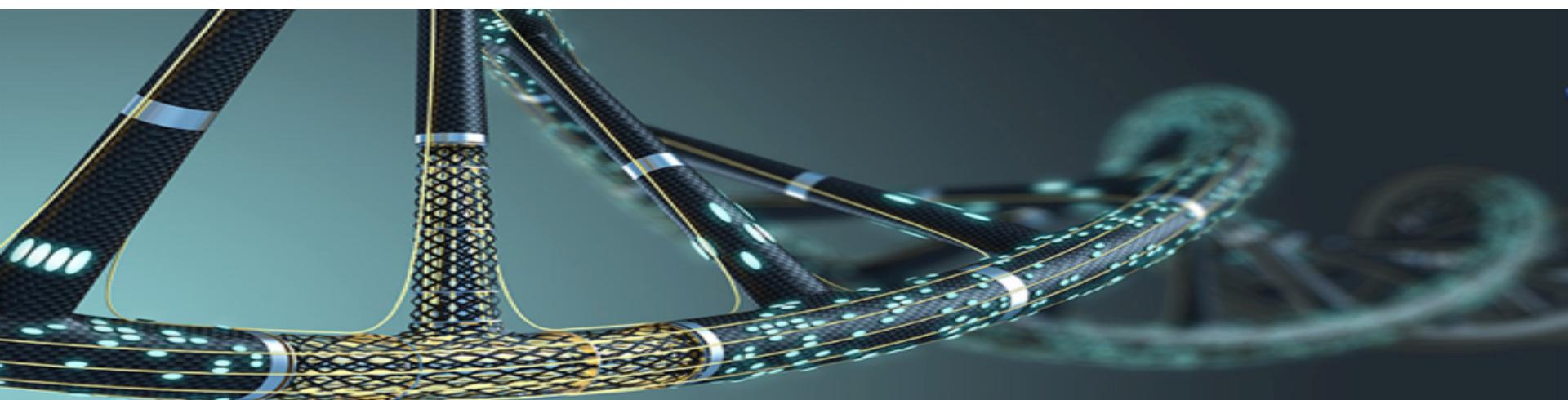




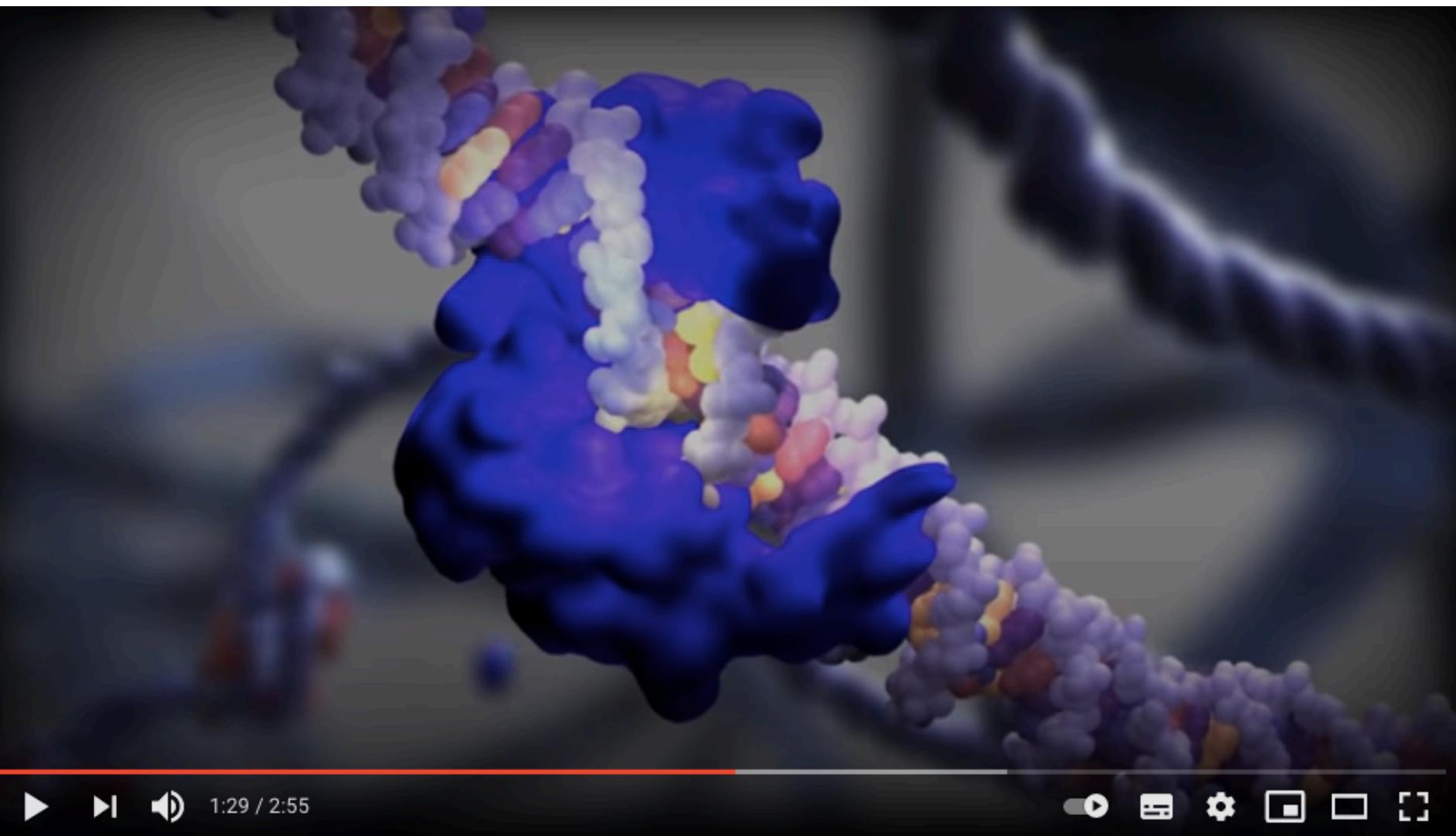
Майнор по биоинформатике

Лекция 13

Мария Попцова



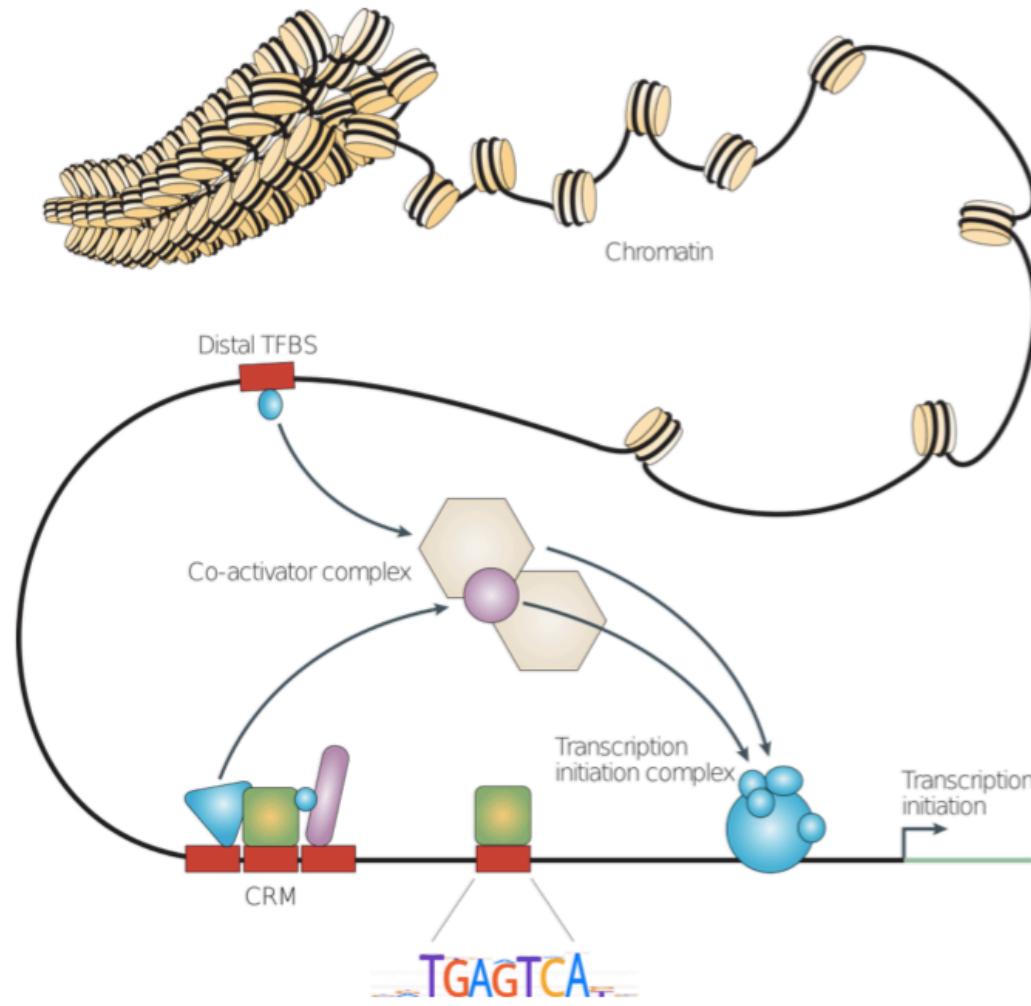
How genes are regulated



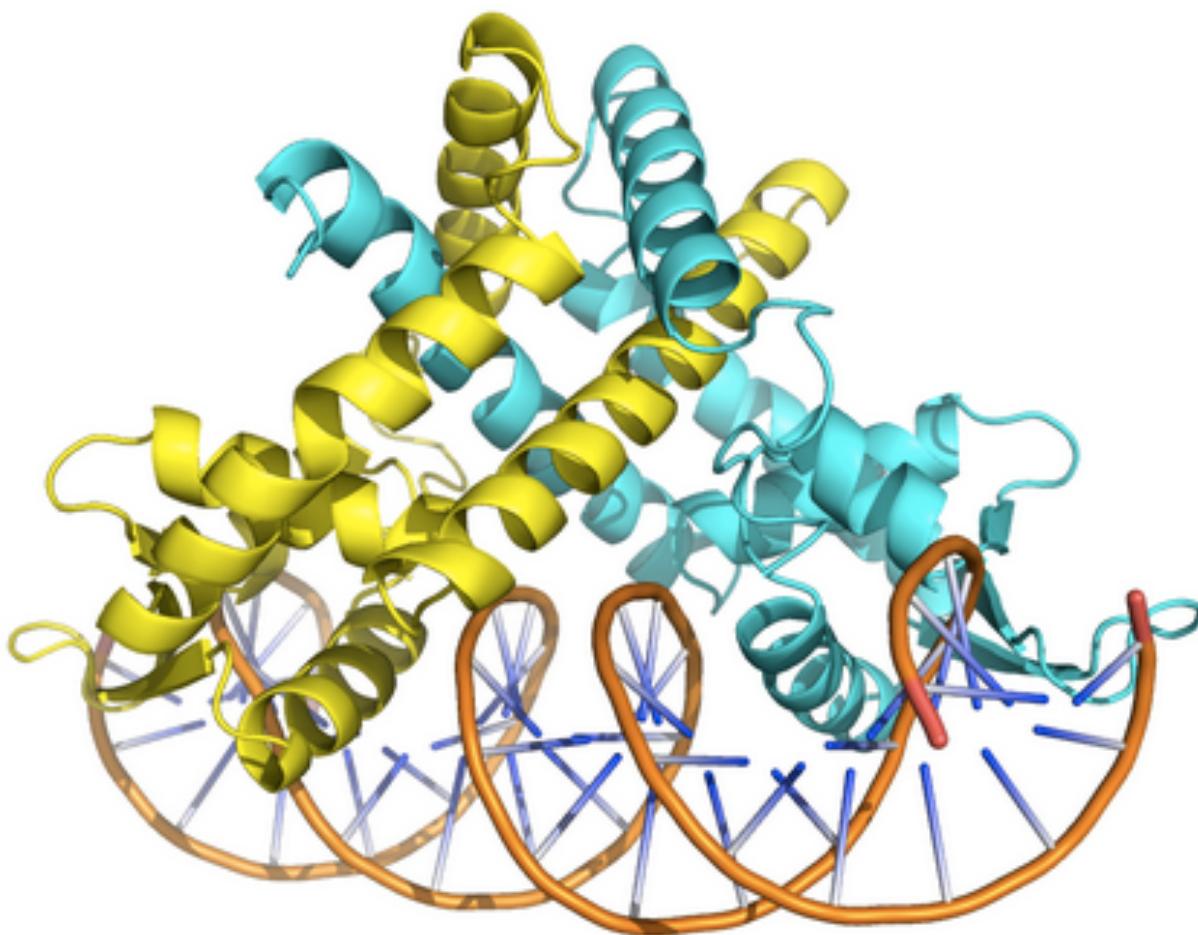
▶ ▶ 🔍 1:29 / 2:55

▢ ⏹ ⚙ ☰ ☱ ☲ ☳

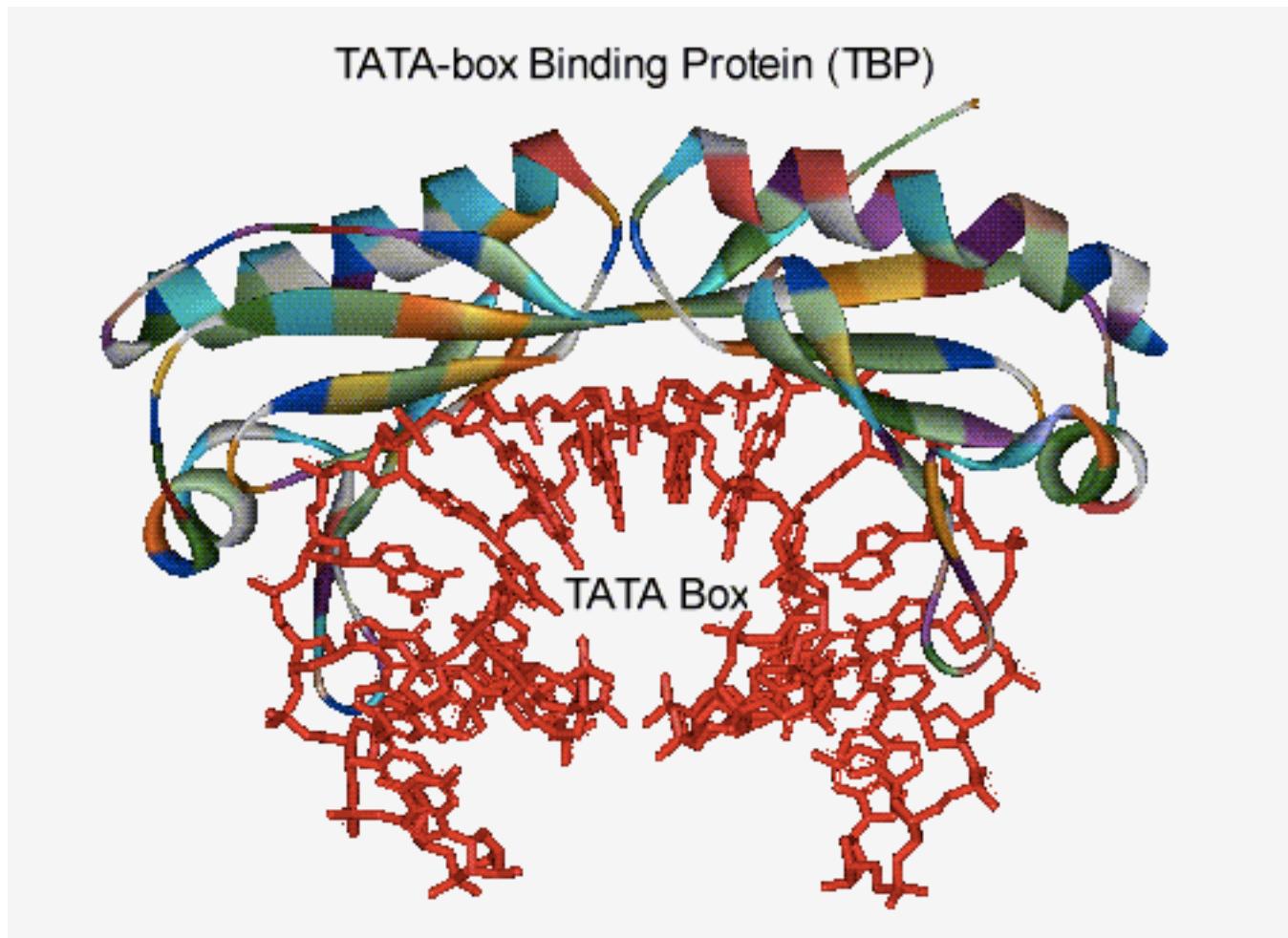
