

Орг. вопросы

- Сайт <http://intbio.org/bioinf2018-2019>
- **Вводный опросник** – дедлайн 12 октября 23:59.
- **Форма отчетности:** в этом семестре зачет

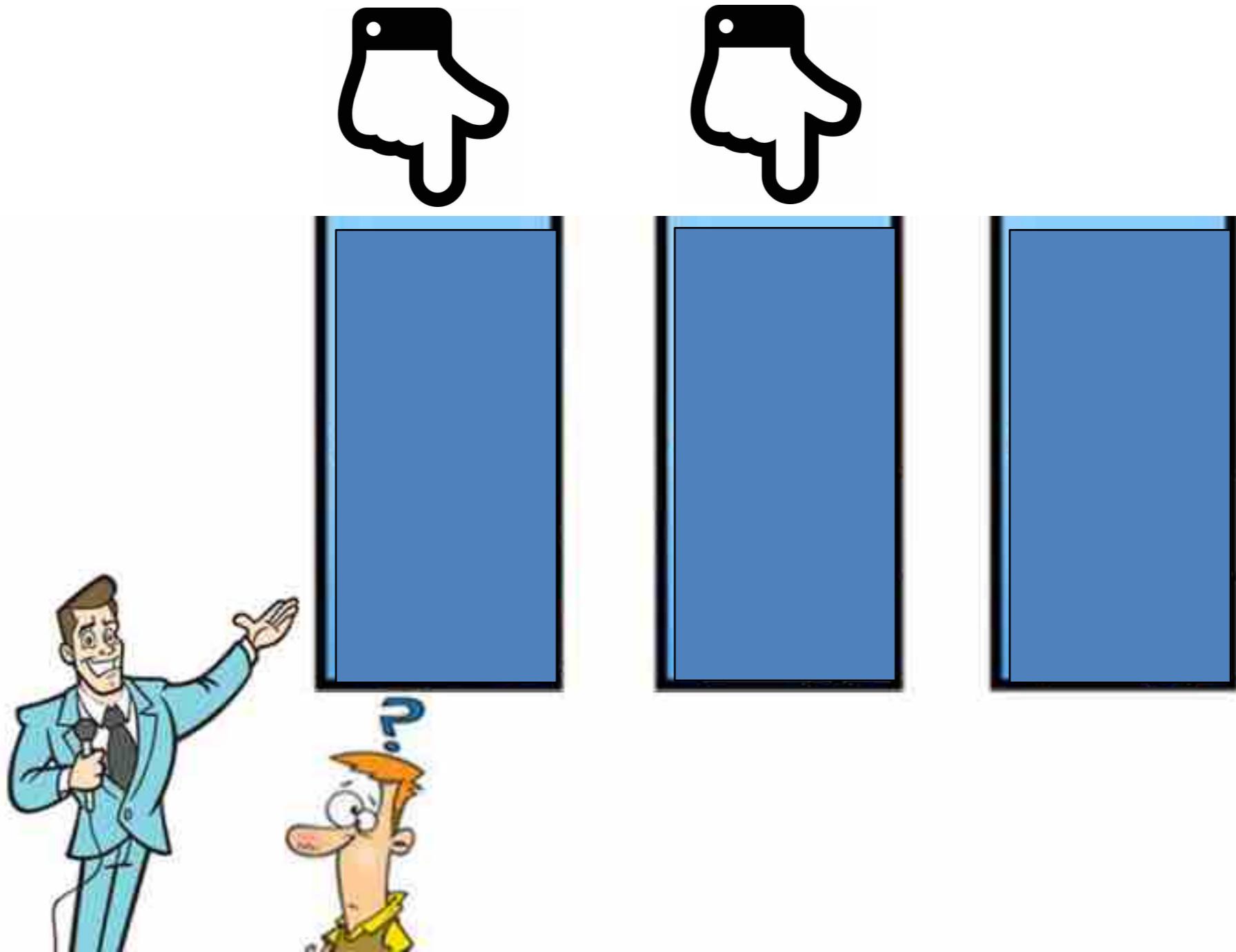
ВВЕДЕНИЕ В БИОИНФОРМАТИКУ

Лекции №2-3

Природа, передача и хранение
информации. Базы данных. Биологические
БД. Базы данных NCBI.

Алексей Константинович Шайтан, к.ф.-м.н.

Сайт курса: <http://intbio.org/bioinf2018-2019>



- В результате игрок получает **0.67 бита информации**
- Если бы ведущий открыл дверь в самом начале – только **0.58 бита информации**

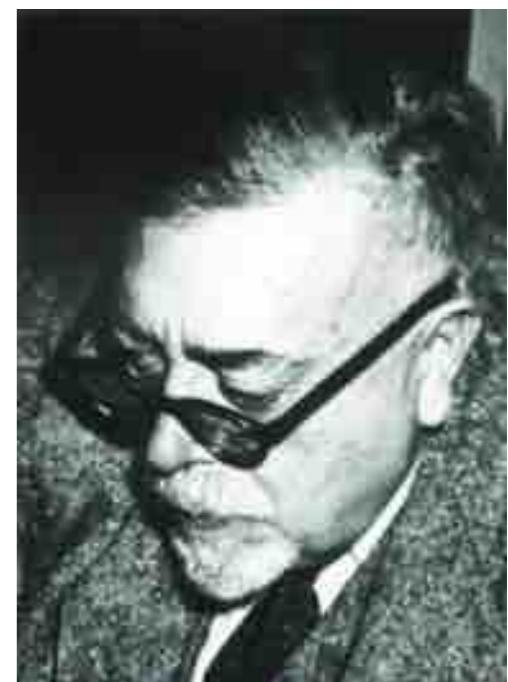
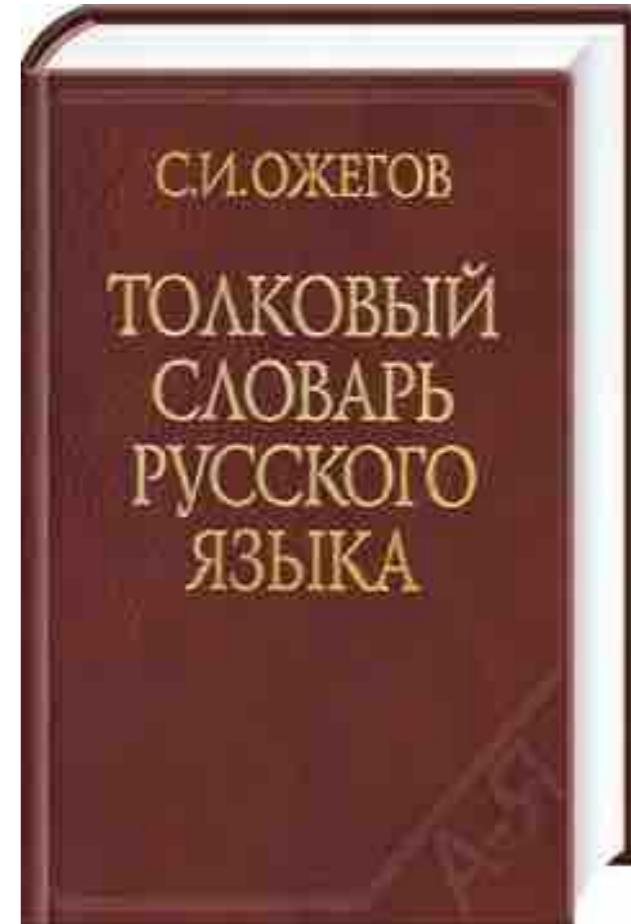
Природа информации

Информация

ИНФОРМАЦИЯ, 1. Сведения об окружающем мире и протекающих в нем процессах, воспринимаемые человеком или специальным устройством.



... универсального определения информации не только нет, но и быть не может из-за широты этого понятия.



Н.Н. Моисеев

Information is information, not matter or energy.

Norbert Wiener

Информация



[iPhone 7 Plus](#)

[Серебристый](#)

[Ёмкость](#)

Теперь выберите ёмкость.

32 ГБ¹

52 990,00 руб.

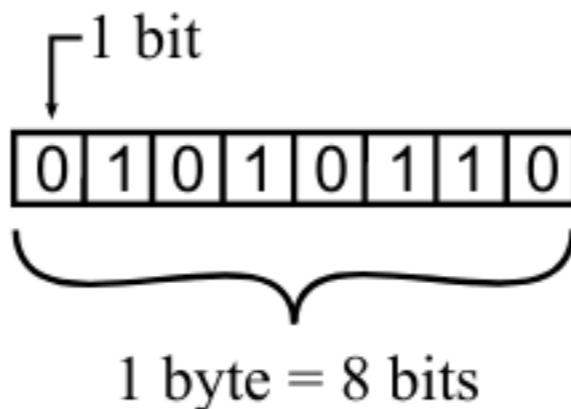
128 ГБ¹

60 990,00 руб.

Доставка: на складе

Доставка: на складе

Измерение информации



Измерения в байтах				
ГОСТ 8.417—2002		Приставки СИ		
Название	Обозначение	Степень	Название	Степень
байт	Б	10^0	-	10^0
килобайт	кбайт	10^3	кило-	10^3
мегабайт	Мбайт	10^6	мила-	10^6
гигабайт	Гбайт	10^9	гига-	10^9
терабайт	Тбайт	10^{12}	тера-	10^{12}
петабайт	Пбайт	10^{15}	пета-	10^{15}
эксабайт	Эбайт	10^{18}	экса-	10^{18}
зеттабайт	Збайт	10^{21}	зетта-	10^{21}
иоттабайт	Ибайт	10^{24}	иотта-	10^{24}

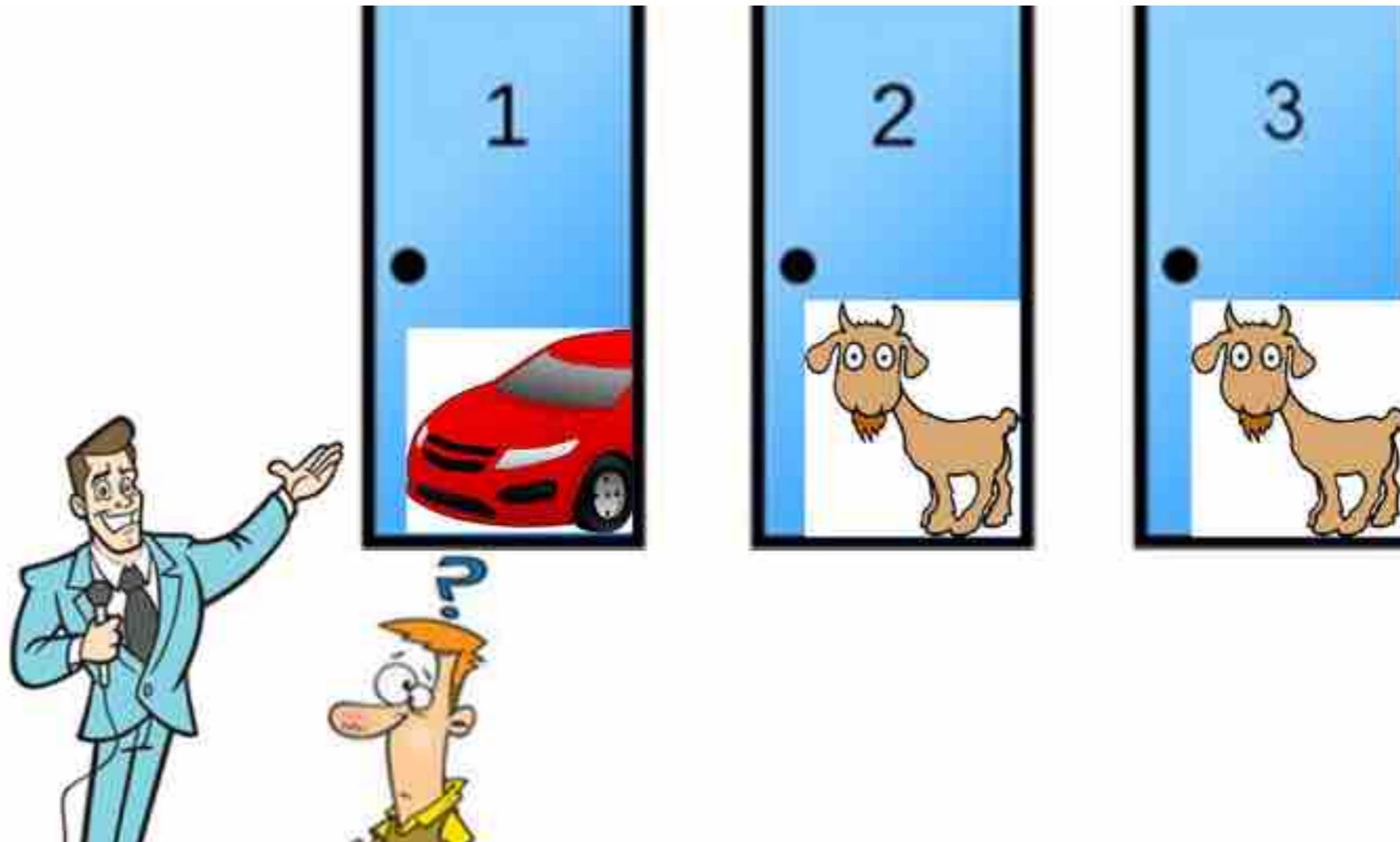
Decimal	Binary
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
9	1001

Вероятность того,
что машина за
этой дверью:

1/3

1/3

1/3

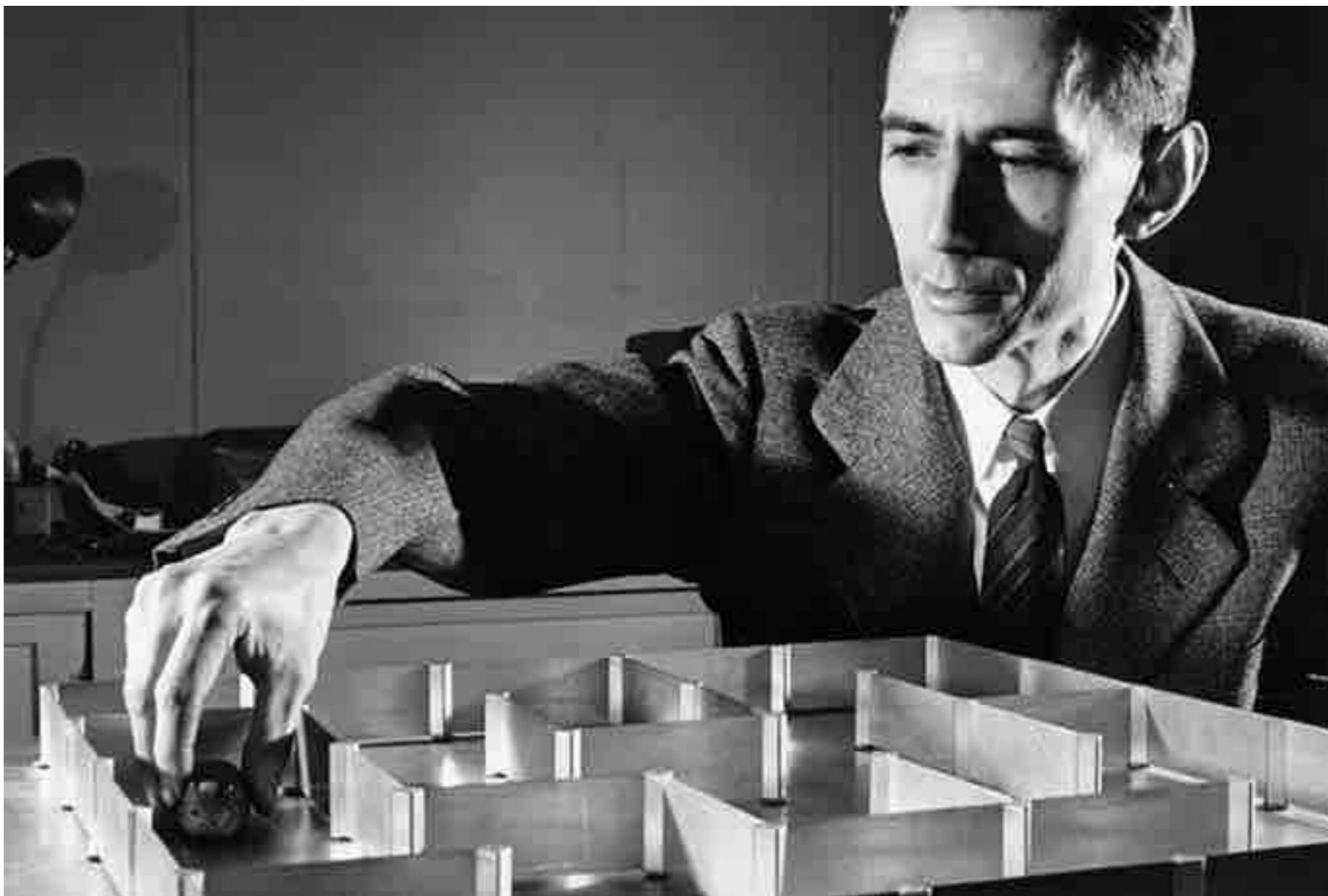


- Сколько информации нужно, чтобы закодировать положение машины?

