report	Completed	1998-07-31	2003-03-31	ClinicalTrials.gov processed this data on 2023-04-26
Breast Neoplasms	Phase 3	Combination Chemotherapy and Peripheral Stem Cell Transplantation in Treating Patients With Stage II or Stage IIIA Breast Cancer View study report	Completed	1998-07-31	2003-03-31	ClinicalTrials.gov processed this data on 2023-04-26
Neoplasm, Residual	Phase 2	Autologous Hematopoietic Stem Cell Transplantation as Adjuvant Treatment for Triple Negative Breast Cancer Patients View study report	Unknown status	2018-02-01	2021-11-30	ClinicalTrials.gov processed this data on 2023-04-27
Neoplasms	Phase 2	Autologous Hematopoietic Stem Cell Transplantation as Adjuvant Treatment for Triple Negative Breast Cancer Patients View study report	Unknown status	2018-02-01	2021-11-30	ClinicalTrials.gov processed this data on 2023-04-27
Breast Neoplasms	Phase 2	Autologous Hematopoietic Stem Cell Transplantation as Adjuvant Treatment for Triple Negative Breast Cancer Patients View study report	Unknown status	2018-02-01	2021-11-30	ClinicalTrials.gov processed this data on 2023-04-27

Showing 1 to 5 of 18 entries (filtered from 2,985 total entries) First Previous 1 2 3 4 Next Last

Toxicity Bioassays Tested what is this?

Toxicity endpoints for which Busulfan has been tested (40 endpoints)

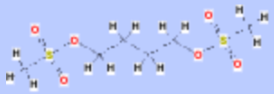
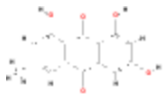
As reported by the FDA

Show 5 entries Search:

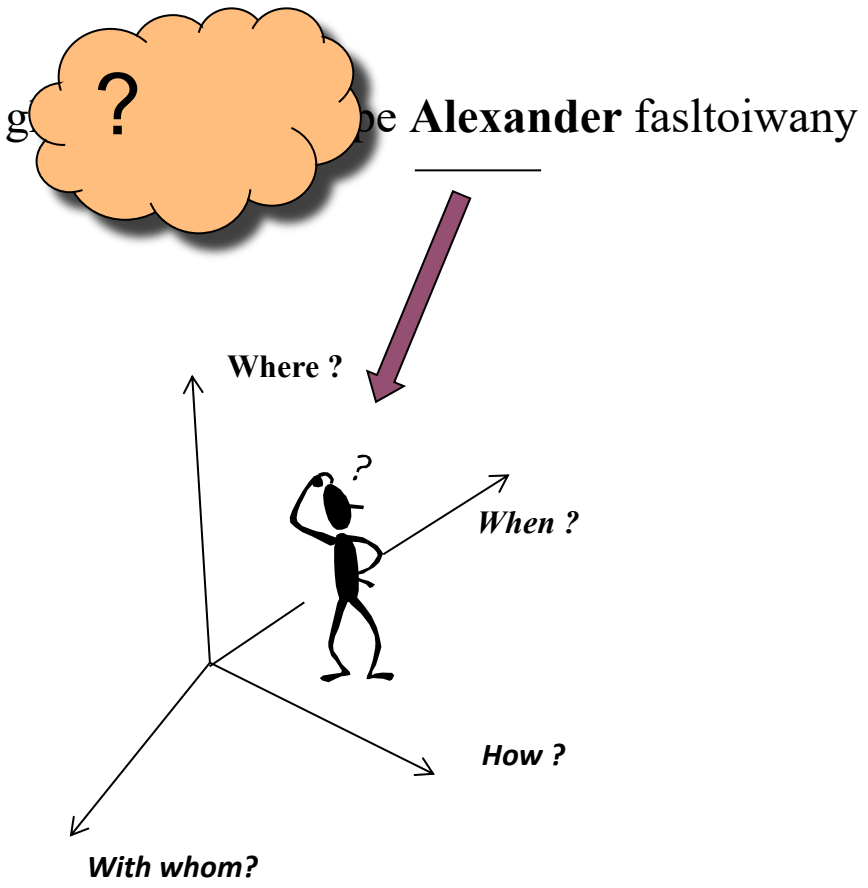
Toxicity endpoint	Toxicity category
Chromosome aberration test in CHL cells	Genetic toxicity
Chromosome aberrations in vitro composite	Genetic toxicity
Developmental toxicity in rodent fetus composite	Reproductive and developmental toxicity

Start page | Drugs PASS repurposed X | Report X | Keynodes for best mode... X

First Previous Page 1 of 7 Next Last Showing 1 to 50 of 318 entries Show 50 entries