Transcription factors and their binding sites (TFBS)

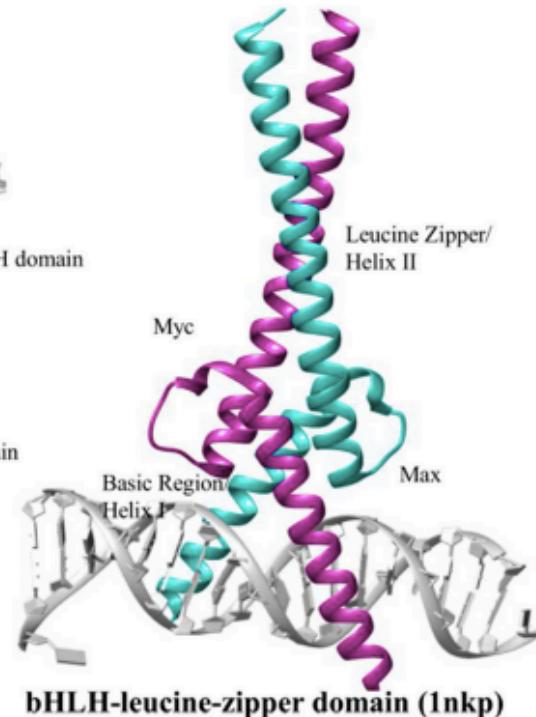
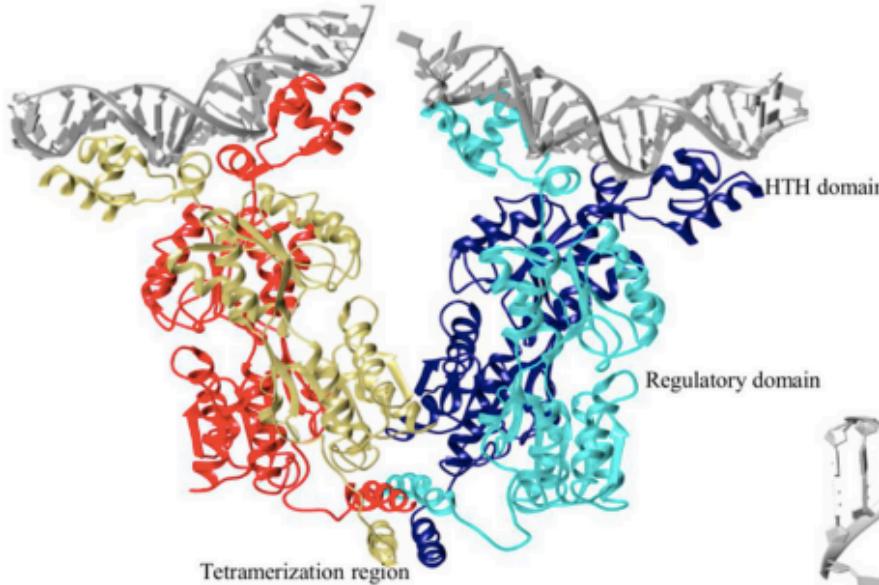
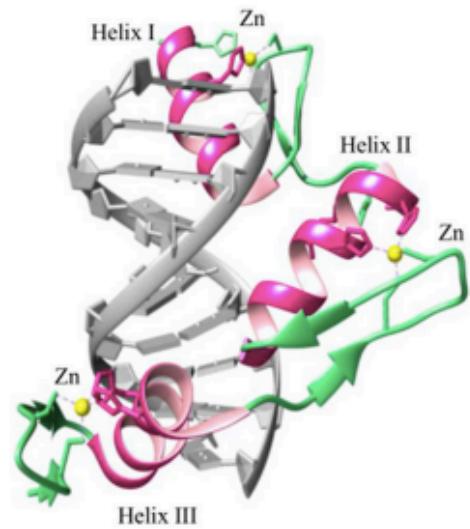
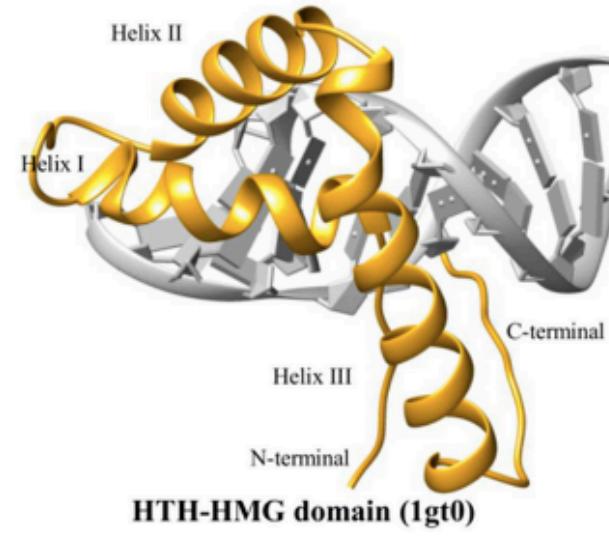
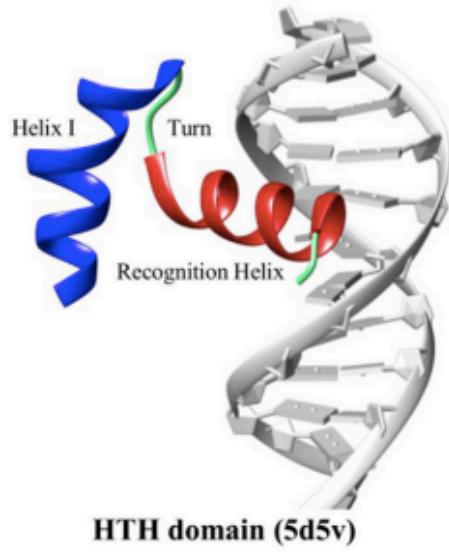
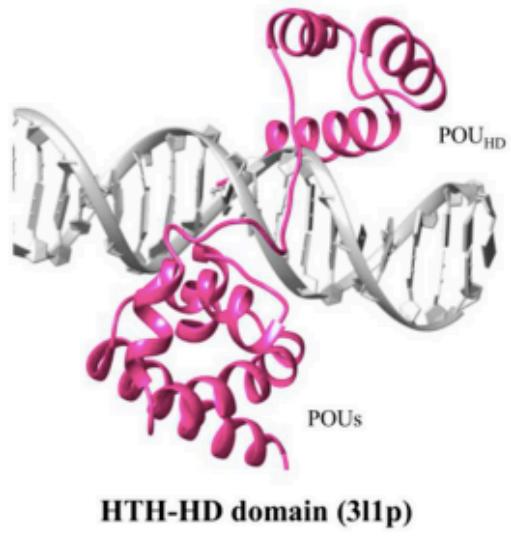