1.5849625007211563... бит

Теория информации

“the father of information theory”



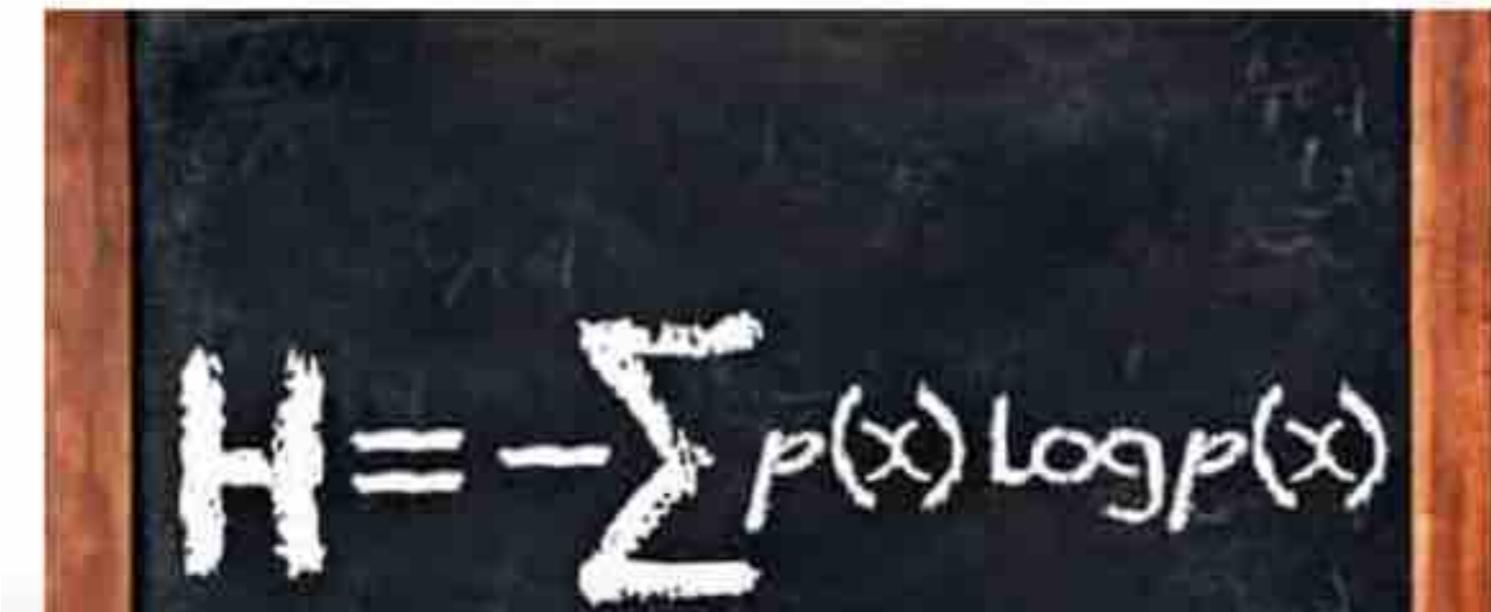
Claude Elwood Shannon
(April 30, 1916 – February 24, 2001)

Теория информации

Science A short history of equations

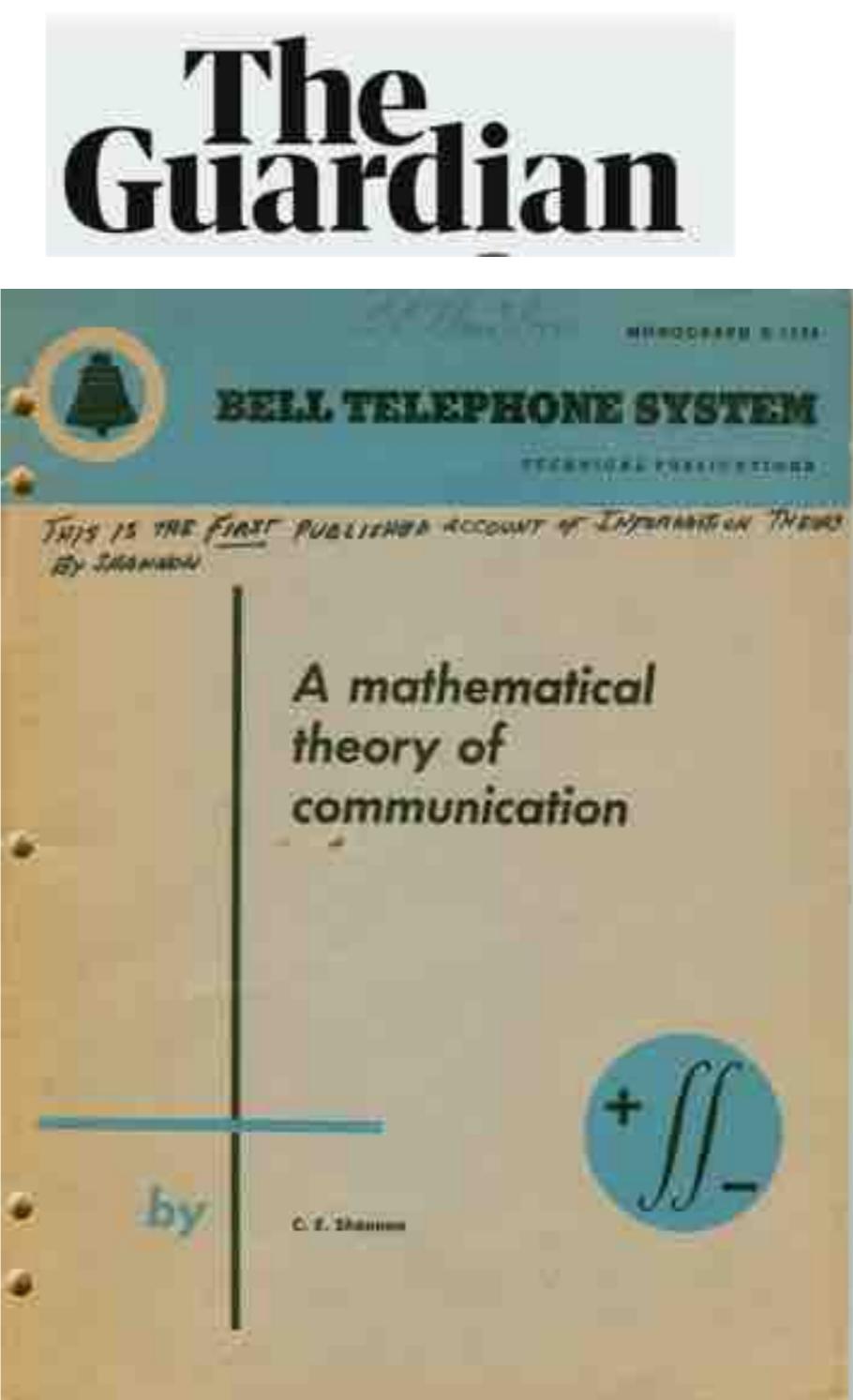
Without Claude Shannon's information theory there would have been no internet

It showed how to make communications faster and take up less space on a hard disk, making the internet possible



Введена мера информации(!)

КГ, метр, секунда + БИТ



1948

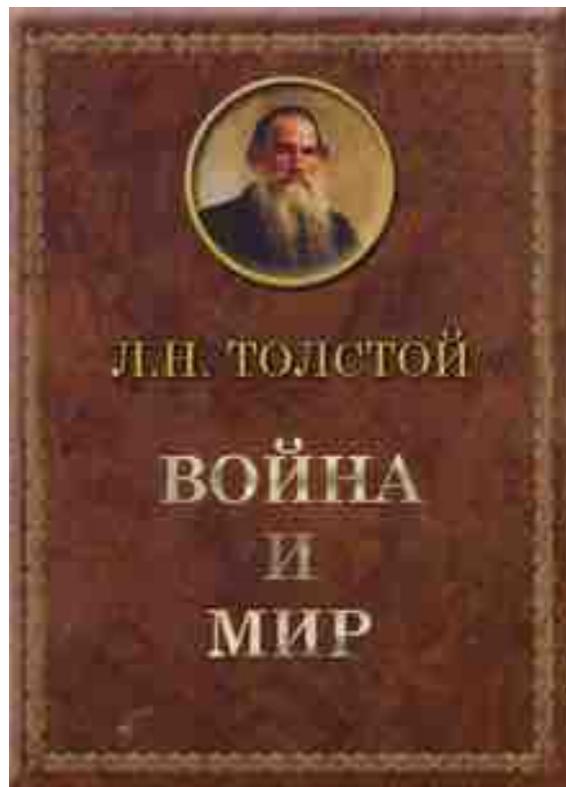
Информационная энтропия

[Клод Шеннон](#) предположил, что прирост информации равен утраченной неопределённости, и задал требования к её измерению:

1. мера должна быть непрерывной; то есть изменение значения величины вероятности на малую величину должно вызывать малое результирующее изменение функции;
2. в случае, когда все варианты (буквы в приведённом примере) равновероятны, увеличение количества вариантов (букв) должно всегда увеличивать значение функции;
3. должна быть возможность сделать выбор (в нашем примере букв) в два шага, в которых значение функции конечного результата должно являться суммой функций промежуточных результатов.

$$H = -\sum p(x) \log p(x)$$

Сжатие информации



Tom 1 2.txt 1.5 MB
Modified: Today, 4:01 PM
Add Tags...

▼ General:
Kind: Plain Text Document
Size: 1,473,547 bytes (1.5 MB on disk)

Tom 1 2.txt.zip 602 KB
Modified: Today, 4:02 PM
Add Tags...

▼ General:
Kind: ZIP archive
Size: 602,098 bytes (602 KB on disk)

Буква в тексте не независимы – одни встречаются чаще других в разных контекстах.

Сжатие информации

Кодирование длин серий

```
WWWWWWWWWWWWBWWWWWWWWWWWWBBBWWWWWWWWWWWWWWWWWWWWWWWWBWWWWWWWWWWWWWWWWWW
```

Посчитаем количество повторяющихся символов:

1. 12 символов «W»;
2. 1 символ «B»;
3. 12 символов «W»;
4. 3 символа «B»;
5. 24 символа «W»;
6. 1 символ «B»;
7. 14 символов «W».

Итого найдено 7 серий. Заменим серии на число повторов и сам повторяющийся символ:

```
12W1B12W3B24W1B14W
```

Информационная энтропия

$$H = -\sum p(x) \log p(x)$$



$$H = -\frac{1}{3} \log_2 \frac{1}{3} - \frac{1}{3} \log_2 \frac{1}{3} - \frac{1}{3} \log_2 \frac{1}{3} = -\log_2 \frac{1}{3}$$

1.5849625007211563... бит

Энтропия

Информационная и физическая энтропии
имеют глубинную связь



$$H \stackrel{\text{def}}{=} \int P(\ln P) d^3v = \langle \ln P \rangle$$

В [термодинамике](#) и [кинетической теории](#), Н -
теорема, полученная [Больцманом](#) в [1872 году](#),
[описывает неубывания](#) [энтропии идеального газа](#) в
необратимых процессах,
исходя из [уравнения Больцмана](#).

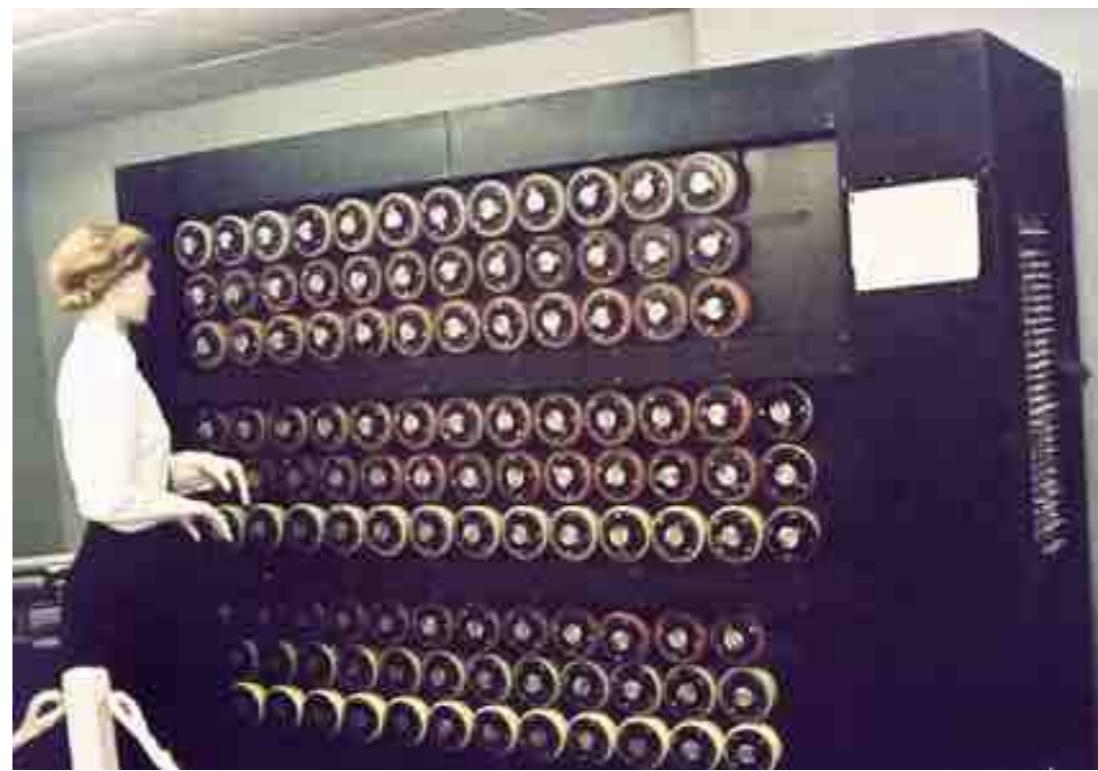
$$S \stackrel{\text{def}}{=} -NkH,$$

Ludwig Eduard Boltzmann

Передача информации

Передача информации

- Комплексная область: **теоретические, практические, физические аспекты**
- Вопросы сжатия данных
- Вопросы надежности
- Вопросы **шифрования и защиты данных**
(особенно в медицине и биологии)



Передача информации

Связь частоты сигнала и пропускной способности

Владимир Александрович
Котельников



Harry Nyquist

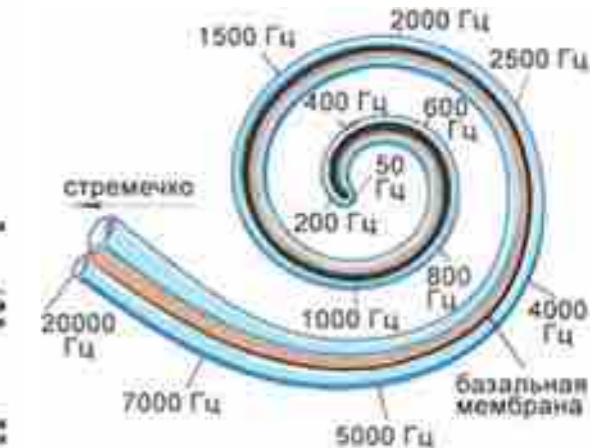
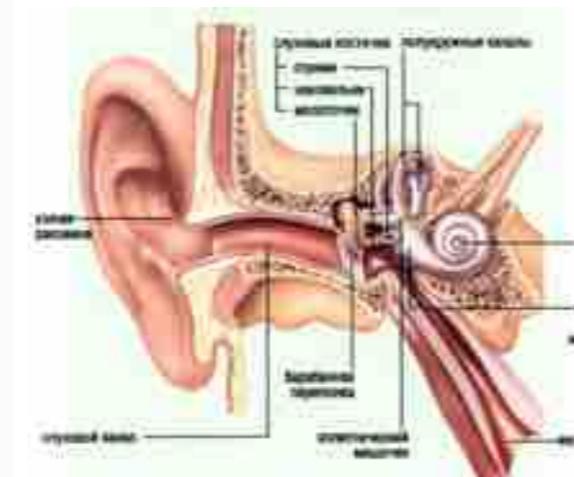


[1908 - 2005](#)

1889 – 1976

Теорема Котельникова-(Найквиста-Шенона)

«любую функцию $F(t)$, состоящую из частот от 0 до f , можно непрерывно передавать с любой точностью при помощи чисел, следующих друг за другом через $1/(2f)$ секунд