ID	Accession	Name	Structure	Target names	Target activity score	Target activity rank	Toxicity score	Disease activity score	Disease activity rank	Drug rank	Drug score
PC:4555	DR000001283	2,5,7-Trihydroxynaphthoquinone		MAPK8, DUSP23, MAPK9, SENP6, EPM2A, MAPK6, PTEN, TNS2, CDC25A, DUSP2, MAPK12, POR, CDKN3, CDC25B,... (more)	0.91996	14	0.865	0.714	5	19	97
PC:2848	DR000001002	Busulfan		DUSP23, EPM2A, PTPN2, PTPRS, EYA2, PTEN, CDC25C, PTPRU, TNS2, UBASH3B, PPM1B, SOD1, CDC25A, DUSP2... (more)	1.73716	10	0.943	0.637	17	27	96
PC:3715	DR000000635	Pentabromophenol		LAP3, DUSP23, EPM2A, PTPN2, PTPRS, EYA2, HDAC3, PTEN, CDC25C, PTPRU, UBASH3B, PPM1B, CDC25A, ... (more)	1.99847	8	0.745	0.63	19	27	96
PC:4873	DR000003237	7-[4-(Dimethylamino)Phenyl]-N-Hydroxy-4,6-Dimethyl-7-Oxo-2,4-Heptadienamide		HDAC4, HDAC2, HDAC3	1.62235	11	0.536	0.645	16	27	96
PC:3707	DR000004081	Iodophenyl		RPS6KA3, ROCK2, CSNK1E, RPS6KA1, AURKB, VRK1, ARAF, NEK2, SGK1, MASTL, NEK6, AKT1, AURKA, CHEK1, ... (more)	5.51612	2	0.744	0.552	32	34	95
PC:2954	DR000010553	3-METHYL-1,6,8-TRIHYDROXYANTHRAQUINONE		MAPK8, DUSP23, MAPK9, EPM2A, PKM, MAPK6, PTEN, TNS2, DUSP2, MAPK12, POR, CDKN3, DUSP9, PIP5K1A, ... (more)	0.60282	25	0.814	0.546	34	59	91
PC:4450	DR000001412	6,4'-Dihydroxy-3-Methyl-3',5'-Dibromoflavone		EPM2A, MAPK6, HDAC3, CDC25C, PTEN, PTPRU, UBASH3B, TNS2, PPM1B, POR, CDKN3, PTPN12, DUSP9, MAPK8,... (more)	2.22117	7	0.839	0.472	55	62	90
PC:3543	DR000001556	2-HYDROXY-1,4-NAPHTHOQUINONE		LAP3, MAPK8, MAPK12, MAPK9, SENP6, POR, MAPK6, CDC25B, PIP5K1A, BRCA1	0.3694	38	0.768	0.586	25	63	90

“Regulatory code”

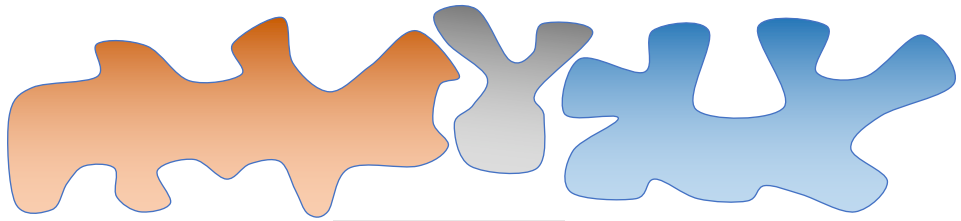


Several regulatory messages could be written in the same sequence

Reading of the messages depends on the cellular context

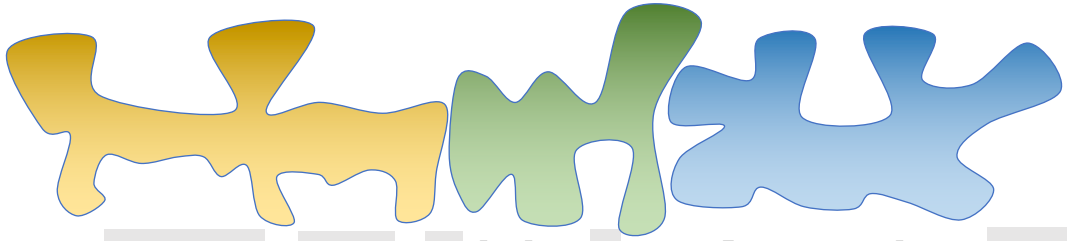
gherllojunomd-bype Alexander fasltoiwany

1)