DNA-binding proteins

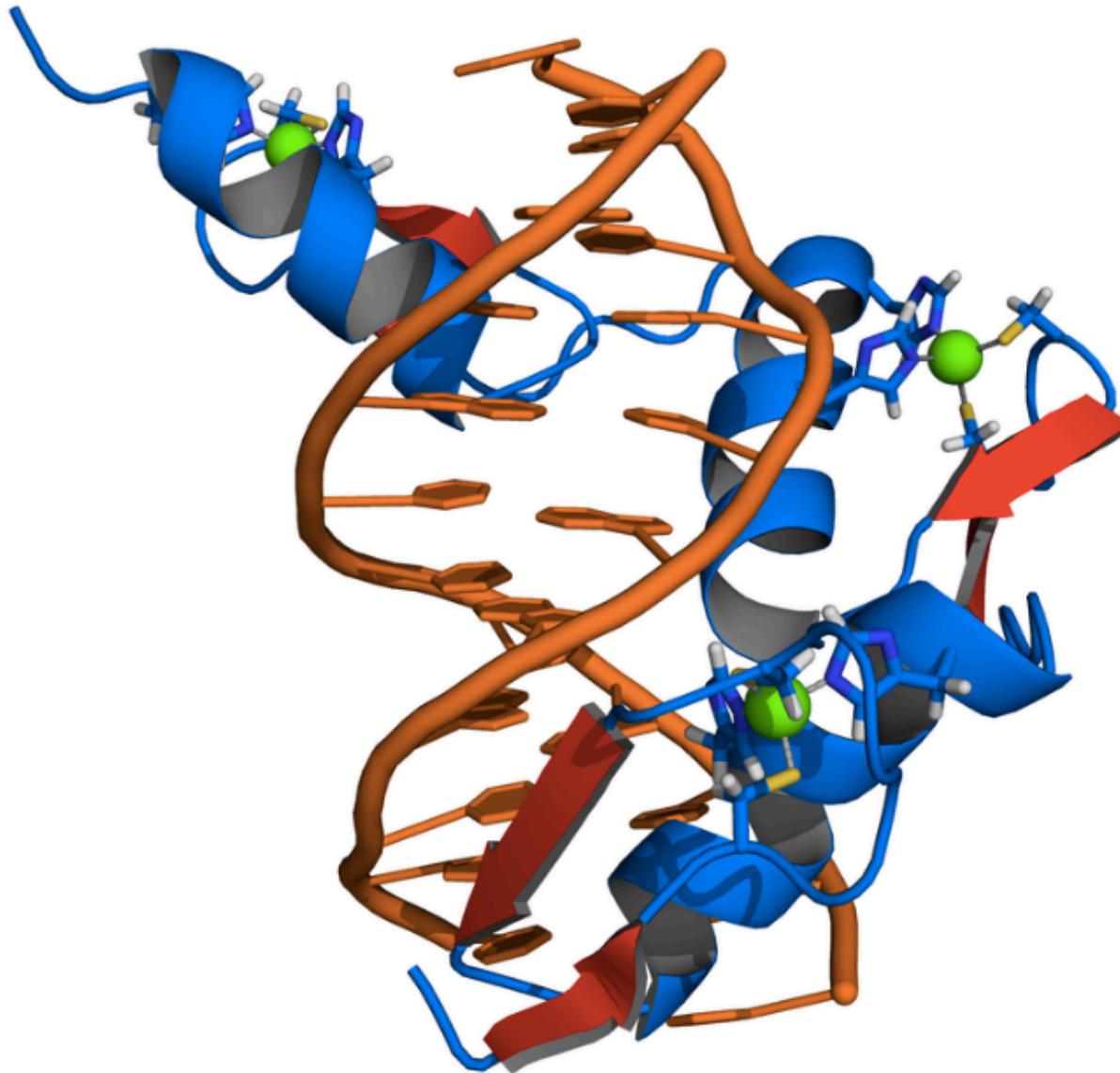


TATA-box binding proteins





Zinc Fingers

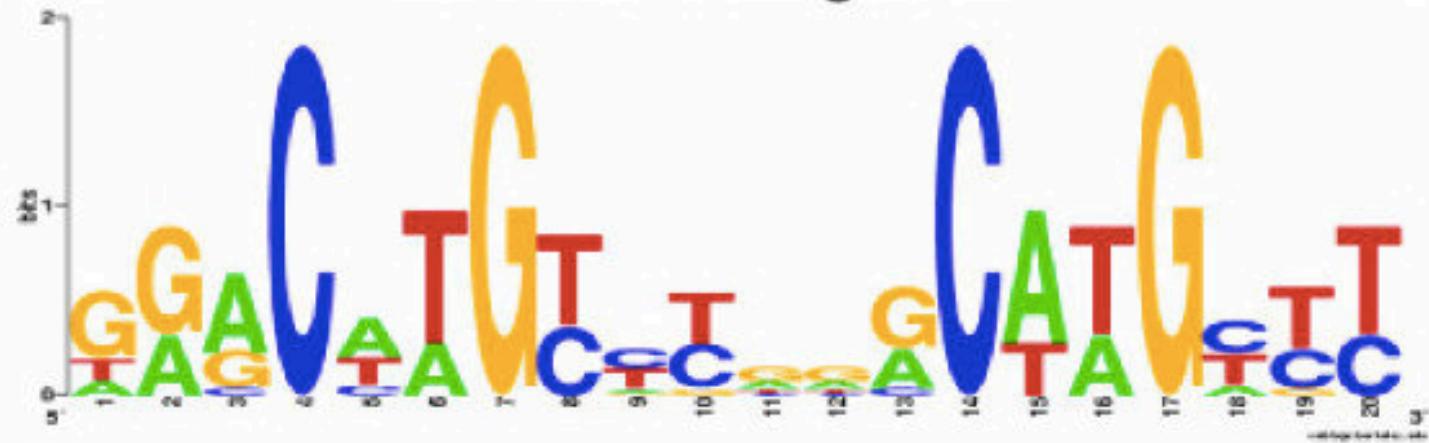


Zif268

DNA Motifs

A.

induced binding sites



B.

repressed binding sites

TP53



ШАГ 1. Подсчёт числа букв $N(b,j)$

1234567890123456

ACGCAAAACGTTTCTT

TCGCAAAACGTTTGCTT

ACGGAAACGTTTCGT

ACCCAAACCCCTTTGCT

ACCCGAAAGCCCTTTTTCCT

ACGCAACCGTTTCTT
TGGATTAAGCTTAACT

ACGCAAAACGTGTGCGT

ACGCAATCGGTACCT

GCGCAAACGTTTCGT

AGGAAAAACGATTGGCT

AAGCAAACGGTGATT

ATGCAATCGGTTACGC

AGGCAAACGTTACCT

ШАГ 2. Частоты букв

$$f(b,j) = N(b,j)/N \text{ в примере } N=13$$

Частоты	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	0.77	0.15	0.00	0.08	1.00	1.00	0.77	0.00	0.00	0.08	0.00	0.00	0.31	0.00	0.08	0.00
G	0.15	0.15	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.31	0.08	0.08	0.23	0.08	0.38	0.00
T	0.08	0.08	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.62	0.92	0.92	0.38	0.08	0.23	0.85
C	0.00	0.62	0.00	0.92	0.00	0.00	0.08	1.00	0.00	0.00	0.00	0.00	0.08	0.85	0.31	0.15
Всего	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

G C C T A C C C A T T A T T T...

ШАГ 3. Логарифм отношения вероятностей

Частота G в позиции 15 равна 0.38

Значит ли это что-нибудь, если GC состав генома равен 0.7,
т.е. частота G в геноме равна 0.35?