44.1 кГц – частота дискретизации при записи звука

Передача информации

Связь частоты сигнала и пропускной способности

Ralph Hartley



$$C = B \log_2 \left(1 + \frac{S}{N} \right),$$

где

C — пропускная способность канала, бит/с;

B — полоса пропускания канала, Гц;

S — полная мощность сигнала над полосой пропускания, Вт или В^2 ;

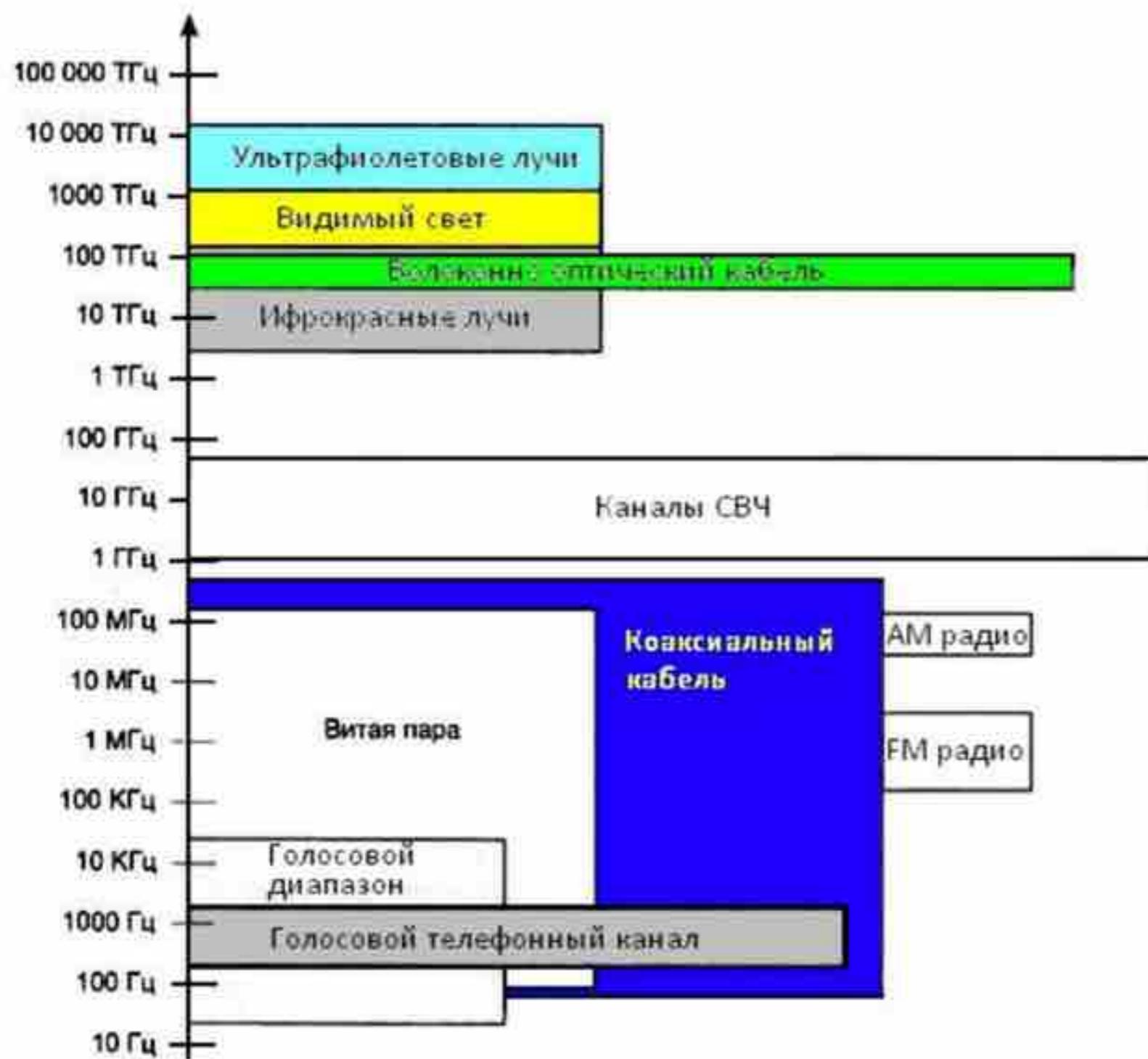
N — полная шумовая мощность над полосой пропускания, Вт или В^2 ;

S/N — отношение мощности сигнала к шуму (SNR).

1888 – 1970

Теорема Шеннона-Хартли

Передача информации

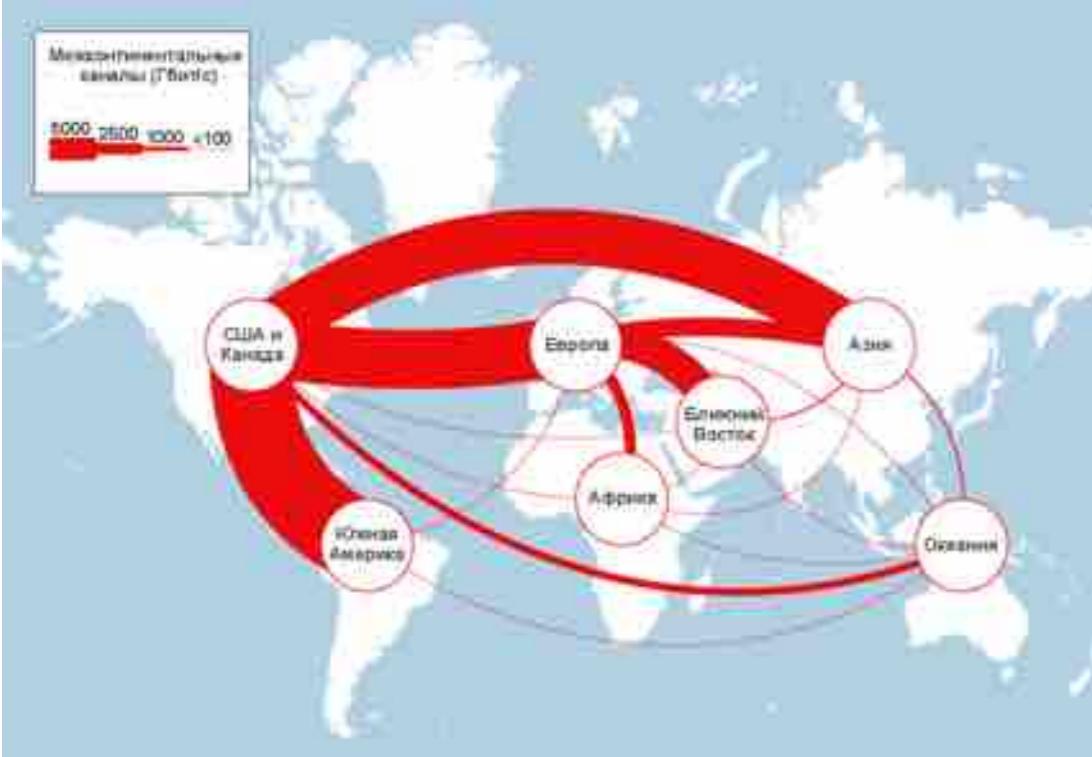


Оптоволокно

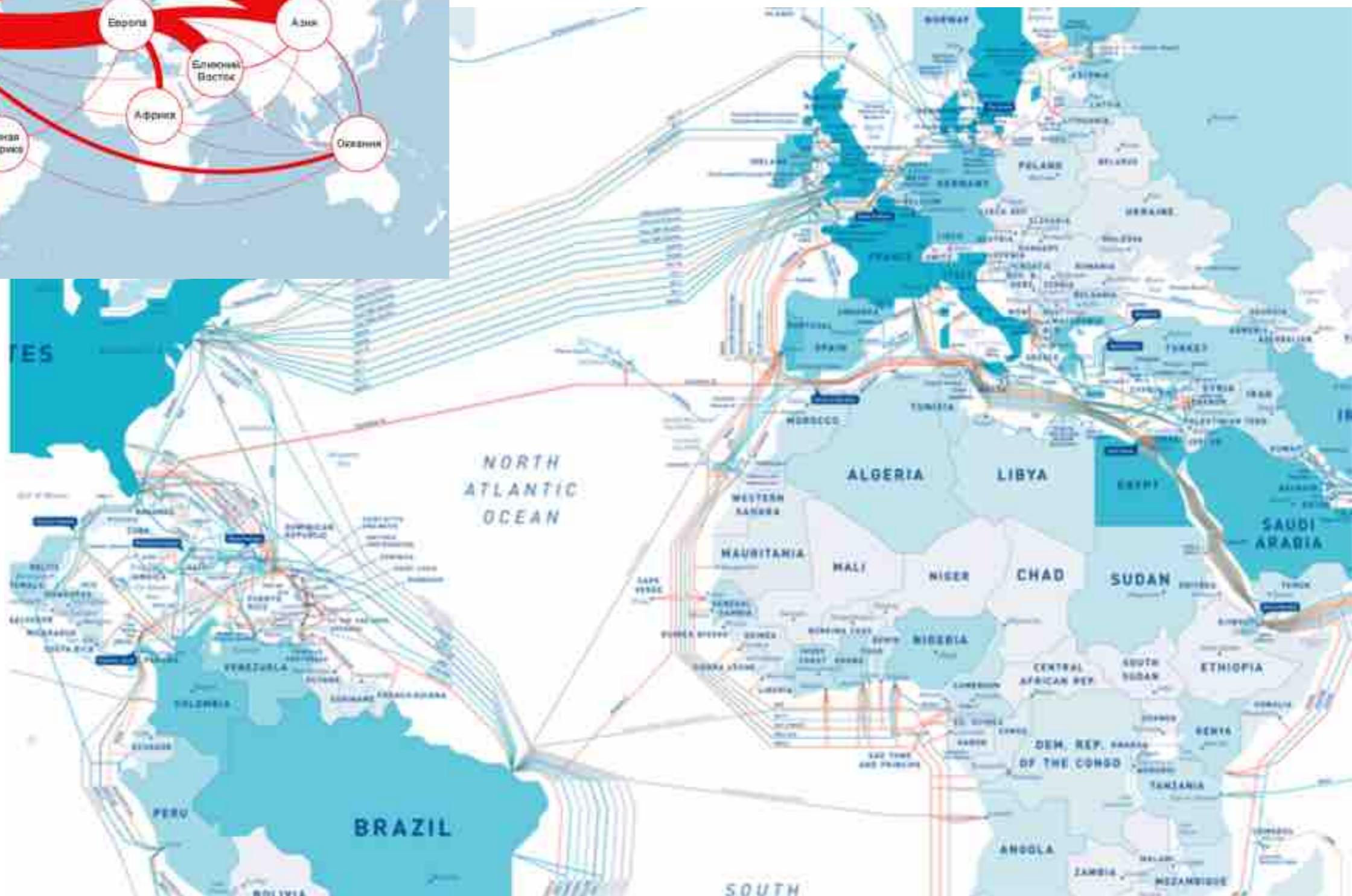


Рис. 1. Полосы пропускания линий связи и популярные частотные диапазоны

Антенны КНЧ



Каналы связи

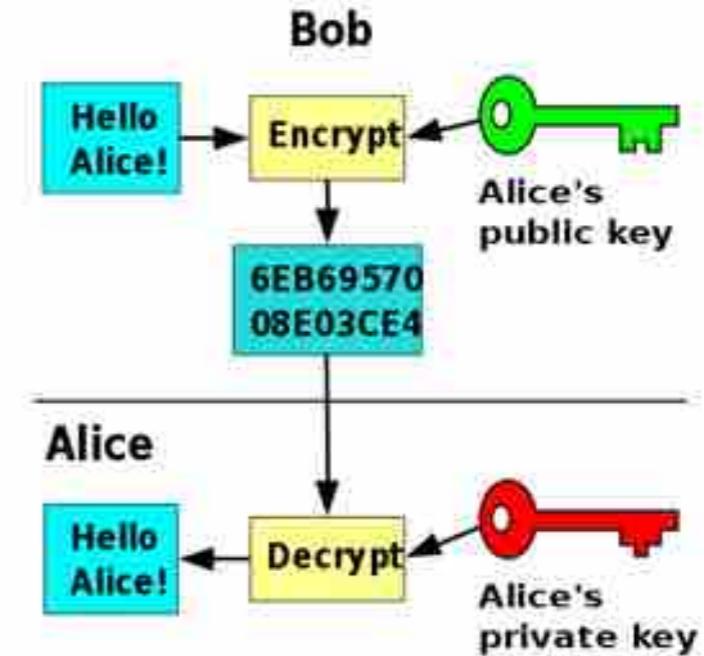
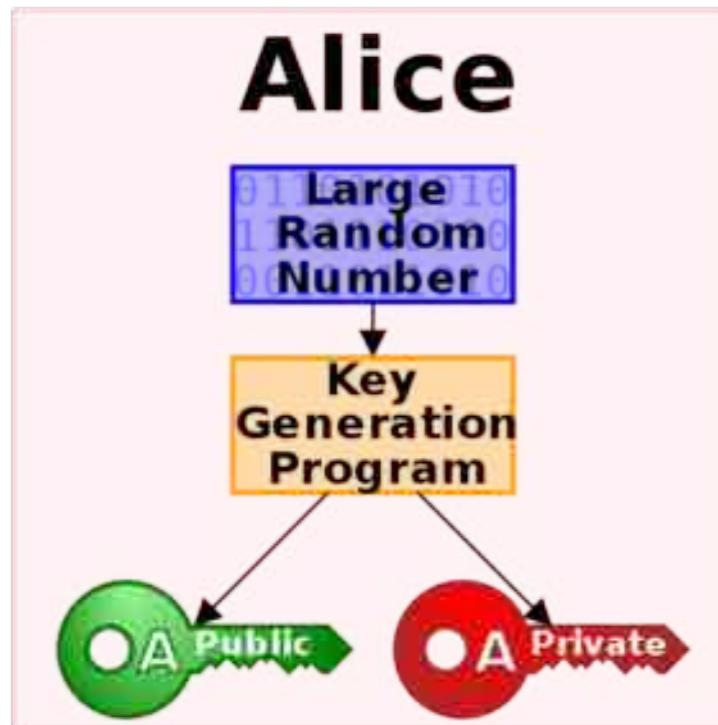


Карта подводных кабелей

Шифрование информации



Криптосистемы с открытым ключом



Необратимая Хэш функция

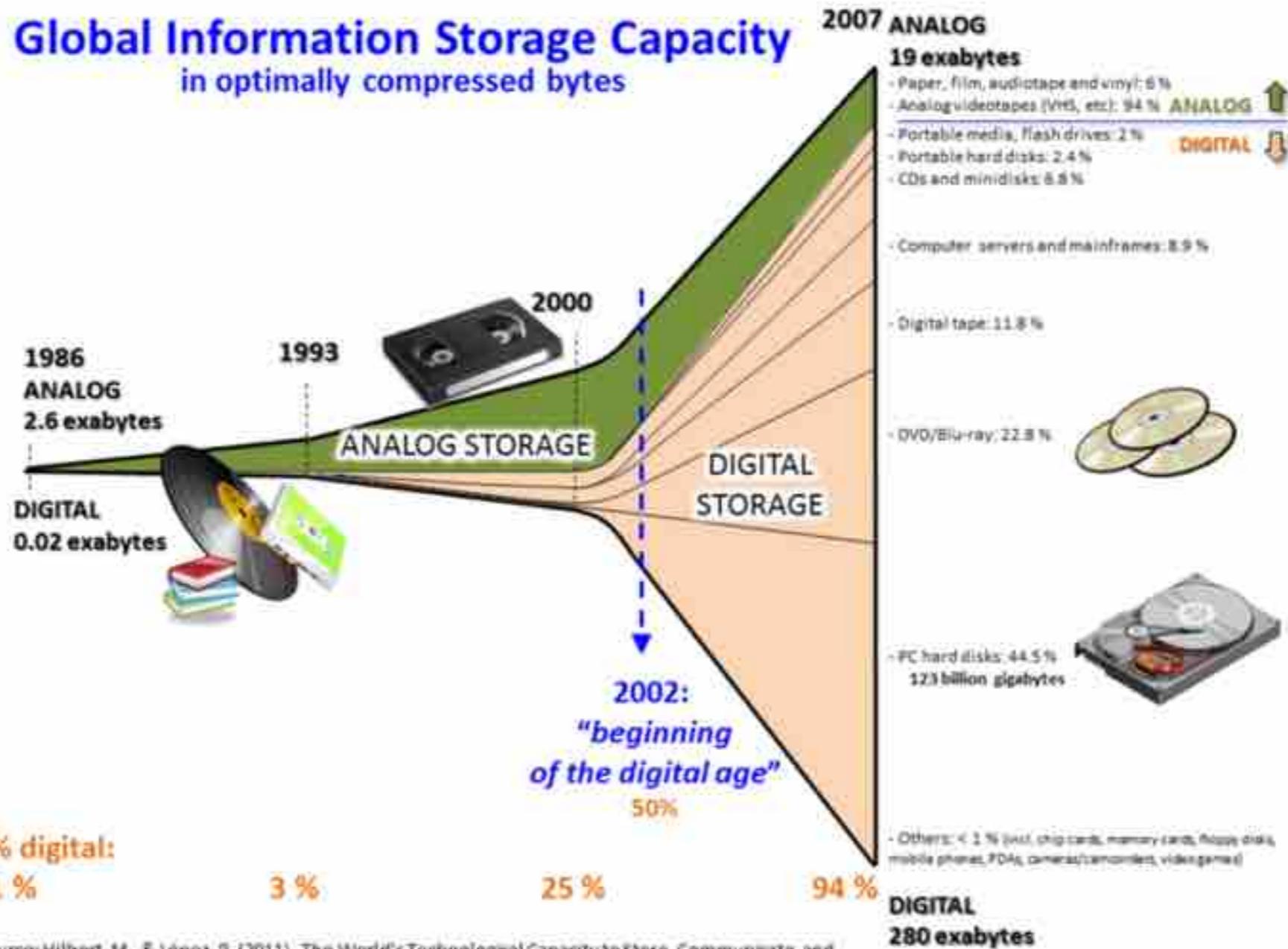
```
mbptb:~ alexsha$ md5 -s 'Hello world!!!'  
MD5 ("Hello world!!!") = 87ee732d831690f45b8606b1547bd09e
```