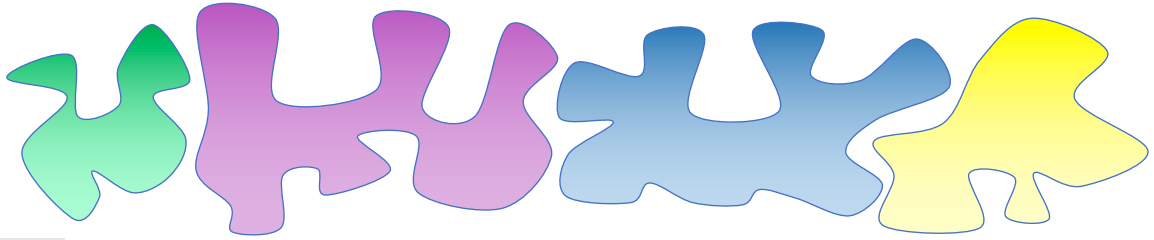
he ll o Alexander

2)



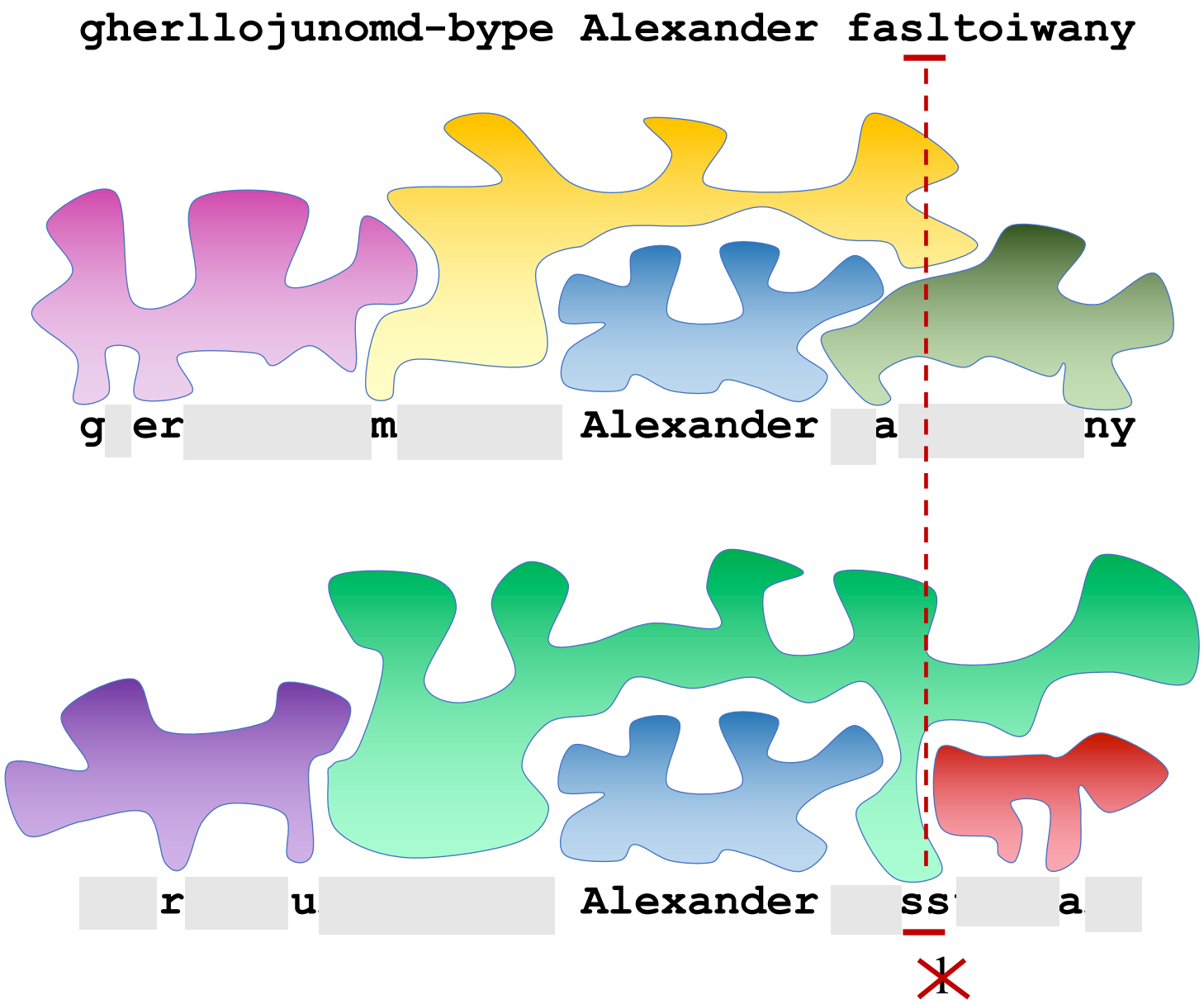
g o o d-by e Alexander

3)



r un Alexander fas t

Even some messages which were not written

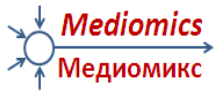


AI should learn to listen good music

Thank you!

Funding

EU



Russia