ЛОГАРИФМ Отношения правдоподобия W как вес различия
наблюдаемой частоты и ожидаемой:

$$w(G, 15) = \ln(0.38/0.35) = 0.1$$

ШАГ 4. Матрица весов PWM

w(b,j)	Фоновые частоты	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	0.15	1.6	0.0	-inf	-0.7	1.9	1.9	1.6	-inf	-inf	-0.7	-inf	-inf	0.7	-inf	-0.7	-inf
G	0.35	-0.8	-0.8	1.0	-inf	-inf	-inf	-inf	-inf	1.0	-0.1	-1.5	-1.5	-0.4	-1.5	0.1	-inf
T	0.15	-0.7	-0.7	-inf	-inf	-inf	-inf	0.0	-inf	-inf	1.4	1.8	1.8	0.9	-0.7	0.4	1.7
C	0.35	-inf	0.6	-inf	1.0	-inf	-inf	-1.5	1.0	-inf	-inf	-inf	-inf	-1.5	0.9	-0.1	-0.8
	1	-inf	-0.9	-inf	-0.3	-inf	-0.3	-inf									

Шаг 5. Псевдоотсчёты

$$F(b,j) = [N(b,j) + \varepsilon(b)] / (N + \varepsilon) \quad \text{вместо}$$

$$f(b,j) = N(b,j)/N$$

Здесь $\varepsilon = \varepsilon(A) + \varepsilon(G) + \varepsilon(T) + \varepsilon(C)$

Все $\varepsilon(b)$ маленькие в сравнении с N

Подбираются опытным путём

ШАГ 6. Частоты с псевдоотсчётоми

ШАГ 7. Матрица РWM с псевдоотсчётаами

$W(b,j)$	баз. Частоты	$e(b)$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	0.15	0.10	1.6	0.0	-3.0	-0.6	1.9	1.9	1.6	-3.0	-3.0	-0.6	-3.0	-3.0	0.7	-3.0	-0.6	-3.0
G	0.35	0.10	-0.8	-0.8	1.0	-3.8	-3.8	-3.8	-3.8	-3.8	1.0	-0.1	-1.5	-1.5	-0.4	-1.5	0.1	-3.8
T	0.15	0.10	-0.6	-0.6	-3.0	-3.0	-3.0	-3.0	0.0	-3.0	-3.0	1.4	1.8	1.8	0.9	-0.6	0.4	1.7
C	0.35	0.10	-3.8	0.5	-3.8	0.9	-3.8	-3.8	-1.5	1.0	-3.8	-3.8	-3.8	-3.8	-1.5	0.9	-0.1	-0.8
	1	0.40	-3.6	-0.8	-8.8	-6.5	-8.8	-8.8	-3.6	-8.8	-8.8	-3.2	-6.5	-6.5	-0.2	-4.2	-0.2	-5.9

Информационное содержание
выравнивания.

Сила сигнала.

LOGO

Содержание информации в выравнивании определяется тем, насколько выравнивание отличается от выравнивания случайных фрагментов той же длины и в том же числе

Как считать содержание информации

ШАГ 1. Подсчёт числа букв $N(b,j)$

1234567890123456
ACGCAAAACGTTTCTT
TCGCAAAACGTTGCTT
ACGCAAAACGTTTCGT
ACGCAAAACGGTTTCGT
ACGCAACC GTTTCCCT
ACGCAAAACGTGTGCGT
ACGCAATCGGTTACCT
GCGCAAAACGTTTCGT
AGGAAAACGATTGGCT
AAGCAAACGGTGATT
ATGCAAATCGGTTACGC
AGGCAAACGTTACCT
GAGCAAACGTTCCAC

ШАГ 2. Частоты букв $f(b,j)$

Частоты	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
A	0.77	0.15	0.00	0.08	1.00	1.00	0.77	0.00	0.00	0.08	0.00	0.00	0.31	0.00	0.08	0.00
G	0.15	0.15	1.00	0.00	0.00	0.00	0.00	0.00	1.00	0.31	0.08	0.08	0.23	0.08	0.38	0.00
T	0.08	0.08	0.00	0.00	0.00	0.00	0.15	0.00	0.00	0.62	0.92	0.92	0.38	0.08	0.23	0.85
C	0.00	0.62	0.00	0.92	0.00	0.00	0.08	1.00	0.00	0.00	0.00	0.00	0.08	0.85	0.31	0.15
Всего	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

G C C T A C C C A T T A T T T ...

Информационное содержание

$$IC(b,j) = f(b,j) * \log_2[f(b,j)/p(b)] = f(b,j) * w(b,j)$$

$\log_2[f(b,j)/p(b)] = w(b,j)$ – вес из матрицы PWM без псевдоотсчётов.

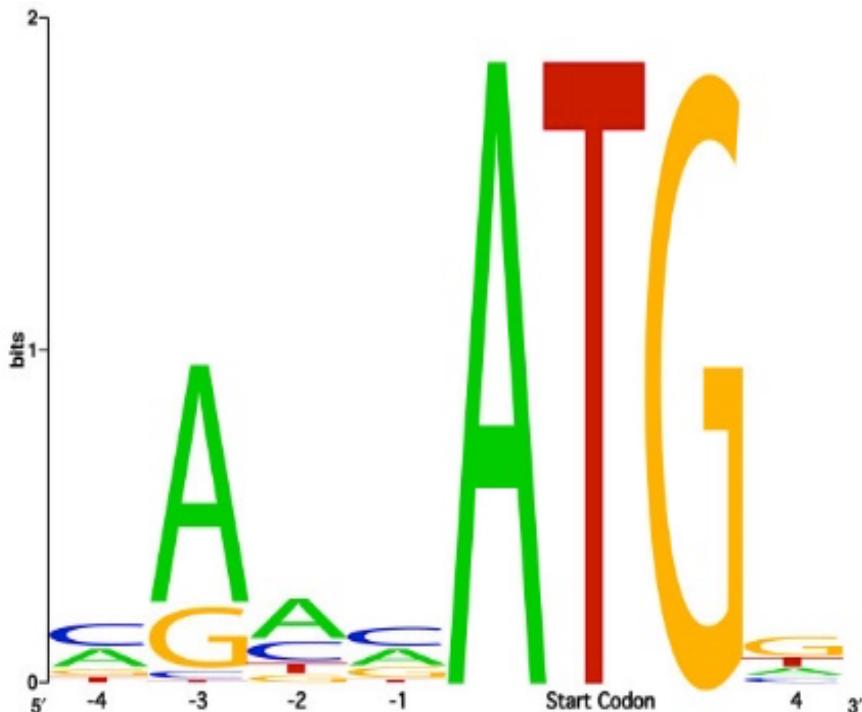
$IC(b,j) = 0$ при $f(b,j) = 0$

Также $IC(b,j) = 0$ если частота $f(b,j) = p(b)$

Максимум $IC(b,j) = \log_2[1/p(b)]$ для минимальной $p(b)$

Информационное содержание IC выравнивания равно

$$IC = \sum_j IC(j)$$



В LOGO сигнала буквы
имеют высоту, равную
информационному
содержанию букв.
 $IC(j)$ = высота j-го столбца

Сервис webLOGO: <http://weblogo.berkeley.edu/logo.cgi>

POSITION WEIGHT MATRIX

	frequency matrix						probability matrix					
123456	1	2	3	4	5	6	1	2	3	4	5	6
TATAAT	1	2	3	4	5	6	1	2	3	4	5	6
TAAAAT	A	0	4	2	4	4	0	0.0	0.8	0.4	0.8	0.8
TAATAT	C	0	0	0	0	1	0	0.0	0.0	0.0	0.0	0.2
TGTAAT	G	0	1	0	0	0	0	0.0	0.2	0.0	0.0	0.0
TATACT	T	5	0	3	1	0	5	1.0	0.0	0.6	0.2	0.0

$$M_{p,n} = \log_2 \left(\frac{p_{p,n}}{b_n} \right)$$

$p_{p,n}$ is probability of nucleotide n in position p

b_n is probability of nucleotide n in background

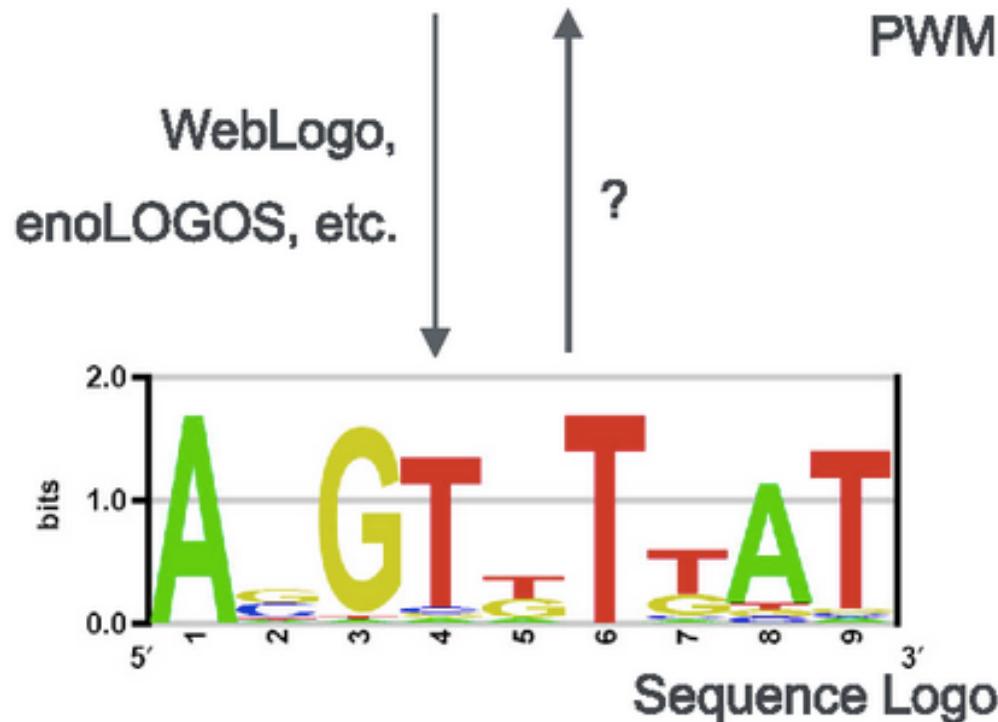
Add pseudocounts (for example, 1), to frequency matrix to evade infinity in PWMS. Pseudocounts reflect the fact, that any sequence can be bound by the protein. But some of them are bound with very low probability

	1	2	3	4	5	6
A	-Inf	1.6	0.6	1.6	1.6	-Inf
C	-Inf	-Inf	-Inf	-Inf	-0.3	-Inf
G	-Inf	-0.3	-Inf	-Inf	-Inf	-Inf
T	2	-Inf	1.2	-0.3	-Inf	2

	1	2	3	4	5	6
A	-1.2	1.2	0.4	1.2	1.2	-1.2
C	-1.2	-1.2	-1.2	-1.2	-0.2	-1.2
G	-1.2	-0.2	-1.2	-1.2	-1.2	-1.2
T	1.4	-1.2	0.8	-0.2	-1.2	1.4

PWM <-> LOGO

	1	2	3	4	5	6	7	8	9
A	.97	.10	.02	.03	.10	.01	.05	.85	.03
C	.01	.40	.01	.04	.05	.01	.05	.05	.03
G	.01	.40	.95	.03	.40	.01	.3	.05	.03
T	.01	.10	.02	.90	.45	.97	.6	.05	.91