Хранение информации

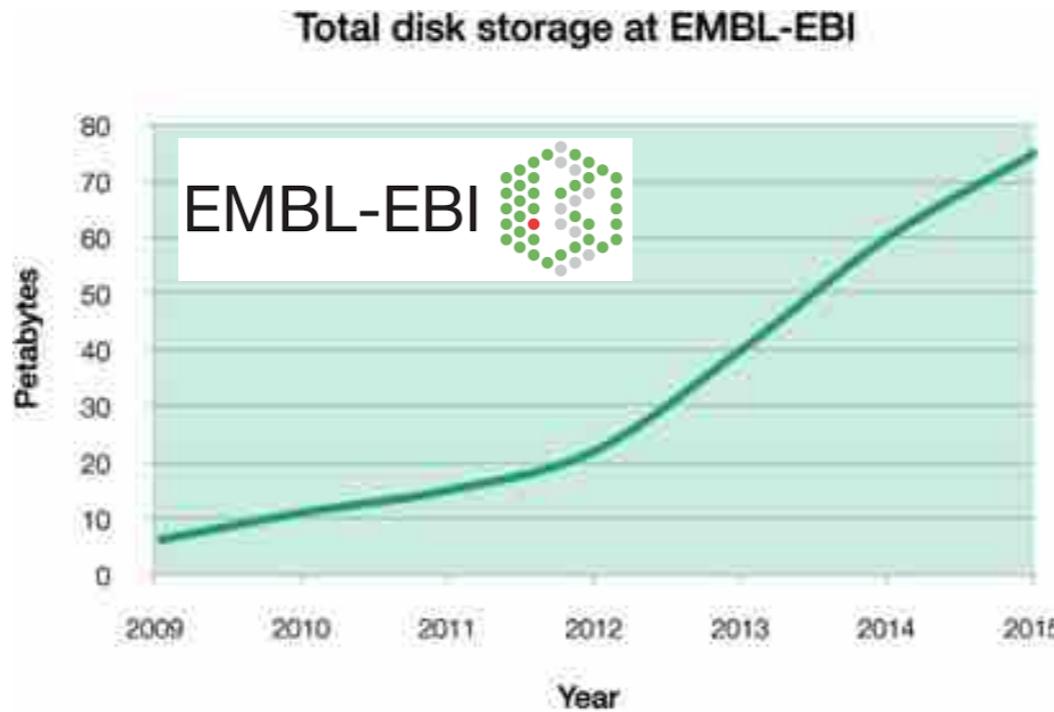
Хранение информации



Хранение информации



Большие данные в биомедицине



2013-2021
~\$400 МЛН

Table 1. Four domains of Big Data in 2025. In each of the four domains, the projected annual storage and computing needs are presented across the data lifecycle.

Data Phase	Astronomy	Twitter	YouTube	Genomics
Acquisition	25 zetta-bytes/year	0.5–15 billion tweets/year	500–900 million hours/year	1 zetta-bases/year
Storage	1 EB/year	1–17 PB/year	1–2 EB/year	2–40 EB/year
Analysis	In situ data reduction	Topic and sentiment mining	Limited requirements	Heterogeneous data and analysis
	Real-time processing	Metadata analysis		Variant calling, ~2 trillion central processing unit (CPU) hours
	Massive volumes			All-pairs genome alignments, ~10,000 trillion CPU hours
Distribution	Dedicated lines from antennae to server (600 TB/s)	Small units of distribution	Major component of modern user's bandwidth (10 MB/s)	Many small (10 MB/s) and fewer massive (10 TB/s) data movement

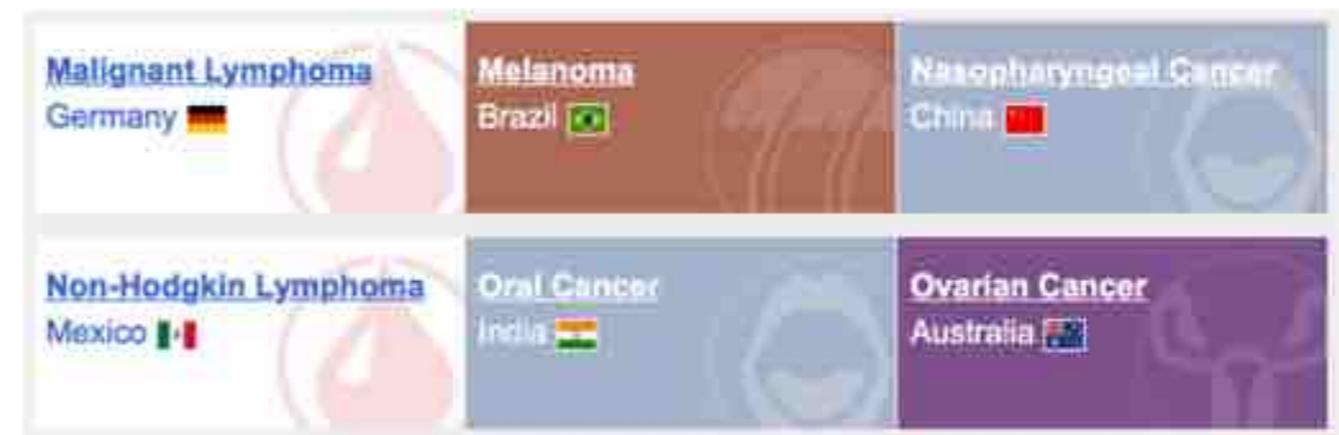
Источники больших данных в биомедицине

- **Омиксные технологии**
 - Секвенирование, геномика, транскриптомика, протеомика, метаболомика и т.д.
 - Коннектом мозга
- **Медицинская информация**
 - Электронные медицинские карты, результаты клинических исследований и т.д.
 - Медицинские изображения, МРТ и т.д.
- **Структурная биология и моделирование**
 - Данные с лазеров на свободных электронах (XFEL)
 - Моделирование структуры и динамики белков.

Данные секвенирования, пример Геномы раковых опухолей



Геном человека ~ 3.3 Gb
x100 секвенирование ~300Gb

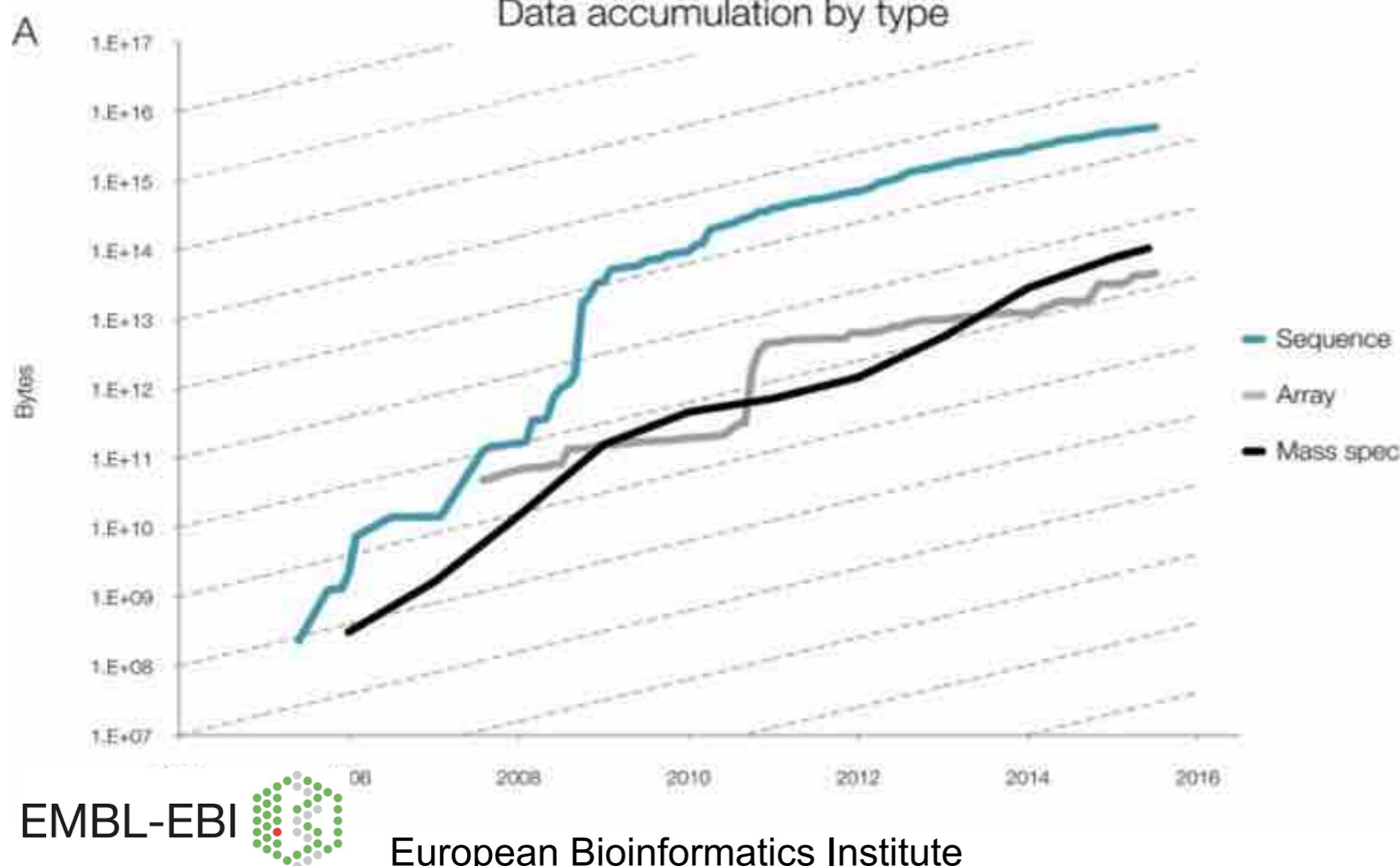


25000 образцов опухолей

Международный проект, данные распределены по миру

File ID	Donor	Repository	Project	Study	Data Type	Strategy	Format	Size	URL
<input type="checkbox"/> R9995	DO217962	PCAWG - London, PCAWG - Barcelona, Collaboratory - Toronto, EGA - Hinxton	BRCA-EU	PCAWG	Aligned Reads	WGS	BAM	128.72 GB	Link
<input type="checkbox"/> R9994	DO217962	PCAWG - London, PCAWG - Barcelona, Collaboratory - Toronto, EGA - Hinxton	BRCA-EU	PCAWG	Aligned Reads	WGS	BAM	107.27 GB	Link
<input type="checkbox"/> R9974	DO46390	PCAWG - London, PCAWG - Barcelona, EGA - Hinxton, Collaboratory - Toronto, AWS - Virginia	OV-AU	PCAWG	Aligned Reads	WGS	BAM	134.05 GB	Link
<input type="checkbox"/> R9973	DO46390	PCAWG - London, PCAWG - Barcelona, Collaboratory - Toronto, AWS - Virginia, EGA - Hinxton	OV-AU	PCAWG	Aligned Reads	WGS	BAM	101.45 GB	Link
<input type="checkbox"/> R9956	DO222303	PCAWG - Chicago (TCGA), PDC - Chicago	DLBC-US	PCAWG	Aligned Reads	WGS	BAM	202.00 GB	Link
<input type="checkbox"/> R9955	DO222303	PCAWG - Chicago (TCGA), PDC - Chicago	DLBC-US	PCAWG	Aligned Reads	WGS	BAM	100.77 GB	Link

Централизованные репозитории омиксных данных



华大基因
BGI

Genomes en masse



5 years ~ 100 000
genomes

nature International weekly journal of science

Home | News & Comment | Research | Careers & Jobs | Current Issue | Archive | Audio & Video | For Authors

Archive > Volume 532 > Issue 7600 > News > Article

NATURE | NEWS

AstraZeneca launches project to sequence 2 million genomes

Drug company aims to pool genomic and medical data in hunt for rare genetic sequences associated with disease.

Heidi Ledford

22 April 2016



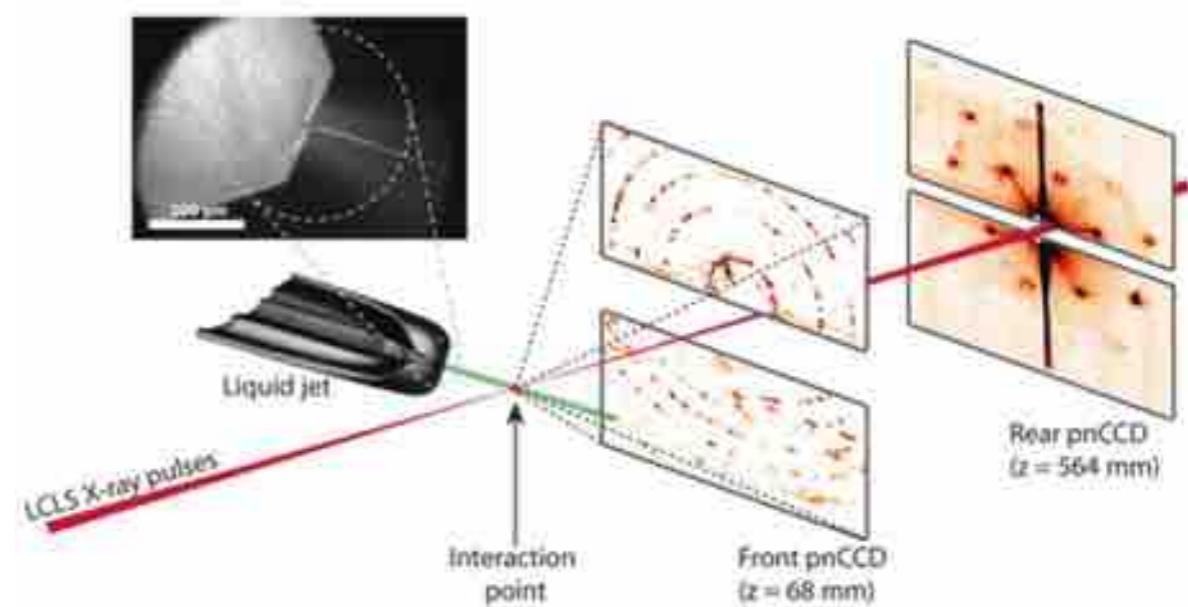
华大基因
BGI



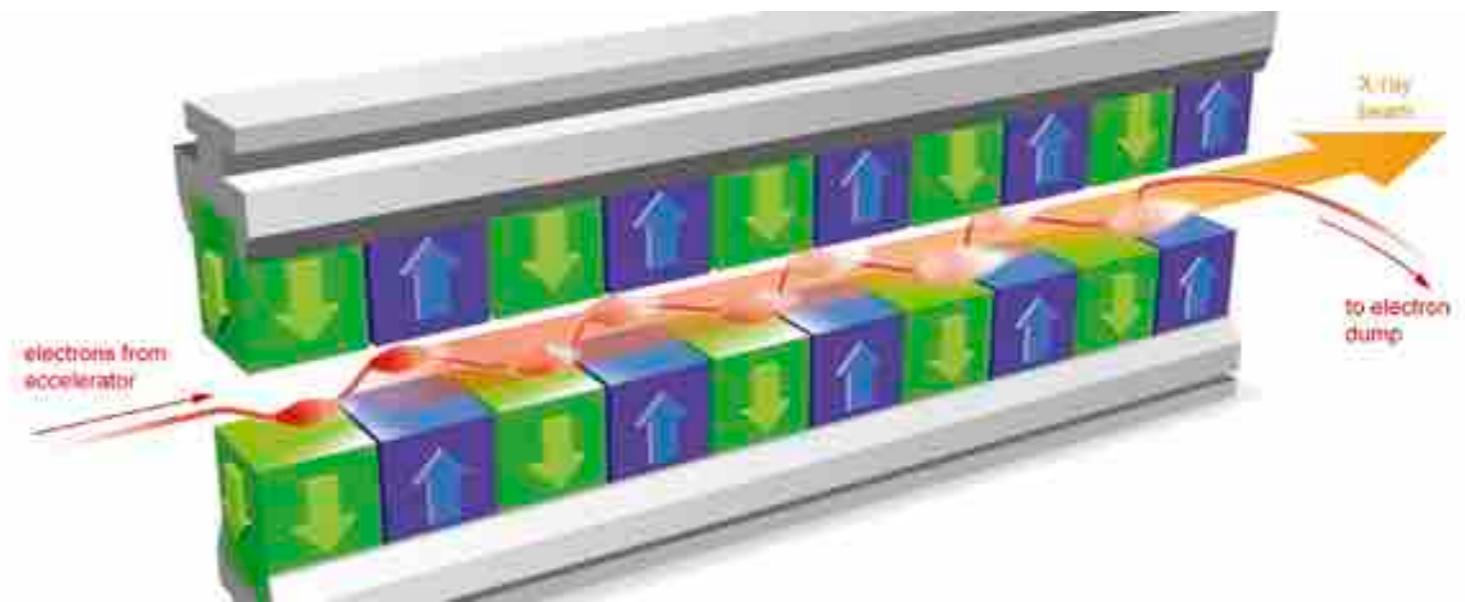
Структурная биология и моделирование



European XFEL, Hamburg



27000 импульсов в секунду



Передача больших данных

Выделенные научно-образовательные сети 100Gbit/s

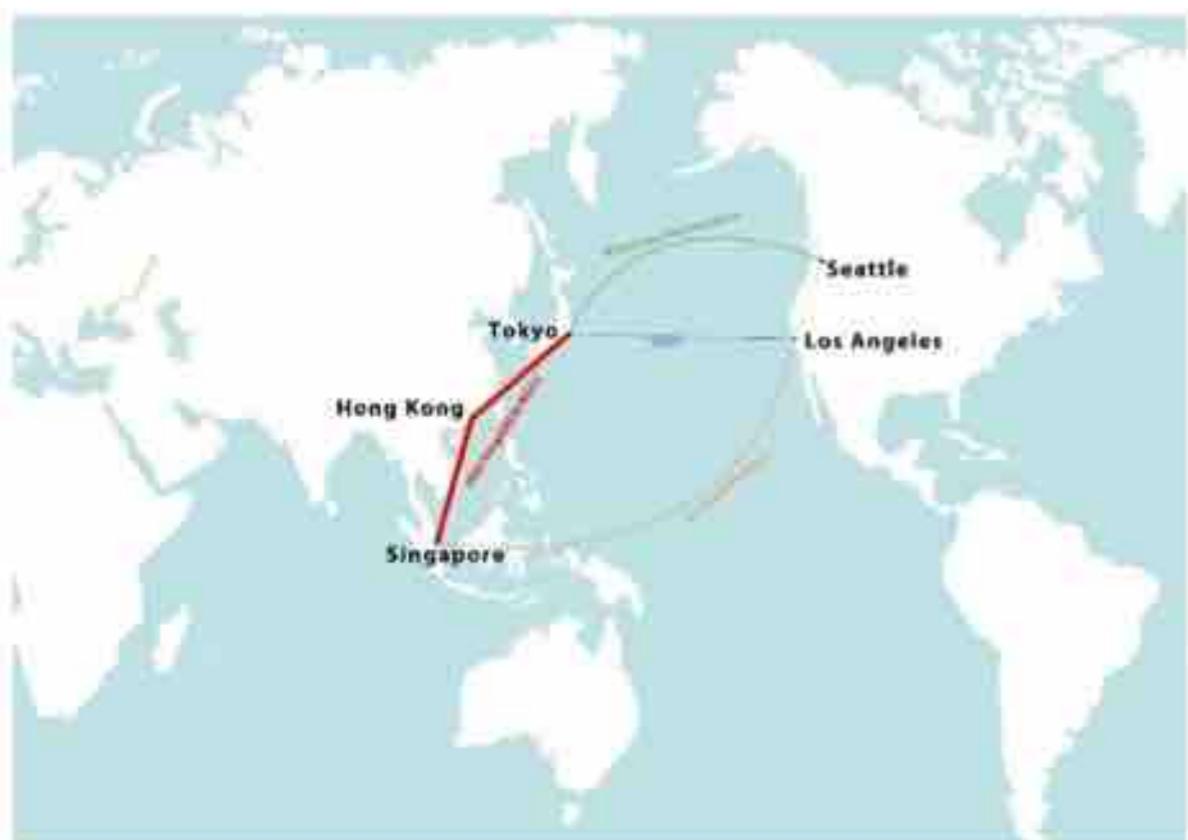


Figure 1: Asia Pacific Ring (APR)



Программные решения

 **aspera**
an IBM® company

 **globus toolkit**

Базы данных

Базы данных

- Реляционные базы данных, объектно-ориентированные, RDF
- Системы управления базами данных СУБД
- Языки и стандарты SQL, SPARQL, RDF



Реляционные базы данных

Клиент

Id_кл	Фамилия	Имя	Отчество
15	Иванов	Иван	Иванович
16	Петров	Петр	Петрович
17	Николаев	Николай	Николаевич

Товар

Id_тov	Название
1	Шкаф
2	Стул
3	Стол

Заказ

Id_зак	Клиент	Товар	Дата	Количество
1	15	1	15.09.2003	1
2	17	1	17.09.2003	2
3	15	2	20.09.2003	12

SQL

**Целостность
данных**

Транзакции

**Соответствие
требованиям ACID**

Атомарность

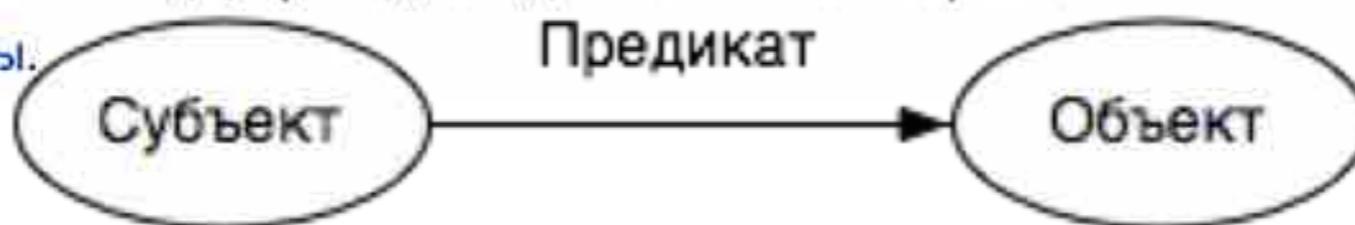
Изолированность

Надежность

Resource Description Framework

Язык SPARQL

Resource Description Framework (RDF, «среда описания ресурса»^[1]) – это разработанная консорциумом Всемирной паутины модель для представления данных, в особенности – метаданных^[2]. RDF представляет утверждения о ресурсах в виде, пригодном для машинной обработки. RDF является частью концепции семантической паутины.



```
1 PREFIX up:<http://purl.uniprot.org/core/>
2 PREFIX taxon:<http://purl.uniprot.org/taxonomy/>
3 PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>
4 SELECT ?protein ?text
5 WHERE
6 {
7   ?protein a up:Protein .
8   ?protein up:organism taxon:9606 .
9   ?protein up:annotation ?annotation .
10  ?annotation a up:Natural_Variant_Annotation .
11  ?annotation rdfs:comment ?text .
12  FILTER (CONTAINS(?text, 'loss of function'))
13 }
```

Select all human UniProt entries with a sequence variant that leads to a 'loss of function'

Биологические базы данных

Biology is a data-intensive science!

- Нужно уметь хранить данные
- Нужно уметь обрабатывать данные
- Нужно уметь обмениваться данными
- Данные должны быть максимально открыты и доступны научному сообществу.
- Data provenance ("происхождение данных")

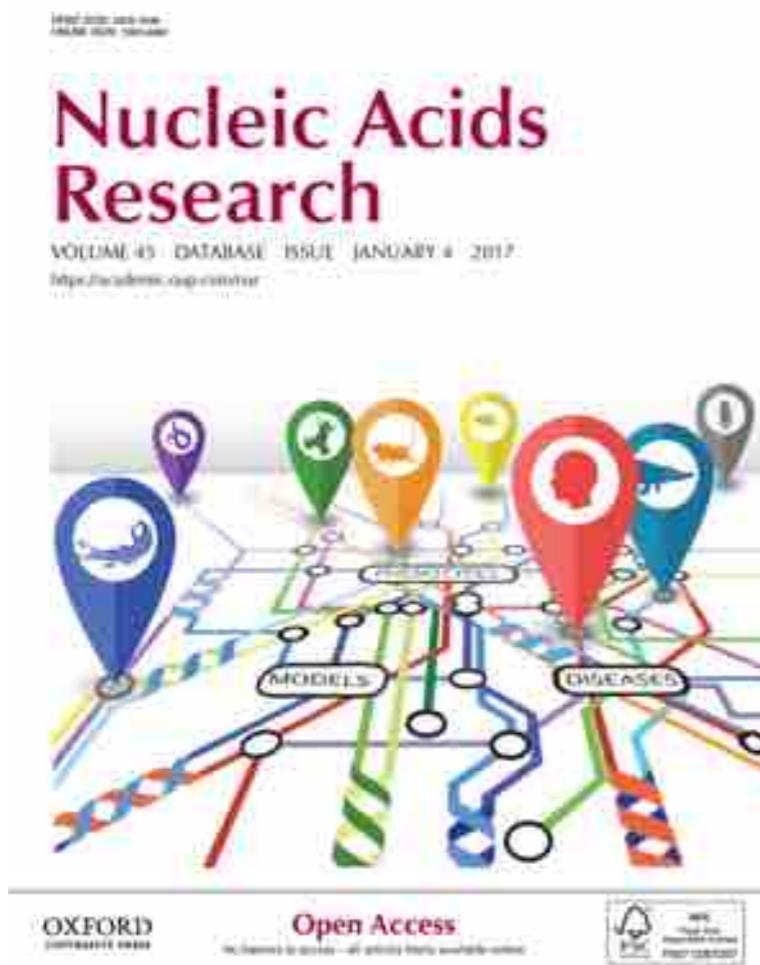
Data provenance [edit]

Scientific research is generally held to be of good provenance when it is documented in detail sufficient to allow reproducibility.^{[27][28]} Scientific workflow systems assist scientists and programmers with tracking their data through all transformations, analyses, and interpretations. Data sets are reliable when the process used to create them are

- Кризис воспроизводимости результатов в науке!?

Базы данных для биологии

- На данный момент количество не возможно сосчитать – очень много – важно не запутаться и не потеряться при их использовании
- Надежные источники информации о базах данных – научные журналы



Annual Database Issue – информация о ~200 БД каждый год.

Базы данных для биологии

- Бесплатные vs Платные (по подписке)
- Свободно доступные vs Ограниченно доступные
- Большие ресурсы (NCBI, EBI/EMBL, etc.)
интегрирующие многие базы данных -
поддерживаются государством
- Коллaborации между университетами (напр. PDB)
- Коммерческие компании
- Локальные базы данных, поддерживаемые силами
научных групп
- База данных vs Web Server – граница размыта.
- Хорошие БД - информационные ресурсы с
возможностями сложного поиска и моделирования.

Крупные центры биологических БД



National
Center for
Biotechnology
Information



- Bethesda, MD USA
- Более 60 БД включая PubMed, GenBank, DBGap, SRA



- European Bioinformatics Institute, Cambridge, UK + Switzerland

Что храниться?

- БД статей, абстрактов, патентов
- Последовательности ДНК
- Последовательности белков
- 3D структуры молекул
- Геномы
- Данные экспрессии
- Сырые данные с секвенаторов
- Информация о химических соединениях и их активности
- Информация о болезнях, информация о пациентах
- Информация о видах живых организмов
- Информация о метаболических и сигнальных путях
- Информация о взаимодействии молекул
- Много производной информации: базы гомологичных последовательностей, аннотация отдельных классов белков и т.д.

План

- **Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)**
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Реферативные базы данных

Clinical/Biomedical

PubMed – US National Library of Medicine database (Medline); refers to >25M articles from 5600 biomedical journals, 1940s to present, with some older items, in medicine, nursing, dentistry, veterinary medicine, allied health & pre-clinical sciences

- bibliographic database with author-provided abstracts, added indexing terms from **MeSH** (Medical Subject Headings) thesaurus, & links to other resources

www.pubmed.gov



FREE

Реферативные базы данных

Clinical/Biomedical

Embase – European based, includes all of Medline (database behind PubMed) and more; > 29M records, >8,500 journals, 1940s to present; includes coverage of more basic science journals & pre-clinical topics - especially useful for drug pipeline information, biotechnology, medical devices, conference coverage, toxicology, health policy/management, & alternative/complementary medicine
EMTREE thesaurus includes almost twice as many terms as PubMed

<https://www.elsevier.com/solutions/embase-biomedical-research>

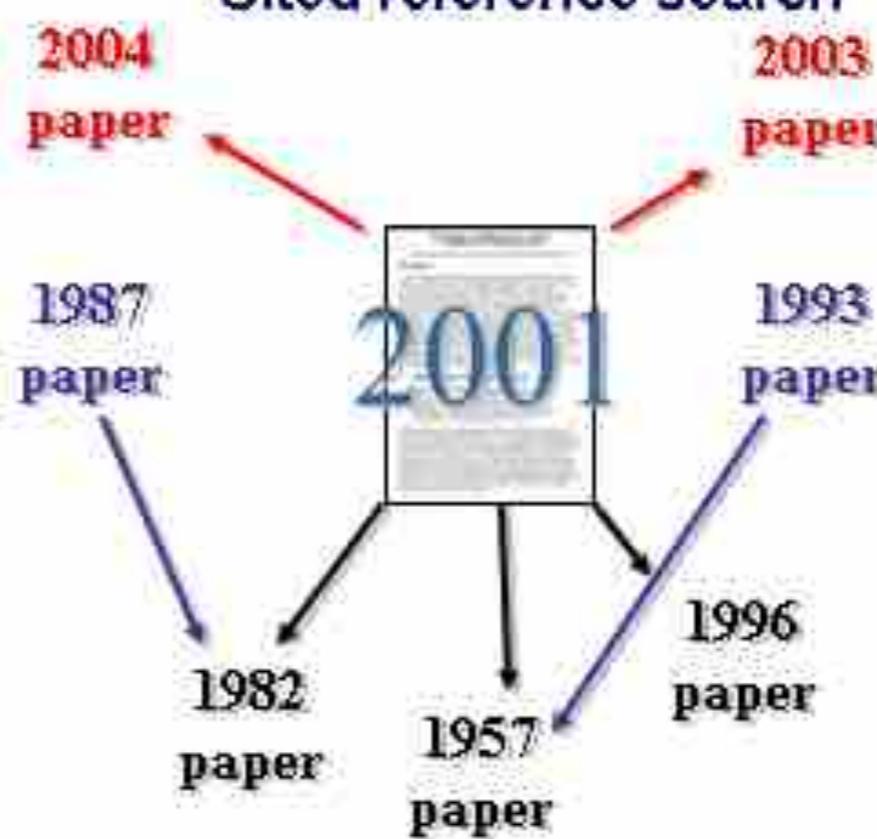
Реферативные базы данных

Cited Reference Searching

Traditional search



Cited reference search



Source: ISI

Реферативные базы данных

Общенаучные базы данных цитирований

Web of Science - covers >12,000 journals from 1900 to present;
useful for cited reference, **conference information & affiliations** (institutions)



PAID

<https://webofknowledge.com/>

Scopus – covers >18,500 journals from 1823 to present, complete citation counts for indexed articles 1996 to present; a general science database, not a specialized database – useful for cited reference, **conference information & affiliations** (institutions)



<https://www.scopus.com/>

ELSEVIER

PAID

Реферативные базы данных

Общенаучные базы данных цитирований



НАУЧНАЯ ЭЛЕКТРОННАЯ БИБЛИОТЕКА

eLIBRARY.RU

ЧИТАТЕЛЯМ | ОРГАНИЗАЦИЯМ | ИЗДАТЕЛЬСТВАМ | АВТОРАМ | БИБЛИОТЕКАМ



Российская Академия Наук

Открытый доступ к журналам
Российской академии наук

ОСНОВНЫЕ ПРОЕКТЫ НА ПЛАТФОРМЕ eLIBRARY.RU

Научная электронная библиотека eLIBRARY.RU - это крупнейший российский информационно-аналитический портал в области науки, технологии, медицины и образования, содержащий рефераты и полные тексты более 26 млн научных статей и публикаций, в том числе электронные версии более 5600 российских научно-технических журналов, из которых более 4800 журналов в открытом доступе. Подробнее...