Transfac Database

G H

genexplain.com



Log in

Home

Products ▾

Services

Free ▾

Research ▾

News

About us ▾

Contact

Upstream analysis

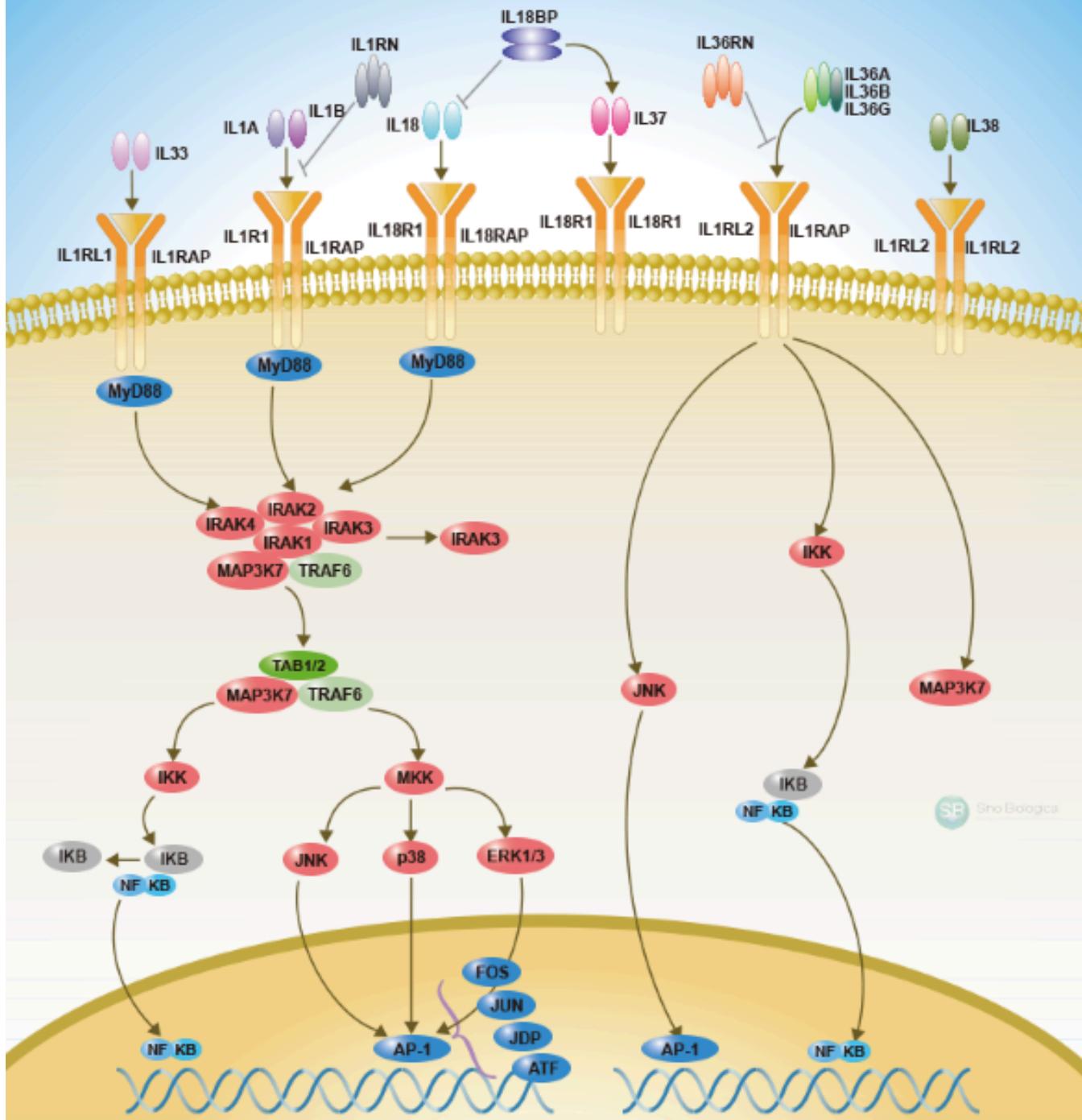
Discover unanticipated causal relationships in your data with our unique upstream analysis.



TRANSFAC 2.0 to be launched in summer 2021

News

Signaling networks in a cell





Reagents Search Google Search

Search



產品

CRO 服務

支持

最新活動

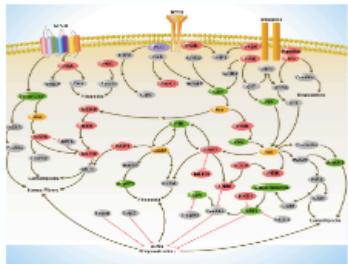
經銷商

關於我們

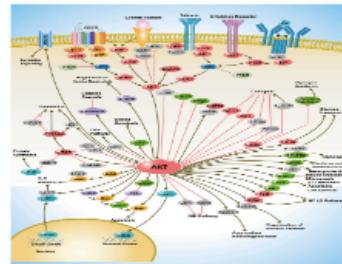
詢價

[Home](#) | [Pathways](#)

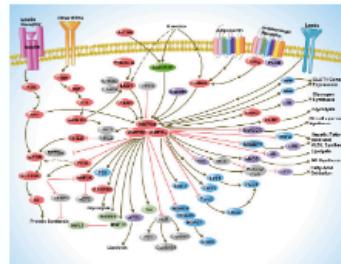
All Cell Signaling Pathways



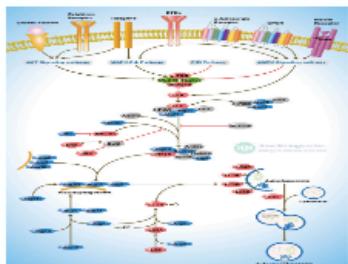
Actin Dynamics Signaling Pathway



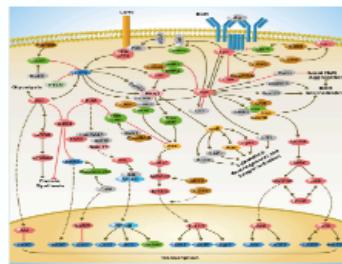
AKT Signaling Pathway



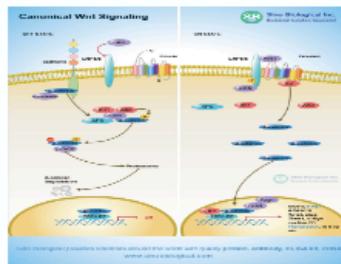
AMPK Signaling Pathway



Autophagy Pathway



B Cell Receptor Signaling Pathway



Canonical Wnt Pathway

Signaling Pathways

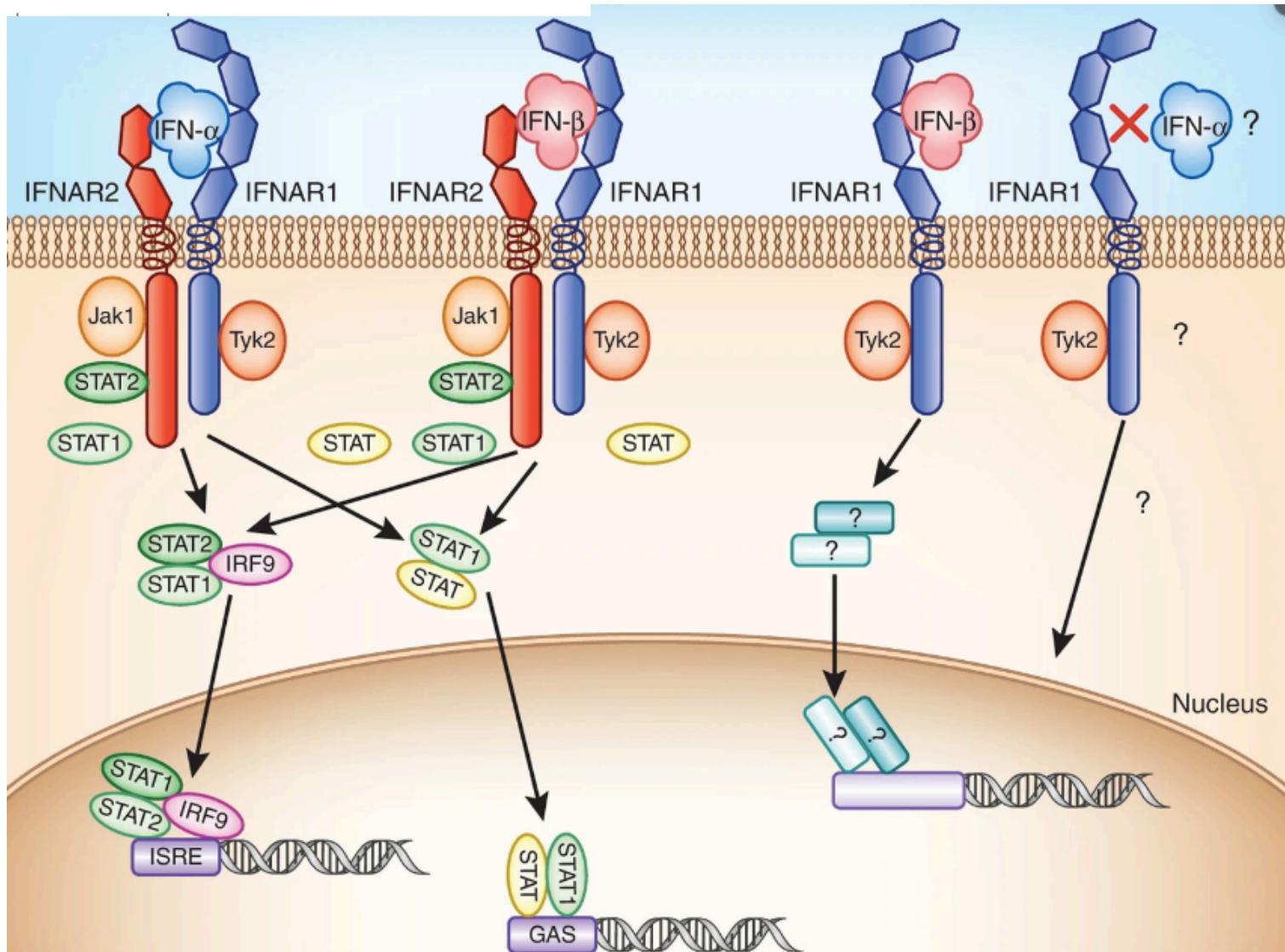
- [Actin Dynamics Signaling Pathway](#) >
- [AKT Signaling Pathway](#) >
- [AMPK Signaling Pathway](#) >
- [Autophagy Signaling Pathway](#) >
- [B Cell Receptor Signaling Pathway](#) >
- [Canonical Wnt Signaling Pathway](#) >
- [Common Cytokine Receptor G Chain](#) >