● РОССИЙСКИЙ ИНДЕКС НАУЧНОГО ЦИТИРОВАНИЯ

Национальная библиографическая база данных научного цитирования, аккумулирующая более 12 миллионов публикаций российских ученых, а также информацию о цитировании этих публикаций из более 6000 российских журналов

● SCIENCE INDEX ДЛЯ ОРГАНИЗАЦИЙ

Информационно-аналитическая система Science Index для анализа публикационной активности и цитируемости научных организаций

● SCIENCE INDEX ДЛЯ АВТОРОВ

Инструменты и сервисы, предлагаемые зарегистрированным авторам научных публикаций

● RUSSIAN SCIENCE CITATION INDEX

Совместный проект компаний Osnovate Analytics и Научная электронная библиотека eLIBRARY.RU

НОВОСТИ И ОБЪЯВЛЕНИЯ

- 25.01 Опубликована программа конференции SCIENCE ONLINE 2018
- 19.01 Открыта регистрация на семинар: "Использование РИНЦ и Science Index для анализа и оценки научной деятельности" 27 февраля
- 12.01 Открыт свободный доступ к архивам журналов Российской академии наук
- 05.10 Продолжение конкурсной поддержки программ развития научных журналов

● Другие новости:

ТЕКУЩЕЕ СОСТОЯНИЕ

Число наименований журналов:	62015
- из них российских журналов:	15351
- из них выходящих в настоящее время:	13115
Число журналов, индексируемых в РИНЦ:	5200
Число журналов с полными текстами:	10721
- из них в открытом доступе:	5537



Подписка научных организаций на
информационно-аналитическую
систему
SCIENCE INDEX

Реферативные базы данных

Общенаучные базы данных цитирований



Recommended articles

[Exploring DNA dynamics within oligonucleosomes with coarse-grained simulations: SIRAH force field extension for protein-DNA complexes](#)

A Brandner, A Schüller, F Melo, S Pantano - Biochemical and biophysical research ..., 2017

Базы данных диссертаций

Open DOAR <http://www.opendoar.org/index.html> ;

OpenThesis <http://www.openthesis.org/> ;

BASE – Bielefeld Academic Search Engine -

<http://www.base-search.net/>

> refine search result > document type > theses

ProQuest Dissertations & Theses

Database <http://www.proquest.com/products-services/pqdt.html> - from 1743

to present; some fulltext since 1990; **fee** with some
free search capability

Базы данных патентов



http://www.lens.org/lens/biological_search – поиск
ДНК последовательностей



Search and read the full text of patents from around the world.

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- **Базы данных последовательностей ДНК**
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Базы данных нуклеотидных последовательностей

Нуклеотидные БД – это хранилища, принимающие данные от научного сообщества и представляющие их широкой общественности. Различные БД отличаются по источнику последовательностей, их надежности, широте аннотирования и т.д. В идеале БД должна содержать все известные последовательности.

The **International Nucleotide Sequence Database Collaboration** – совместный проект EMBL-Bank в Европейском Институте Биоинформатики (EBI), японского банка данных ДНК (DDBJ) в Центре Информационной Биологии (CIB) и GenBank в Национальном Центре Биотехнологической Информации (NCBI).



База данных GenBank

Открытая БД нуклеотидных последовательностей, учреждена в 1982 г.

2017: > 300 000 организмов, ~ 203 млн. последовательностей,

~ 240 млрд. пар оснований

The screenshot shows the NCBI GenBank homepage. At the top, there's a blue header bar with the NCBI logo, 'NCBI Resources' dropdown, and 'How To' dropdown. Below the header, the word 'GenBank' is displayed. A search bar is positioned above a horizontal menu bar containing links for 'GenBank', 'Submit', 'Genomes', 'WGS', 'Metagenomes', 'TPA', 'TSA', 'INSDC', and 'Other'. The main content area is titled 'Sample GenBank Record'.

Sample GenBank Record

This page presents an annotated sample GenBank record (accession number [U49845](#)) in its *GenBank Flat File* format. You can see the corresponding [live record for U49845](#), and see [examples of other records](#) that show a range of biological features.

<u>LOCUS</u>	SCU49845	5028 bp	DNA	PLN	21-JUN-1999
<u>DEFINITION</u>	Saccharomyces cerevisiae TCP1-beta gene, partial cds, and Axl2p (AXL2) and Rev7p (REV7) genes, complete cds.				
<u>ACCESSION</u>	U49845				
<u>VERSION</u>	U49845.1	GI:1293613			
<u>KEYWORDS</u>	.				
<u>SOURCE</u>	Saccharomyces cerevisiae (baker's yeast)				
<u>ORGANISM</u>	Saccharomyces cerevisiae Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
<u>REFERENCE</u>	1 (bases 1 to 5028)				
<u>AUTHORS</u>	Torpey,L.E., Gibbs,P.E., Nelson,J. and Lawrence,C.W.				
<u>TITLE</u>	Cloning and sequence of REV7, a gene whose function is required for DNA damage-induced mutagenesis in <i>Saccharomyces cerevisiae</i>				
<u>JOURNAL</u>	<i>Yeast</i> 10 (11), 1503-1509 (1994)				
<u>PUBMED</u>	7871890				
<u>REFERENCE</u>	2 (bases 1 to 5028)				
<u>AUTHORS</u>	Roemer,T., Madden,K., Chang,J. and Snyder,M.				
<u>TITLE</u>	Selection of axial growth sites in yeast requires Axl2p, a novel plasma membrane glycoprotein				
<u>JOURNAL</u>	<i>Genes Dev.</i> 10 (7), 777-793 (1996)				
<u>PUBMED</u>	8846915				

База данных GenBank. Структура файла

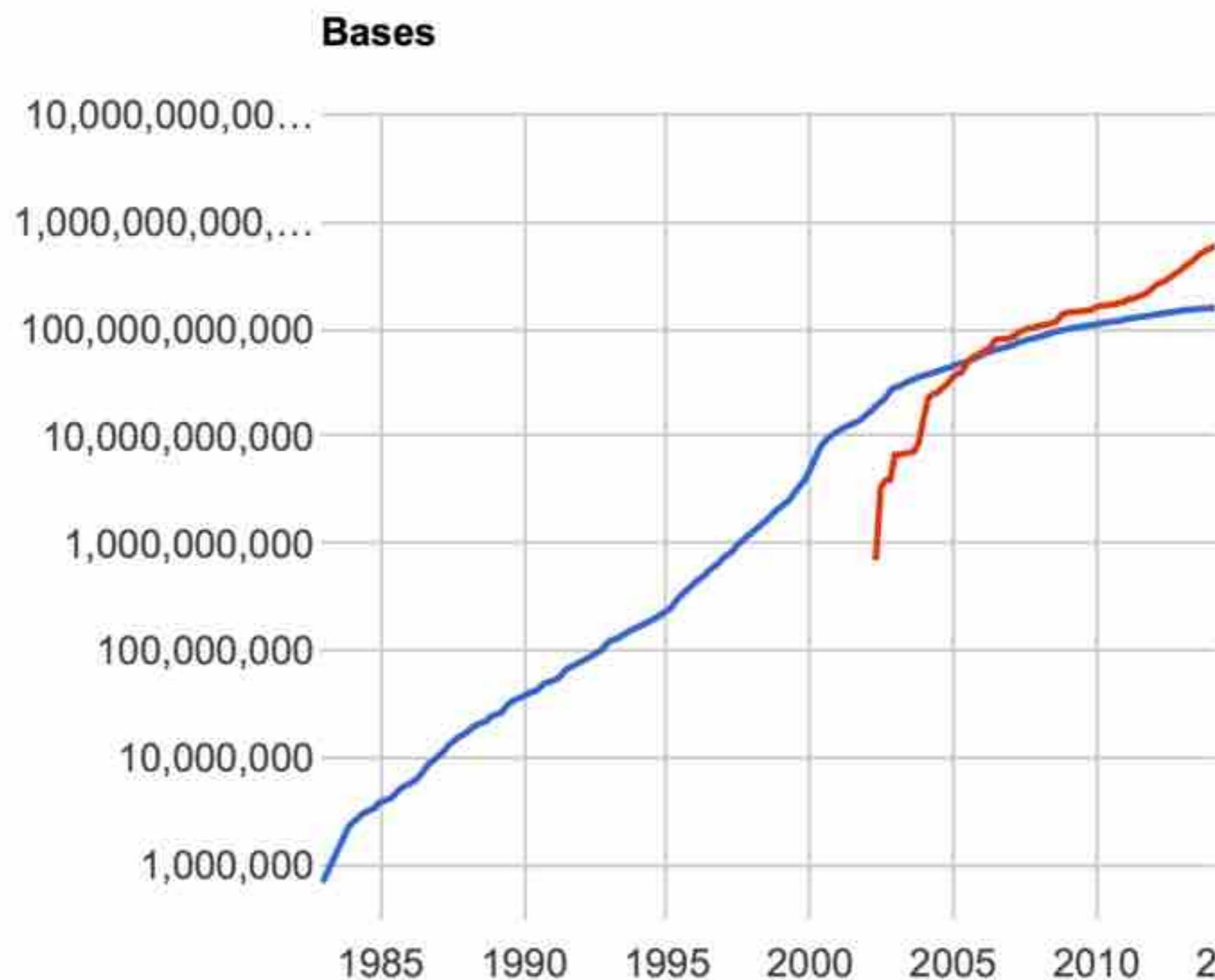
```
...  
FT      /translation="MGQPGNGSAFL LAPNGSHAPDHDTQQRDEVWVVG  
FT      AIVFGNVLVITAIAKFERLQTNTYFITSACADLVMGLAVVPFGAAHIL  
FT      MKMWTFGNF  
FT      WCEFWTSIDVLCVTASIETLCVIAVDRYFAITS  
FT      SPFKYQSLLTKNKARVII  
FT      ILMVWIVSGL  
FT      TSFLPIQMHWYRATHQEAINCYANETCCDFFT  
FT      NQAYAIASSIVSFYVPLVIMVFVYSRV  
FT      FQEAKRQLQKIDKSEG  
FT      RHFHVQNLSQVEQDGRTGHGLRRSSKFCLKEH  
FT      KALKTLGIIMGT  
FT      FTLCWLPFFIVNIVHVIQDNLIRKEVYILL  
FT      NWIGYVNNSGFNPLIYCRSPDFRIAFQELL  
FT      CLRRSSLKAYGNGYSSNGNTGEQSGYHVEQEKEN  
FT      KLLEDLPGTEDFVGHQGT  
FT      VPSDNI  
FT      DSQGRNCSTND  
FT      SLL<<  
FT      variation      46  
FT      /gene="ADRB2<<  
FT      /replace="a<<  
FT      /note="Arg16 to Gly polymorphism<<  
XX  
...  
...
```

База данных GenBank. Структура файла

...

SQ Sequence 1242 BP; 275 A; 331 C; 326 G; 310 T; 0 other;
atggggcaac ccgggaacgg cagcgccttc ttgctggcac ccaatggaag ccatgcgccg 60
gaccacgacg tcacgcagca aaggacgag gtgtgggtgg tggcatggg catcgatcg 120
tctctcatcg tcctggccat cgtgtttggc aatgtgctgg tcacacagc cattgccaag 180
ttcgagcgtc tgcaacggt caccaactac ttcatcactt cactggcctg tgctgatctg 240
gtcatgggcc tggcagtggc gcccattggg gccgcccata ttcttatgaa aatgtggact 300
tttggcaact tctggtgca gttttggact tccattgatg tgctgtgcgt cacggccagc 360
attgagaccc tgtgcgtgat cgcaatggat cgctactttg ccattacttc accttcaag 420
taccagagcc tgctgaccaa gaataaggcc cgggtgatca ttctgatggc gtggatttg 480
tcaggccta ccccttctt gcccattcag atgcactggc accgggcccc ccaccaggaa 540
gccatcaact gctatgcca tgagacactgc tgtgacttct tcacgaacca agcctatgcc 600
attgcctctt ccatcggtgc cttctacgtt cccctggta tcatggtctt cgtctactcc 660
agggtcttc aggaggccaa aaggcagctc cagaagattt acaaatctga gggccgcttc 720
catgtccaga accttagcca ggtggagcag gatggcgga cggggcatgg actccgcaga 780
tcttccaagt tctgcttgaa ggagcacaaa gccctcaaga cgttaggcat catcatggc 840
actttcaccc tctgctggct gcccttcttc atcgtaaca ttgtgcgt gatccaggat 900
aacctcatcc gtaaggaagt ttacatcctc ctaaatttggc taggctatgt caattctgg 960
ttcaatcccc ttatctactg ccggagccca gatttcagga ttgccttcca ggagcttctg 1020
tgccctgcgca ggtcttctt gaaggcctat gggaaatggct actccagcaa cggcaacaca 1080
ggggagcaga gtggatatca cgtggAACAG gagaaAGAAA ataaACTGCT gtgtGAAGAC 1140
ctcccaggca cggaaagactt tggggccat caaggactg tgcctagcga taacattgt 1200
tcacaaggga ggaattttag tacaaatgac tcactgctgt aa 1242

//



GenBank and WGS Statistics

<https://www.ncbi.nlm.nih.gov/genbank/statistics/>

Genbank – is an archive! Contains everything.

Nicotiana tabacum chloroplast JLA region, sequence 2

GenBank: Z71230.1

[FASTA](#) [Graphics](#)

FEATURES	Location/Qualifiers
source	1..124 /organism="Nicotiana tabacum" /organelle="plastid:chloroplast" /mol_type="genomic DNA" /isolate=" <u>Cuban cahibo cigar, gift from President Fidel Castro</u> " /db_xref="taxon:4097"

RefSeq – is a reference sequence database!

RefSeq – is a reference sequence database!

Если нужен список последовательностей всех генов человека – это вопрос к RefSeq, а не GenBank!

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- **Базы данных последовательностей белков**
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

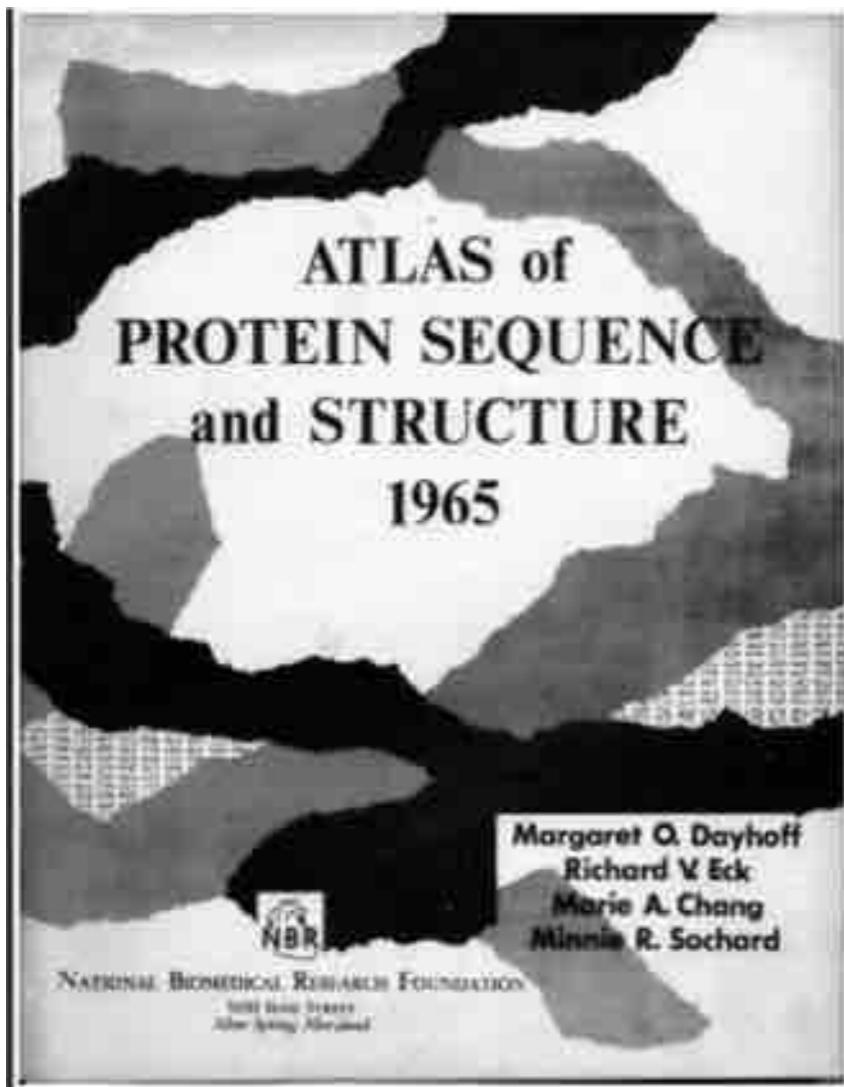
GenBank/RefSeq is nucleotide centric, but

...

```
FT      /translation="MGQPGNGSAFL LAPNGSHAPDHDTQQRDEVVVGMGIVMSLIVL
FT      AIVFGNVLVITAIKFERLQT VTVNYFITS LACADLVMGLAVV PFGAAH ILMKMWTFGNF
FT      WCEF WTSIDVLCVTASIETLCVIAVDRYFAITS PFKYQSLLTKNKARVIIILMVWIVSGL
FT      TSFLPIQMHWYRATHQEAINCYANETCCDFFTQNQAYAIASSIVSFYVPLVIMVFVYSRV
FT      FQEAKRQLQKIDKSEGRFH VQNLSQVEQDGRTGHGLRRSSKFCLKEHKALKTLGIIMGT
FT      FTLCWL PFFIVNIVHVIQDNLIRKEVYILLNWIGYVN SGNPLIYCRSPDFRIAFQELL
FT      CLRRSSLKAYGNGYSSNGNTGE QSGYHVEQEKENKL CEDLPGTEDFVGHQGTVP SDNI
FT      DSQGRNCSTND SLL<<
FT  variation 46
FT      /gene="ADRB2<<
FT      /replace="a<<
FT      /note="Arg16 to Gly polymorphism<<
XX
...
...
```

Protein sequences are annotate within GB records

Protein Centric Sequence Databases



Margaret Oakley Dayhoff
1925-1983

Margaret Dayhoff, a founder of the field of bioinformatics

Invented one-letter amino acid code, substitution matrices, etc.

https://en.wikipedia.org/wiki/Margaret_Oakley_Dayhoff

Protein Centric Sequence Databases



<http://pir.georgetown.edu>

In 2002, PIR along with its international partners, EBI (European Bioinformatics Institute) and SIB (Swiss Institute of Bioinformatics), were awarded a grant from NIH to create UniProt, a single worldwide database of protein sequence and function, by unifying the PIR-PSD, Swiss-Prot, and TrEMBL databases. As of 2010, PIR offers a wide variety of resources mainly oriented to assist the propagation and standardization of protein annotation: PIRSF,^[8] iProClass, and iProLINK.

The Protein Ontology (PRO) is another popular database released by the Protein Information Resource.^{[9][10]}

Белковые базы данных



Number of entries in UniProtKB/Swiss-Prot



UniProt – наиболее всеобъемлющий каталог информации о белках, объединяющий в себе данные из UniProtKB/Swiss-Prot, UniProtKB/TrEMBL и PIR.

Белковые базы данных

The screenshot shows the UniProt homepage with a navigation bar at the top featuring links for BLAST, Align, Retrieve/ID mapping, Peptide search, Advanced search, Help, and Contact. A yellow banner at the top states: "From June 20, 2018 all traffic will be automatically redirected to HTTPS. More information or view this page using https." Below the banner, a mission statement reads: "The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information." On the left, there's a sidebar for UniProtKB with statistics: Swiss-Prot (558,568) manually annotated and reviewed, and TrEMBL (107,627,435) automatically annotated and not reviewed. The main content area includes sections for UniRef (Sequence clusters), UniParc (Sequence archive), Proteomes (Proteome analysis), Supporting data (Literature citations, Taxonomy, Subcellular locations, Cross-ref. databases, Diseases, Keywords), and a News section with links to recent releases and an archive.

Качественно аннотированную информацию о белках
нужно искать в UniProtKB

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- **Базы данных 3D структур**
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Структурные базы данных

PDB
PROTEIN DATA BANK

A MEMBER OF THE  PDB
An Information Portal to Biological Macromolecular Structures

As of Tuesday Oct 12, 2010 at 5 PM PDT there are 68562 Structures | PDB Statistics

Contact Us | Print | PDB ID or Text | Search | ? | Advanced Search | Customize This Page

MyPDB Hide
Login to your Account
Register a New Account

Home Hide
News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide
All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamline
Related Tools

Search Hide

A Resource for Studying Biological Macromolecules

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the [wwPDB](#), the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

[Hide Welcome Message](#)

Featured Molecules (MotM Category View / Previous) Hide
Features: MotM | PSI

Structural View of Biology  **Protein Synthesis**    

Molecule of the Month:
Riboswitches
Why use two or more molecules when one?

New Features Hide
Improved Domain Display on Sequence Tab
Latest features released:
Website Release Archive

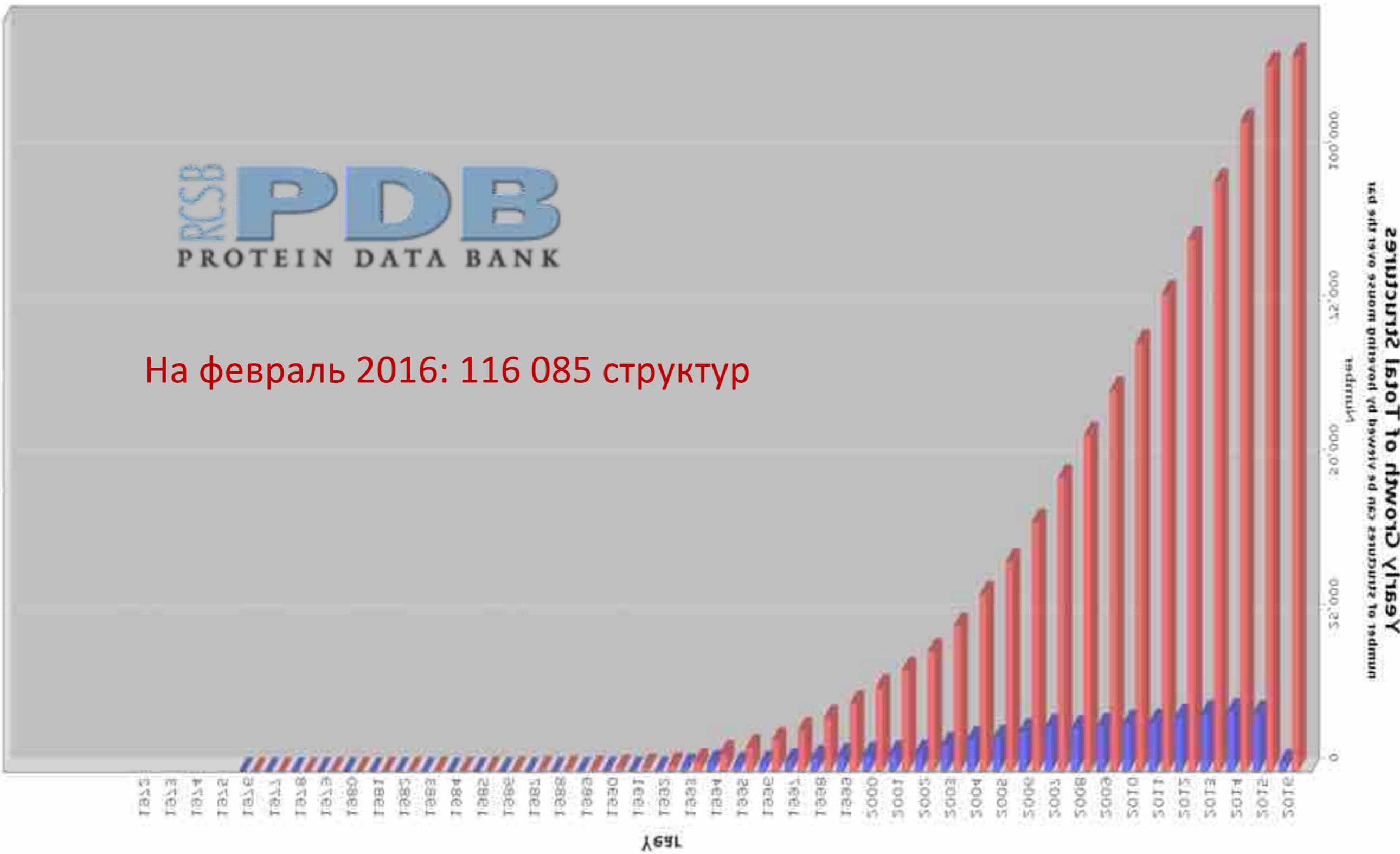
RCSB PDB News Hide
Weekly | Quarterly | Yearly
2010-10-12
Categorized View of Molecule of the Month


Explore new Molecule of the Month feature categories.

- Latest Website Release
- Poster Prize Awarded at ECM
- Analyze small molecule interactions in the PDR with

PDB – содержит информацию об экспериментально определенных структурах белков, нуклеиновых кислот и различных комплексов.

Структурные базы данных



На февраль 2016: 116 085 структур

Структурные базы данных

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLMAMMONIUM IN HIGH K

MyPDB Hide
Login to your Account
Register a New Account

Home Hide
News & Publications
Usage/Reference Policies
Deposition Policies
Website FAQ
Deposition FAQ
Contact Us
About Us
Careers
External Links
Sitemap
New Website Features

Deposition Hide
All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamline
Related Tools

Search Hide
Advanced Search
Latest Release
New Structure Papers
Sequence Search
Chemical Components
Unreleased Entries
Browse Database

Summary Sequence Derived Data Seq. Similarity 3D Similarity Literature Biol. & Chem. Methods Chemistry Links

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLMAMMONIUM IN HIGH K 2JK5
DOI:10.2210/pdb2jk5/pdb

Primary Citation
Structure of a Potassium Channel in Complex with Symmetrical Quaternary Ammonium Compounds Reveals a Novel Hydrophobic Binding Site
Lenaeus, M.J., Focia, P.J., Wagner, T.P., Gross, A.P.
Journal: To be Published
Not in PubMed

Molecular Description
Classification: Immune System/metal Transport
Structure Weight: 61280.07
Molecule: ANTIBODY FAB FRAGMENT LIGHT CHAIN
Polymer: 1 **Type:** polypeptide(L) **Length:** 219
Chains: A

Molecule: ANTIBODY FAB FRAGMENT HEAVY CHAIN
Polymer: 2 **Type:** polypeptide(L) **Length:** 212
Chains: B

Molecule: VOLTAGE-GATED POTASSIUM CHANNEL
Polymer: 3 **Type:** polypeptide(L) **Length:** 124
Chains: C

Biological Assembly 2
Biological Assembly Image for 2JK5
More Images...
View in Jmol SimpleViewer
Other Viewers Protein Workshop
Biological assembly generated by PISA (Software)

Структурные базы данных

POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM IN HIGH K

Deposition Hide
All Deposit Services
Electron Microscopy
X-ray | NMR
Validation Server
BioSync Beamline
Related Tools

Search Hide
Advanced Search
Latest Release
New Structure Papers
Sequence Search
Chemical Components
Unreleased Entries
Browse Database
Histograms

Explorer:
Last Structure: 2K5

Tools Hide
File Downloads
Compare Structures
FTP Services
File Formats
Services: RESTful | SOAP
Widgets

Education Hide
Understanding PDB Data
Molecule of the Month
Educational Resources

Chain Display

Chain A (polymer 1) [help] [fasta] [text/markup]

Description ANTIBODY FAB FRAGMENT LIGHT CHAIN
Chain Type polypeptide(L)
Length 219 residues
dssp secondary structure 5% helical (4 helices; 12 residues)
51% beta sheet (22 strands; 112 residues)
[hide] [reference]
More annotations Select

Sequence Details:

dssp PDB PDB
QVQLQQPGAEVKPGASVKLSCKASGYTFTSDWIHWVKQRPGHGLEWIGEIIIPSYGRANY
PDB 1 10 20 30 40 50 60

dssp PDB PDB
NEKIQKKATLTADKSSSTAFMQLSSLTSEDSAVYYCARERGDGYFAVWGAGTTVTVSSAK
PDB 61 70 80 90 100 110 120

dssp PDB PDB
TTPPSVYPLAPGSAAQTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSDLY
PDB 121 130 140 150 160 170 180

dssp PDB PDB
TLSSSVTVPSSSWPSETVTCNVAHPASSTKVVDKKIVPRD
PDB 181 190 200 210 219

База данных PDB. Структура файла

```
HEADER      IMMUNE SYSTEM/METAL TRANSPORT          15-AUG-08  2JK5
TITLE       POTASSIUM CHANNEL KCSA IN COMPLEX WITH TETRABUTYLAMMONIUM
TITLE       2 IN HIGH K
COMPND      MOL_ID: 1;
COMPND      2 MOLECULE: ANTIBODY FAB FRAGMENT LIGHT CHAIN;
COMPND      3 CHAIN: A;
COMPND      4 ENGINEERED: YES;
COMPND      5 MOL_ID: 2;

...
KEYWDS     IMMUNE SYSTEM METAL TRANSPORT COMPLEX, QUATERNARY AMMONIUM,
...
EXPDTA     X-RAY DIFFRACTION
AUTHOR     M.J.LENAEUS, P.J.FOCIA, T.WAGNER, A.GROSS
REVDAT     1 17-NOV-09 2JK5    0
JRNL        AUTH     M.J.LENAEUS, P.J.FOCIA, T.WAGNER, A.GROSS
JRNL        TITL    STRUCTURE OF A POTASSIUM CHANNEL IN COMPLEX WITH
JRNL        TITL    2 SYMMETRICAL QUATERNARY AMMONIUM COMPOUNDS REVEALS
JRNL        TITL    3 A NOVEL HYDROPHOBIC BINDING SITE
JRNL        REF     TO BE PUBLISHED
JRNL        REFN
REMARK     2
REMARK     2 RESOLUTION.    2.4 ANGSTROMS.
REMARK     3
REMARK     3 REFINEMENT.
REMARK     3 PROGRAM       : REFMAC 5.5.0051
...
```

Структурные базы данных

NDB – основана в 1992 г. для сбора и распространения информации о структуре нуклеиновых кислот. Формат хранения данных идентичен PDB.

ndb
**WELCOME TO THE
NUCLEIC ACID DATABASE**
a repository of three-dimensional structural information about nucleic acids

• Atlas
• Deposit Data
• Download Data
• Search
• Reports
• Education
• Standards
• Tools
• Links

Number of Released Structures:
4923 Structures
Last Update: 13-Oct-2010

Search the NDB by ID
Enter an NDB ID or PDB ID

Search for Released Structures

Nucleic Acids Highlight:


Site Index

About NDB
News

The NDB follows the dictionaries and formats used by the Worldwide Protein Data Bank. Please see www.wwpdb.org for format announcements and documentation.

Archive of NDB newsletters.

The NDB is supported by funds from the National Science Foundation and the Department of Energy.

In citing the NDB please refer to: H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.-H. Hsieh, A. R. Srinivasan, and B. Schneider. (1992) The Nucleic Acid Database: A Comprehensive Relational Database of Three-Dimensional Structures of Nucleic Acids. *Biophys. J.*, 63, 751-759.

ndbadmin@ndbserver.rutgers.edu
©1995-2010 The Nucleic Acid Database Project Rutgers, The State University of New Jersey

Структурные базы данных

EMBL-EBI 

Editor Test Here Terms of Use | Privacy | Cookies

Databases Tools Research Training Industry About Us Help Site Index  

Home EMD-1367 **Mirrors** **Contact us**



Title: Three-dimensional structure of a voltage-gated potassium channel at 2.5 nm resolution.

Authors: Olga Sokolova, Ludmila Kolmakova-Partensky and Nikolaus Grigorieff

Sample: Shaker B channel

Aggregation state: Single particle (25 angstroms resolution)

 **Latest update:** 2011-05-26

• Summary

• Experimental details

• Visualization

• Map information

• Downloads

Summary

Status: Released

Deposition date: 2007-05-24

Header release date: 2007-05-30

Map release date: 2007-05-30

Primary citation: Sokolova O, Kolmakova-Partensky L, Grigorieff N. Three-dimensional structure of a voltage-gated potassium channel at 2.5 nm resolution. *STRUCTURE* (2001) 9, pp. 215-220 [PubMed 11286888]

Sample: Shaker B channel

Resolution: 25 Å (determined by FSC at 0.5 cut-off)

Fitted PDB:	PDB Authors	PubMed Status
1b18	Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., Mackinnon, R.	9525859 Released
1a68	Kreusch, A., Pfaffinger, P.J., Stevens, C.F., Choe, S.	9582078 Released

Структурные базы данных

 Cambridge Crystallographic Data Centre

About CCDC Products Case Studies Free Services News Events Contact Support Search

Location: Home

Products

Free Services

Cambridge Structural Database (CSD)

Deposit a Structure

Request a Structure

New consultancy service

News

Last Updated:
10 September 2010

Positions Available

 日本語

500,000th Crystal Structure Milestone Reached

The Cambridge Crystallographic Data Centre (CCDC) is proud to announce the archiving of the 500,000th small molecule crystal structure to the Cambridge Structural Database (CSD)



"The determination of 500,000 crystal structures is a remarkable achievement. However, the scientific community is hungry for the next 500,000 and the knowledge these will undoubtedly bring. As the CSD grows both in size and in the complexity of structures it contains, the database not only helps us to answer our questions about molecular structure and interactions, but tells us what those questions should be."