Product Type

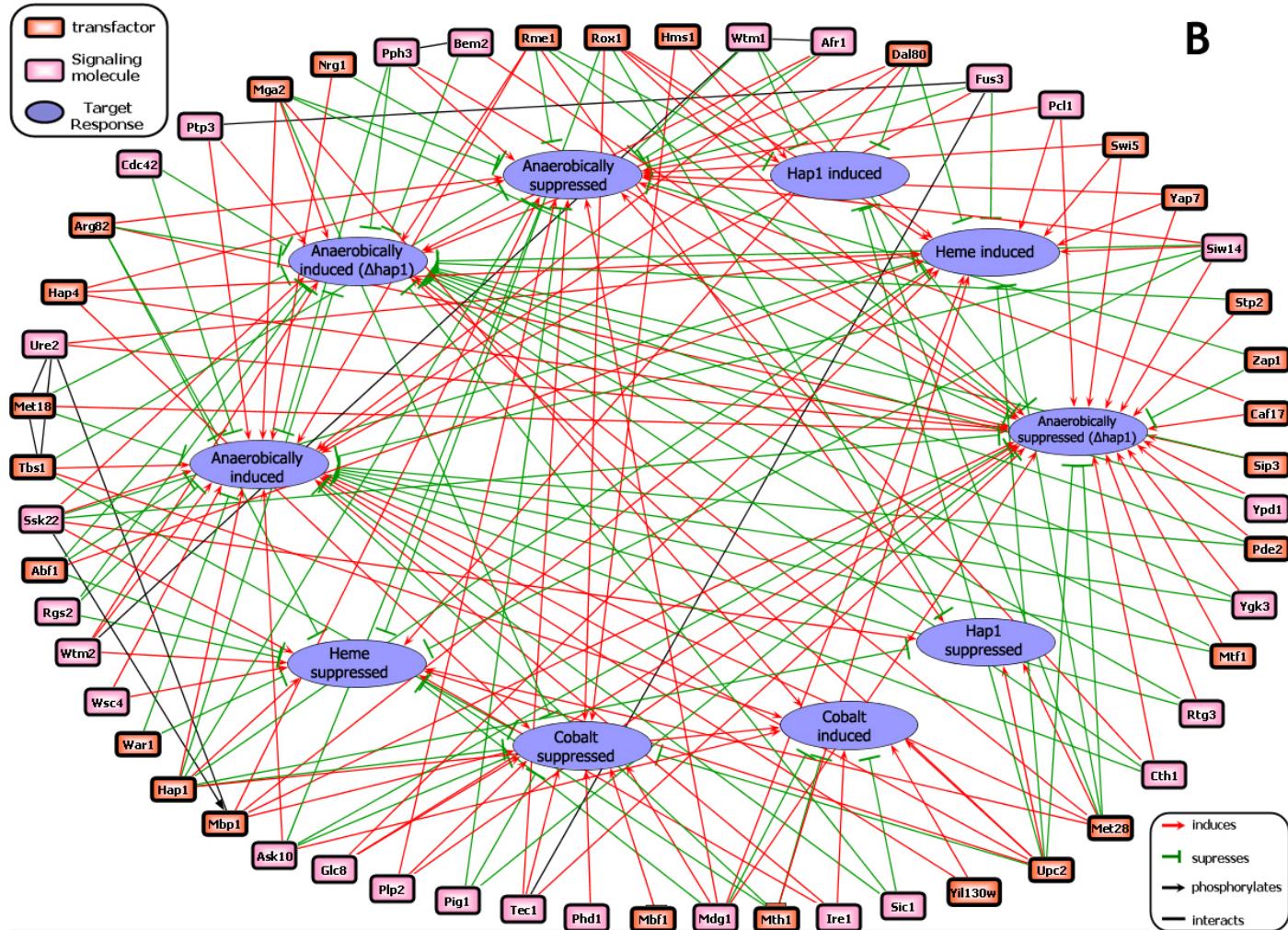
IFN- β -specific signaling via a unique IFNAR1 interaction

Surinder Kaur & Leonidas C Platanias [✉](#)

Nature Immunology 14, 884–885(2013) | [Cite this article](#)

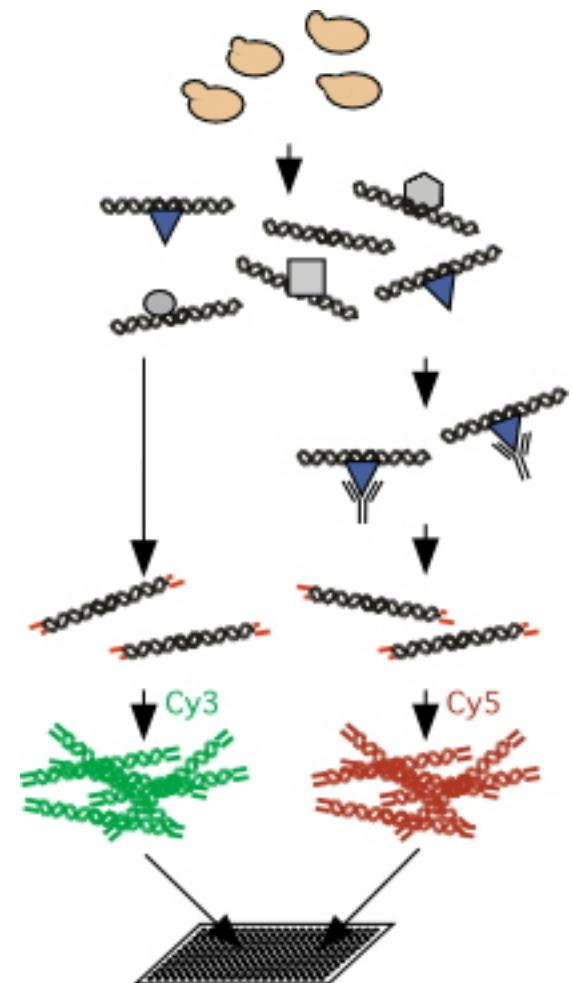
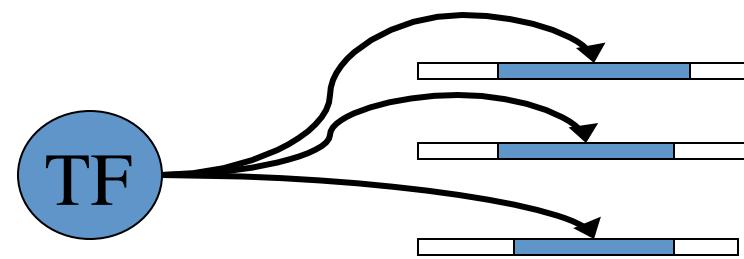


Example: oxygen sensing and regulatory network



Binding data for regulatory networks

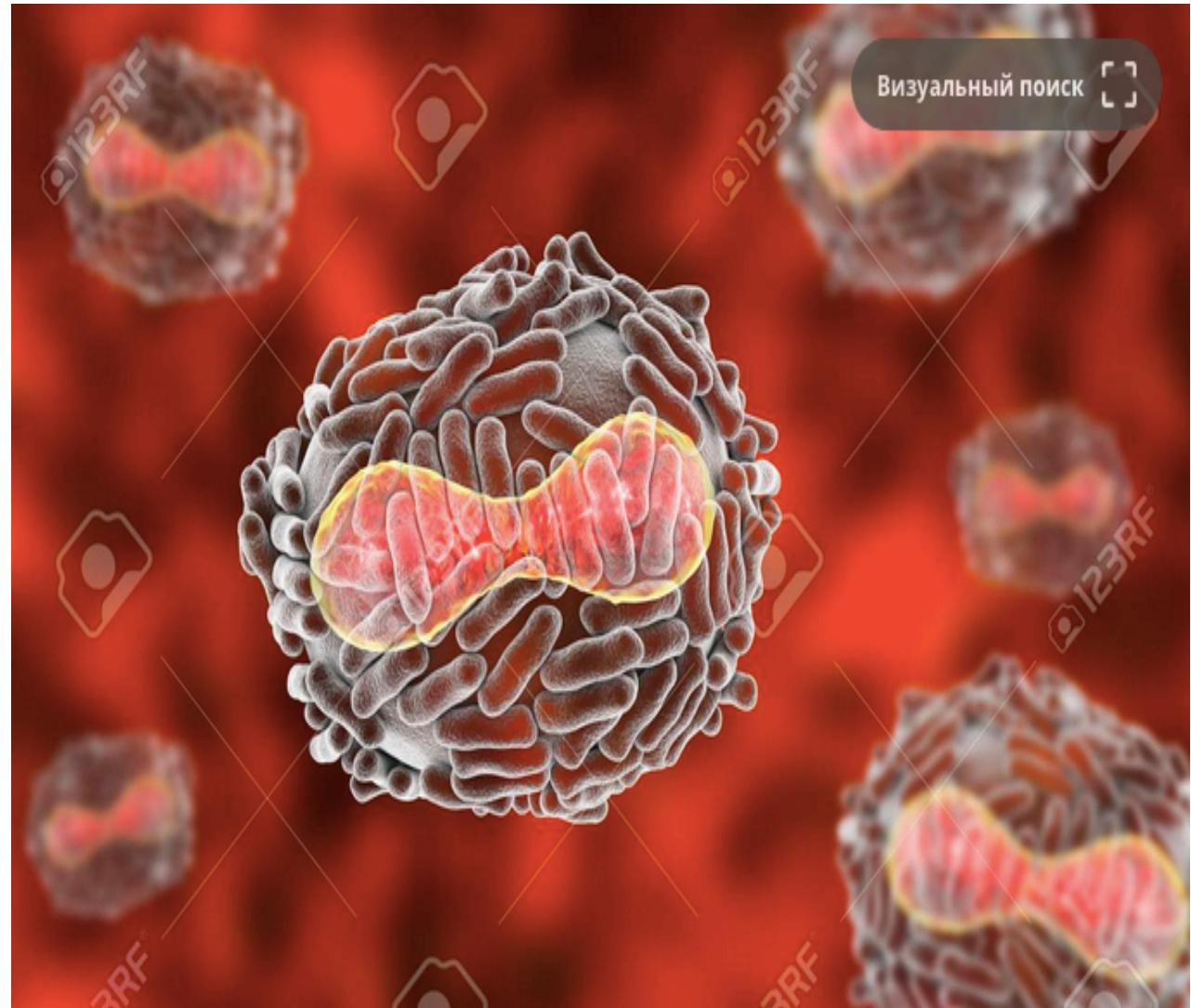
- ChIP-chip: genome-wide protein-DNA binding data, i.e. what promoters are bound by TF?
- Investigate regulatory network model:
use ChIP-chip data in place of motifs
(no motif discovery)
 - Features: (regulator, TF-occupancy) pairs



Вакцинация

Вирус оспы (Variola virus)

Variola is a large brick-shaped virus measuring approximately 302 to 350 nanometers by 244 to 270 nm,[43] with a single linear double stranded DNA genome 186 kilobase pairs (kbp) in size



Вариоляция



Китай, 15 век

Вариоляция в Европе

из осипин выздоравливающих больных извлекали биологический материал, который затем искусственно прививали здоровым людям. В XVIII веке это проделывали так: протягивали зараженную нитку под надрезанной кожей.

Георг 28 мая (7 июня) 1660 - 11 (22) июня 1727



Вариоляция во Франции

Людовик Великий Дофин Людóвик Велíкий Дофíн (фр. Louis le Grand Dauphin; 1 ноября 1661 — 14 апреля 1711) — единственный выживший законный ребёнок Людовика XIV от Марии-Терезии Испанской, его наследник (дофин Франции). Умер от оспы за четыре года до смерти отца и не царствовал.

НЕ ДЕЛАЛИ



Вариоляция в России

В 1730 году от оспы скончался 14-летний император Петр II. Заразил его один из князей Долгоруких, пришедший ко двору в то время, когда в доме его были больные оспой.

"Собою подала пример": на медали изображены Екатерина II и сын ее Павел, а напротив них женщина с детьми - воплощение Российской Империи



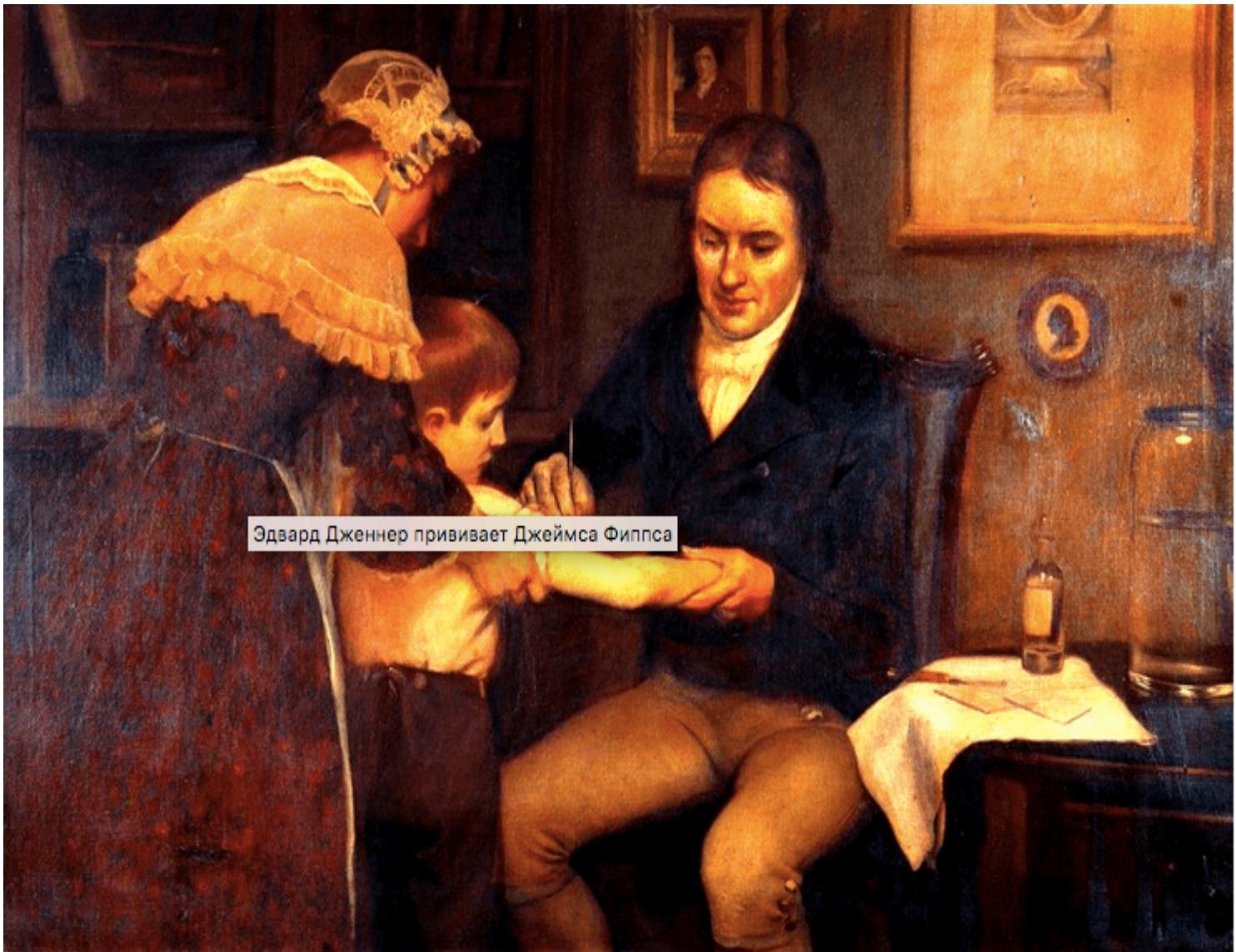
перенесла вариоляцию 12 октября 1768 года в Царском Селе. Материал взяли у семилетнего сына вахмистра Александра Маркова. Позже мальчику был пожалован дворянский титул, герб с изображенной на нем рукой с осиной, и фамилия Оспенныи.

Эдвард Дженнер

Именно
коровьей
оспой

1796

Эдвард Дженнер прививает Джеймса Филпса



Variolae Vaccinae – коровья оспа

Вакцинация =
окоровливание

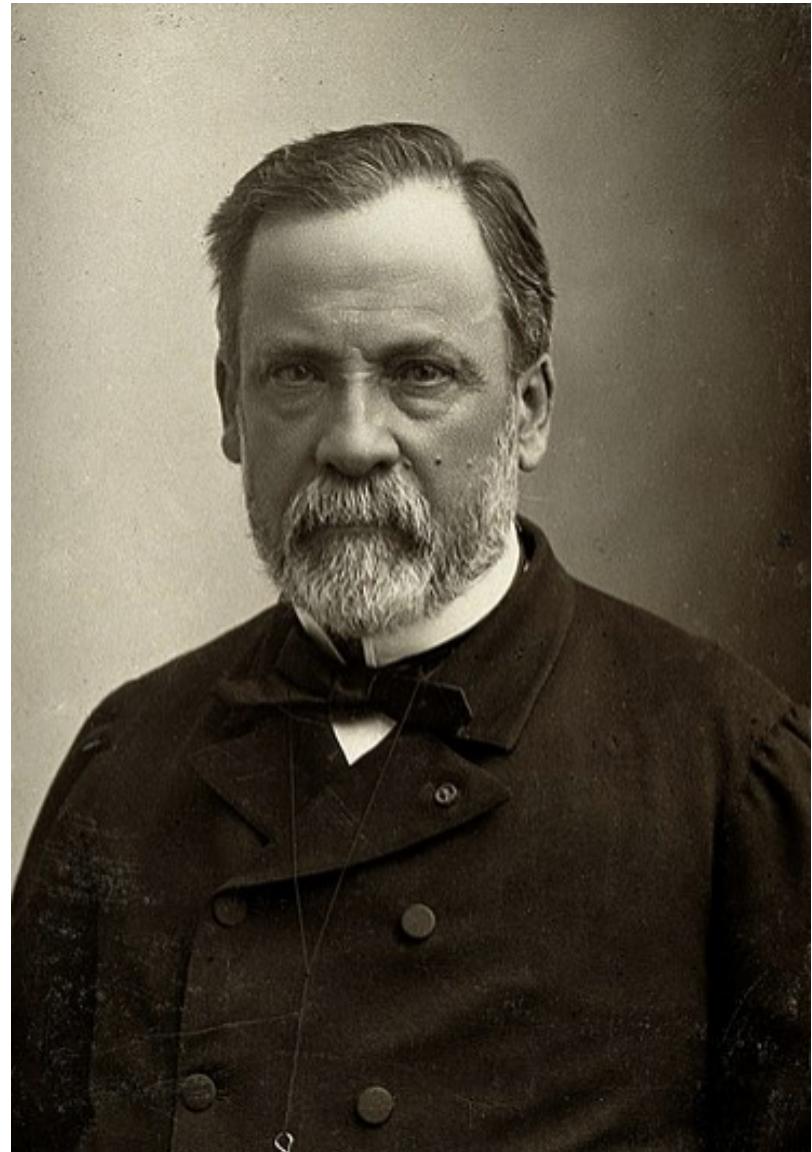


- В XVIII веке от натуральной оспы в России умирал каждый седьмой ребёнок.
- В 1815 году в России был учреждён оспопрививательный комитет
- 10 апреля 1919 г. вышел декрет «Об обязательном оспопрививании»
- прививки против натуральной оспы в СССР прекратились в 1978—1982 годах

"Добрый хранцуз"

- Пастеризация
- Ферментация
- Вакцинация

*«Мы затронули основной принцип вакцинации.
Переболев вирусом в ослабленной форме,
птицы затем не пострадали и после
заражения вирулентным вирусом, и оказались
надежно защищены от птичьей холеры».*



Смоленские крестьяне

- Земцы собрали две тысячи рублей, еще 700 добавил император Александр III с подачи брата Владимира - большого франкофила, который ранее встречался с Пастером в Париже.
- Из 19-ти пациентов 16 выжили и ровно через месяц отправились домой.

- Французы, устыдившись недавнего недоверия и поношения, за считаные недели собрали на строительство и оборудование Института Пастера 2,5 миллиона франков. 100 тысяч из них внес русский царь, наградивший исследователя за спасение подданных орденом Святой Анны.

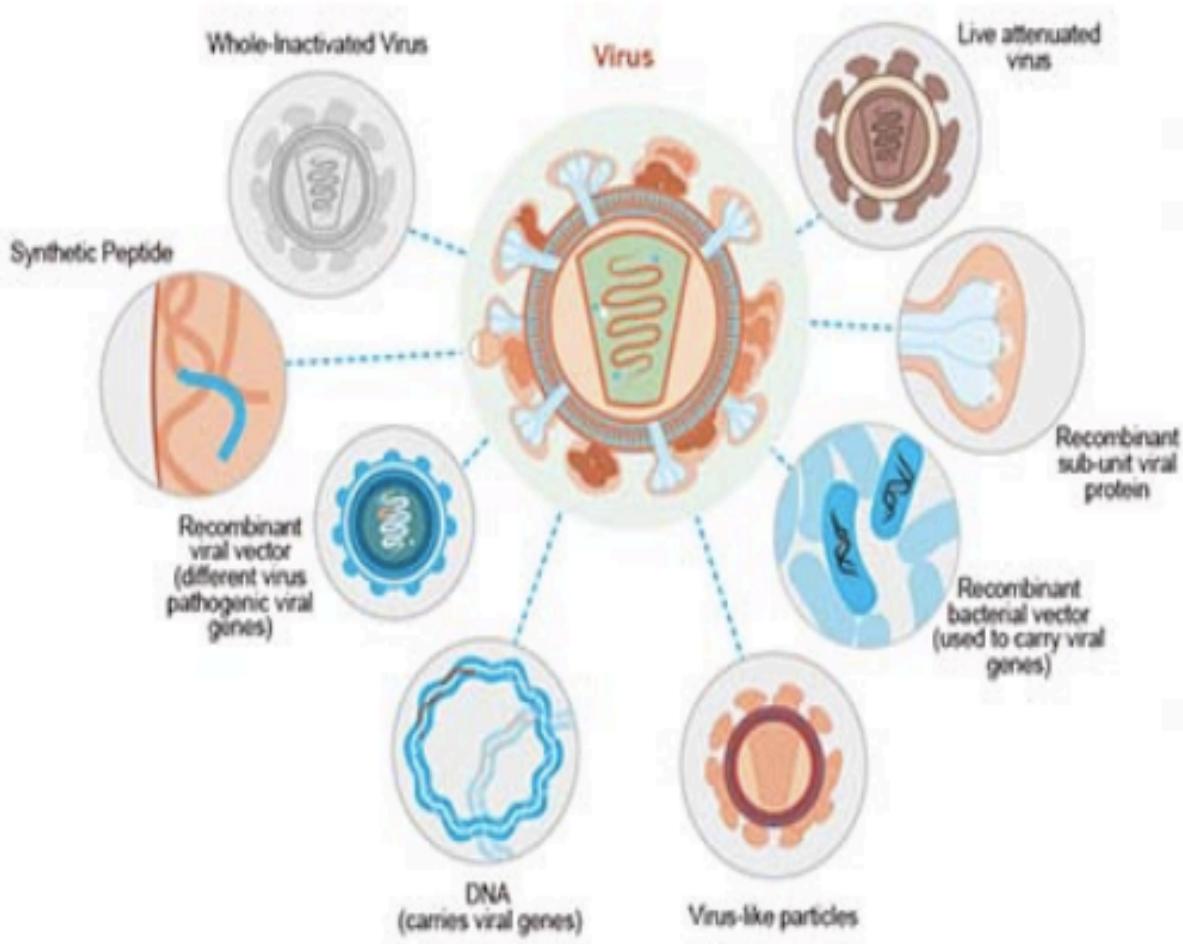
БЦЖ

- *Mycobacterium tuberculosis* – возбудитель человеческого туберкулеза
- *Mycobacterium bovis* – возбудитель бычьего туберкулеза

БЦЖ (бацилла Кальметта—Герена)

- BCG (bacillus Calmette—Guirine) — бацилла Кальметта—Герена (в русскоязычной литературе из-за лингвистического казуса она стала называться БЦЖ)
- Оба — сотрудники института Пастера
- Бацилла бычьего туберкулеза провела 13 лет в крахмале и глицерине и ОСЛАБЛА

Types of Vaccines



Live attenuated (LAV)

- Tuberculosis (BCG)
- Oral polio vaccine (OPV)
- Measles
- Rotavirus
- Yellow fever

Inactivated (killed antigen)

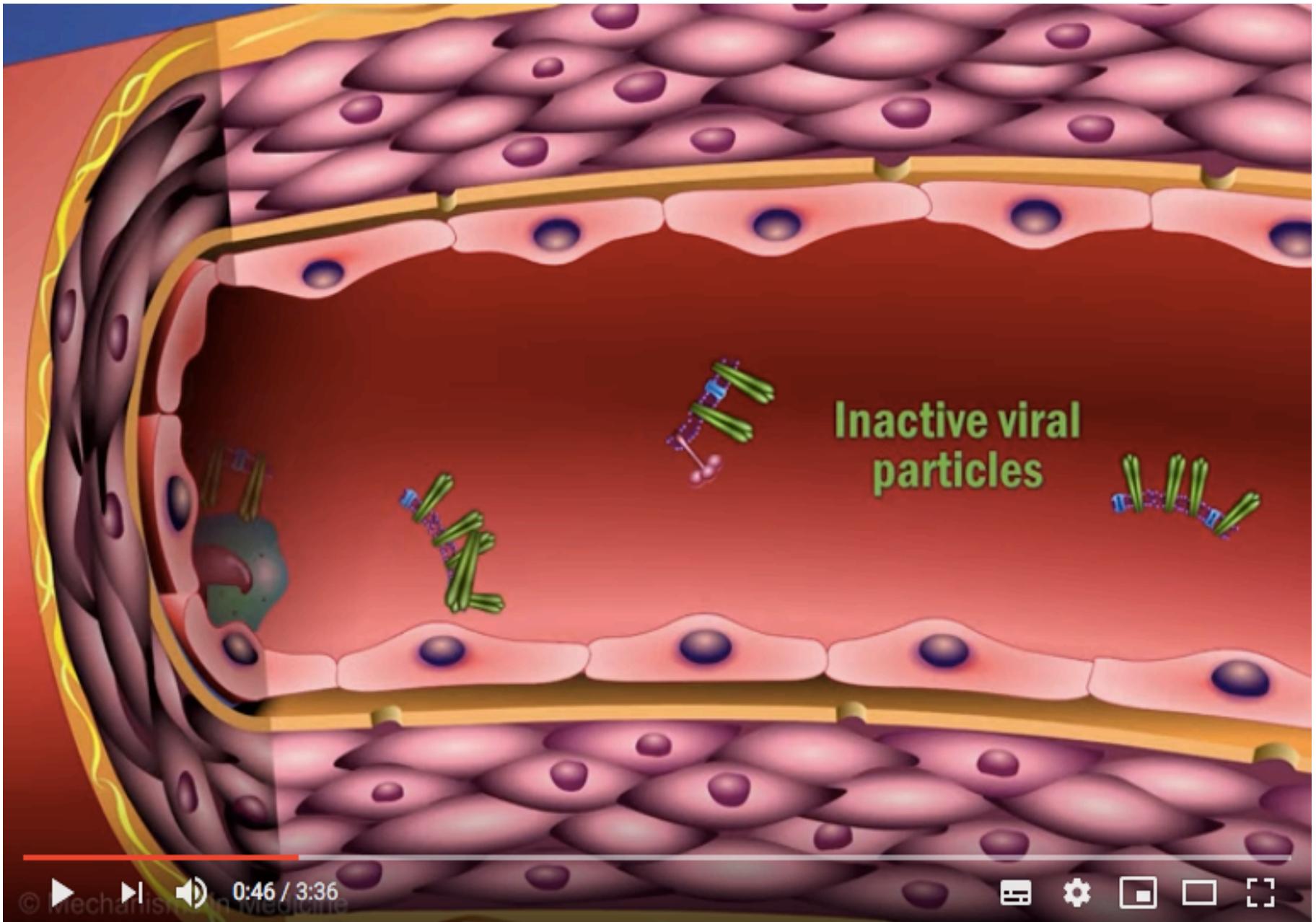
- Whole-cell pertussis (wP)
- Inactivated polio virus (IPV)

Subunit (purified antigen)

- Acellular pertussis (aP).
- *Haemophilus influenzae* type B (Hib).
- Pneumococcal (PCV-7, PCV-10, PCV-13)
- Hepatitis B (HepB)

Toxoid (inactivated toxins)

- Tetanus toxoid (TT),
- Diphtheria toxoid



<https://www.youtube.com/watch?v=Atrx1P2EkiQ>

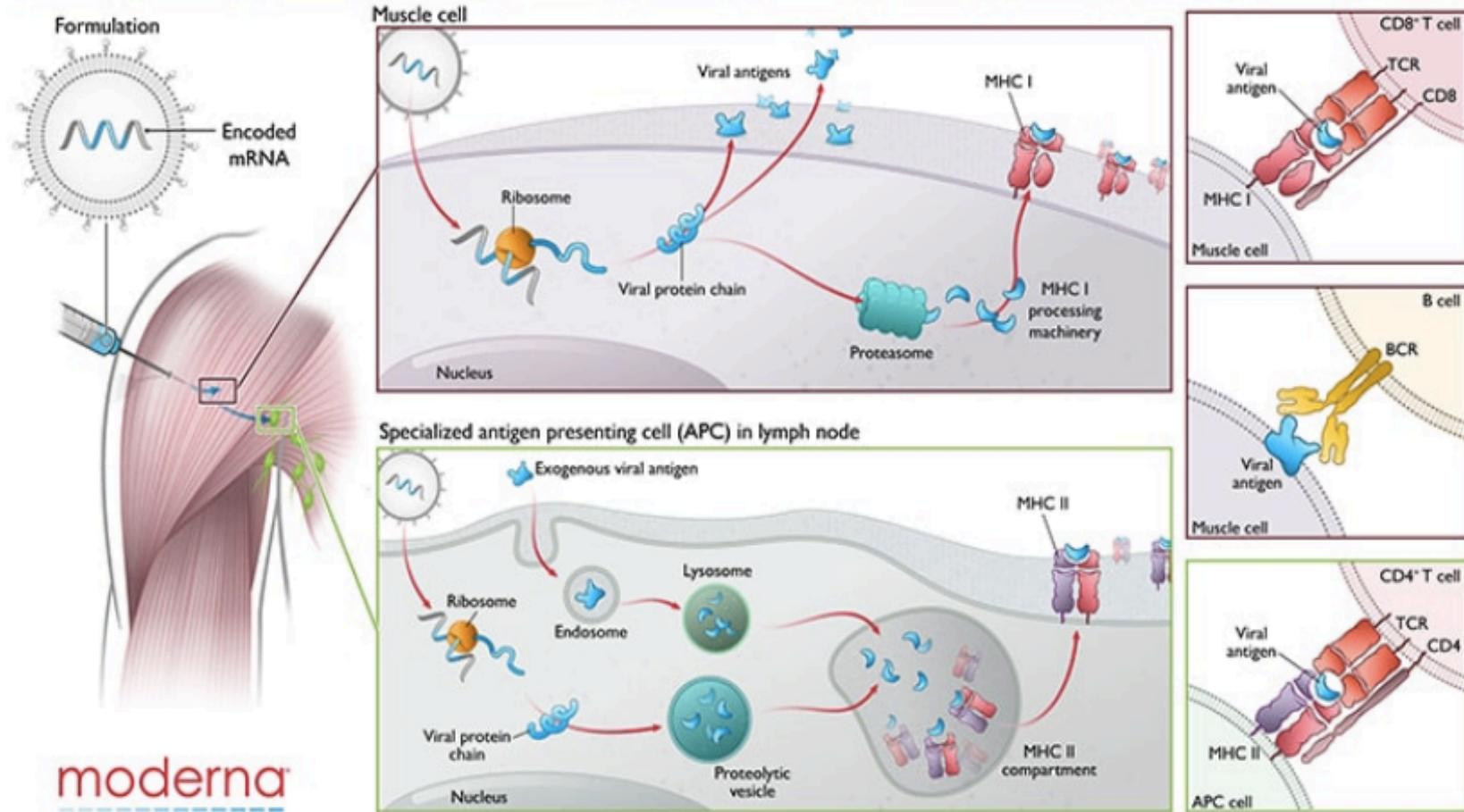
© 2020 Nucleus Medical Media, All Rights Reserved

Затем, антитела присоединяются к вирусу,



Moderna's mRNA Vaccine Approach

Closely mimics a native viral infection leading to B and T cell responses



moderna

© 2017 Moderna Therapeutics