-Dr Colin Groom, Executive Director of the CCDC

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- **Базы данных хим. соединений**
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Базы данных химических соединений

Databases | Deposition | Services | Help | more

PubChem

BioAssay Compound Substance Advanced search

Chemical structure search | BioActivity analysis

New Structures from **Zancheng Functional Chemicals** are now available in PubChem.

more ... 

Write to Helpdesk | Disclaimer | Privacy Statement | Accessibility
National Center for Biotechnology Information
NLM | NIH | HHS

Bioactivity summary 
Bioactivity datatable 
Bioactivity structure-activity 
Chemical structure search 
3D conformer viewer 
Chemical structure clustering 
Deposition gateway 
Structure download 
PubChem FTP 

Базы данных химических соединений

NCBI  PubChem Compound

Search PubChem Compound for epinephrine Go Clear Save Search

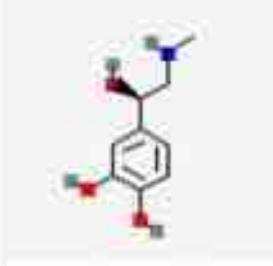
Advanced Search Preview/Index History Clipboard Details

Display Summary Show 20 Sort By Send to

Tools:    Links: Related Structures, BioAssays, BioSystems, Literature, Other Links

All: 120 Rule of 5: 62 

Items 1 - 20 of 120 Page 1 of 6 Next

1: CID: 5816 Related Structures, BioAssays, BioSystems, Literature, Other Links

epinephrine; Adrenalin; adrenaline ...
IUPAC: 4-[(1R)-1-hydroxy-2-(methylamino)ethyl]benzene-1,2-diol
MW: 183.204420 g/mol | MF: C₉H₁₃NO₃
Tested in BioAssays: All: 274, Active: 22; BioActivity Analysis 
Vasoconstrictor Agents... more

2: CID: 838 Related Structures, BioAssays, Literature, Other Links

epinephrine; DL-Adrenaline; Racepinephrine ...
IUPAC: 4-[(1R)-1-hydroxy-2-(methylamino)ethyl]benzene-1,2-diol
MW: 183.204420 g/mol | MF: C₉H₁₃NO₃
Tested in BioAssays: All: 17, Active: 1; BioActivity Analysis 
Vasoconstrictor Agents... more

Selected Compounds  Compound Count

BioActivity Experiments	
BioAssays, Active 	39
BioAssays, Tested 	60
Protein 3D Structures	6
Crystal Structure Of Dipeptide...	1
BioMedical Annotation	
Pharmacological Actions	69
Sympathomimetics	37
BioSystems	16
Depositor Category	
Biological Properties	120
Chemical Vendors	55
Journal Publishers	47
NIH Molecular Libraries	52

Recent activity 

Go to CID: 247704 Related Structures BioAssays Literature Other Links Turn off Close

Базы данных химических соединений

Chemical Abstract Service – в регистре содержится 130 млн соединений (2018)

The screenshot shows the homepage of the Chemical Abstract Service (CAS) website. At the top, there is a navigation bar with links for ACS JOURNALS, C&EN, and CAS. Below the navigation bar, the CAS logo is displayed, followed by the text "A division of the American Chemical Society". A search bar labeled "Site Search" is also present.

On the left side of the page, there are three main promotional boxes:

- SciFinder®**: Increase your productivity and achieve faster breakthroughs.
- STN®**: Access the most comprehensive scientific, technical, & patent information.
- 55 519 680** DEOGENIC AND INORGANIC SUBSTANCES TO DATE

The central feature of the page is a large banner with the headline: "Carbon Bond Formations Win 2010 Nobel Prize with Award-Winning Research in CAS Databases". Below the headline, it says: "Check out the early inspiring efforts of these three great chemists." There is a "LEARN MORE" button and a small navigation bar at the bottom right of the banner.

At the bottom of the page, there are three columns of links:

- QUICK LINKS**
Find all your product resources in SciFinder and STN Support & Training.
SciFinder®
- THE RESEARCH EDGE**
In the race to discovery, access to information is the critical differentiator. See how to ensure your research is on the cutting edge and stay ahead of the competition.
- CAS UPDATES**
 - CAS adds Costa Rica Registry of Industrial Property as 61st patent authority in its patent coverage
 - Access an additional 1.8 million records exclusively

Базы данных химических соединений



SEARCH CART MENU

Reaxys[®]

Reaxys retrieves literature, compound properties and chemical reaction data faster than any other solution.

[Get started](#)

Already a Reaxys customer?

If you already have a Reaxys account, please click the link below to sign in.

[Reaxys sign in](#)



Life is chemistry: Meet the Reaxys team

Our technologists and chemists combine their expertise to create a high-quality information database.

[Watch the video](#)

Teaching chemistry info literacy

Are you doing enough to prepare future chemists for their careers in research?

[Take the assessment](#)

[Why choose Reaxys](#)

[How Reaxys works](#)

[Who we serve](#)

[Customer stories](#)

[Learn & support](#)

Базы данных углеводов



For 2017:

7005 publications for 18924 compounds from 8859 organisms

Database search

- Structures
- Composition
- Organisms
- Publications
- NMR signals

Additional operations are available from the [left menu](#). If you don't see it [click here](#)

Useful tools

- Predict NMR
- Elucidate
- Fragments
- Cluster taxa
- GT activities
- Examples

Структурные базы данных

7009 структур липидов и сходных соединений – не поддерживается в настоящее время ☹

LipidBank

Vitamin A

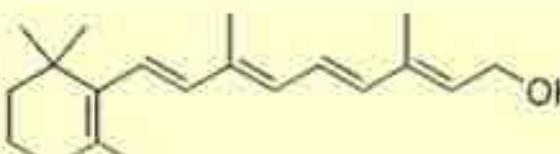
DATA No : VVA0001 INFORMANT : Hiroyuki Kagechika

NAME : ((2E,4E,6E,8E)-3,7-Dimethyl-9-(2,6,6-trimethyl-cyclohex-1-enyl)-nona-2,4,6,8-tetraen-1-ol

COMMON NAME : all-trans-retinol

SYMBOL : ROH

FORMULA : C₂₀H₃₀O MOL.WT (average) : 286.452



Download ChemDraw structure data

BIOLOGICAL ACTIVITY

Relative physiological activity for All trans-retinyl acetate is 100%. ([Ref. 0046](#)) New physiological functions of vitamin A have been identified in addition to its role in vision ([Ref. 0005](#)), namely its role in immune defence reducing morbidity of measles, ([Ref. 0006](#)) of respiratory and possibly HIV infections ([Ref. 0007](#)), in gene regulation ([Ref. 0008](#)), in cell differentiation ([Ref. 0009](#)) and morphogenesis. ([Ref. 0010/0011](#))

PHYSICAL AND CHEMICAL PROPERTIES

MELTING POINT : 62-64°C ([Ref. 0013](#))

LipidBank

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- **Базы данных геномов и аннотаций**
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Genome

Genome

Limits Advanced

Search

Help



Genome

This resource organizes information on genomes including sequences, maps, chromosomes, assemblies, and annotations.

Using Genome

[Help](#)[Browse by Organism](#) UPDATED[Download / FTP](#)[Download FAQ](#)[Submit a genome](#)

Custom resources

[Human Genome](#)[Microbes](#)[Organelles](#)[Viruses](#)

Other Resources

[Assembly](#)[BioProject](#)[BioSample](#)[Map Viewer](#)[Genome Data Viewer](#) NEW

Genome Tools

[BLAST the Human Genome](#)[Microbial Nucleotide BLAST](#)

Genome Annotation and Analysis

[Eukaryotic Genome Annotation](#)[Prokaryotic Genome Annotation](#)[PASC \(Pairwise Sequence Comparison\)](#)

External Resources

[GOLD - Genomes Online Database](#)[Bacteria Genomes at Sanger](#)[Ensembl](#)

Search: for

e.g. [BRCA2](#) or [rat 5:62797383-63627669](#) or [rs699](#) or
[coronary heart disease](#)

Browse a Genome

Ensembl is a genome browser for vertebrate genomes that supports research in comparative genomics, evolution, sequence variation and transcriptional regulation. Ensembl annotates genes from multiple

[Find a Data Display](#)

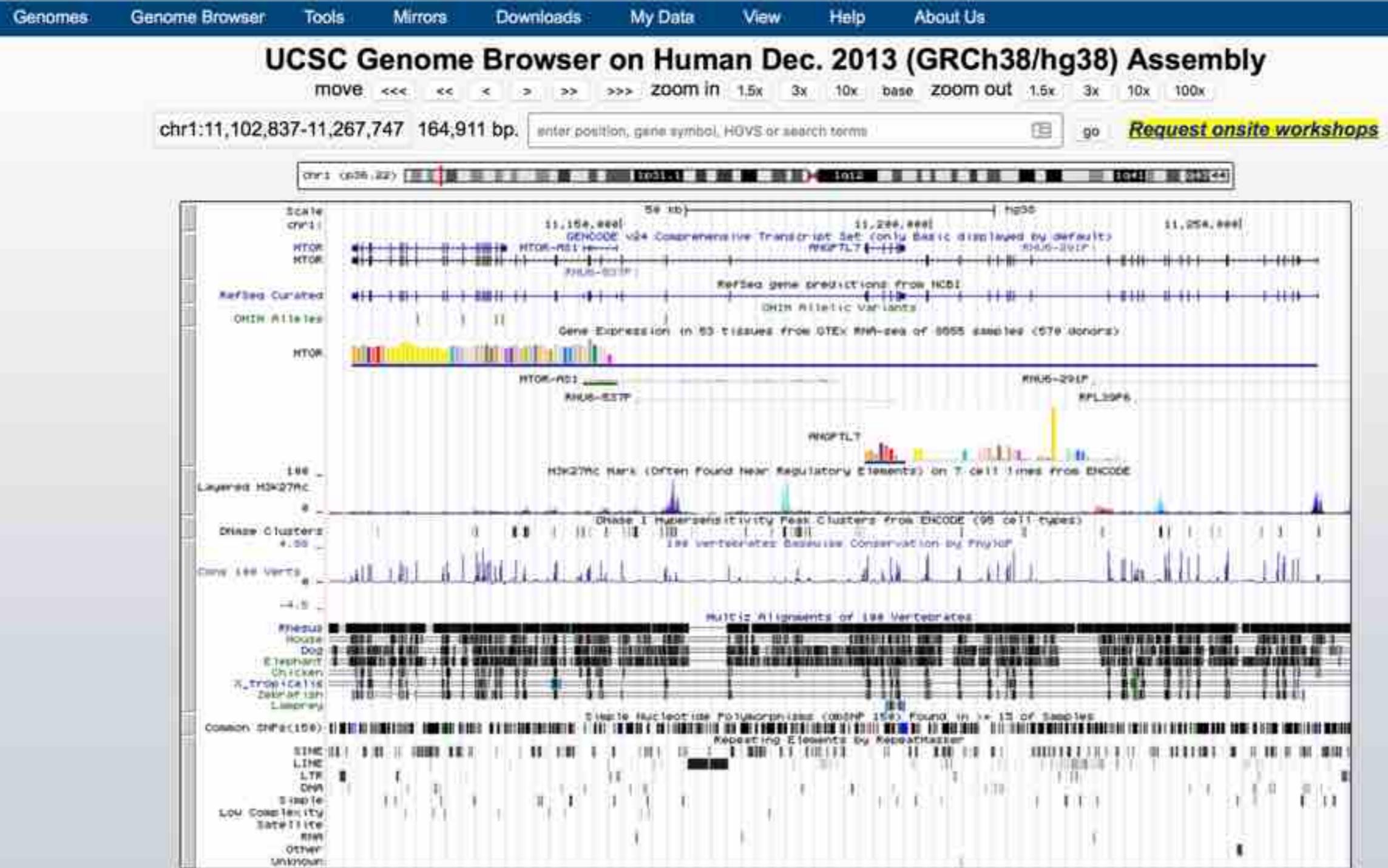


Not sure how to find the data visualisation you need?
With our

[Full details](#) | [All web updates](#), by

What's New in Ensembl Release 91 (December 2017)

- [New Primate Species](#)
- [Update of cat assembly and genebuild to *Felis_catus_8.0*](#)
- [Mouse: update to Ensembl-Havana GENCODE gene set](#)
- [New and updated probe mapping data for primates](#)
- [Microarray Probe Mapping Update](#)



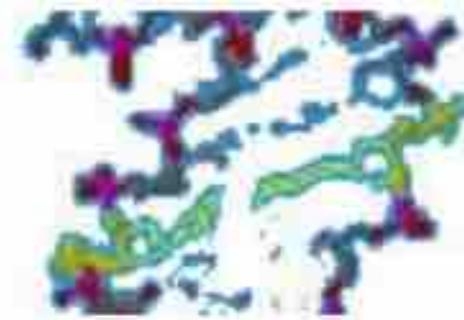
План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- **Базы данных вариаций генома**
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии



dbSNP

Short Genetic Variations

[dbVar](#)[ClinVar](#)[GaP](#)[PubMed](#)[Nucleotide](#)[Protein](#)

Search small variations in dbSNP or large structural variations in dbVar

Search Entrez

dbSNP



for



Go

Have a question
about dbSNP? Try
searching the SNP
FAQ Archive!



Go

ANNOUNCEMENT



dbSNP and dbVar no longer accept submissions
for non-human organism data. Please read more
[here](#).

GENERAL

RSS Feed

Contact Us

Organism Data

dbSNP Homepage

Search by IDs on All Assemblies

Note: `rs#` and `omim#` must be prefixed with "rs" or "omim" respectively. e.g.

dbVar

dbVar



Search

Help

Advanced



dbVar

dbVar is NCBI's database of human genomic structural variation — insertions, deletions, duplications, inversions, mobile elements, and translocations

Getting Started

[Overview of Structural Variation](#)

Accessing Data

[Structural Variation Data Hub](#)

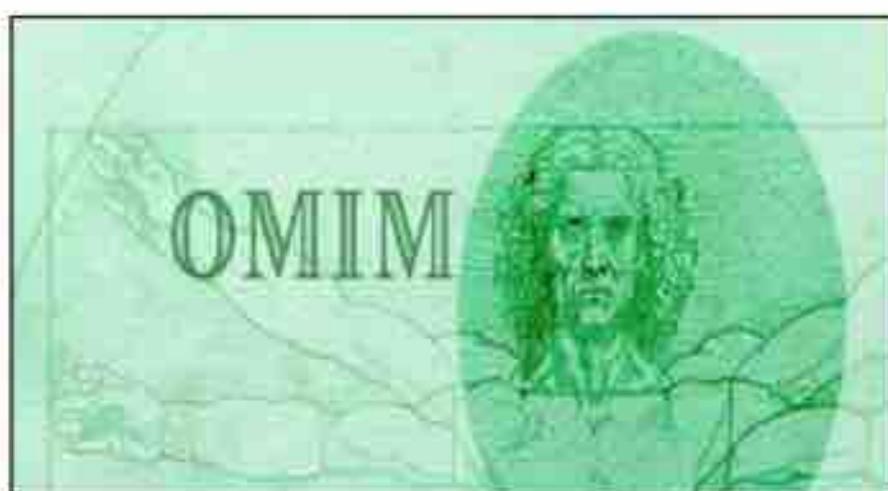
Other NCBI Resources

[dbSNP](#)



OMIM

OMIM

[Search](#)[Limits](#) [Advanced](#)[Help](#)

OMIM

OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, under the direction of Dr. Ada

Using OMIM

[Getting Started](#)[FAQ](#)

OMIM tools

[OMIM API](#)

Related Resources

[ClinVar](#)[Gene](#)[GTR](#)[MedGen](#)

ClinVar

ClinVar



Search ClinVar for gene symbols, HGVS expressions, conditions, and more

Search

Advanced

Help

Home About Access Help Submit Statistics FTP

```
ACTGATGGTATGGGCCAAGAGATATCT  
CAGGTACGGCTGTCACTAGTTAGACCTCAC  
CAGGGCTGGGCATAAAAGTCAGGGCAGAGC  
CCATGGTGCATCTGACTCCTCAGGAGAACT  
GCAGGTTGGTATCAAGGTTACAAGACAGGT  
GGCACTGACTCTCTGCCTATTGGTCTAT
```

ClinVar

ClinVar aggregates information about genomic variation and its relationship to human health.

Using ClinVar

[About ClinVar](#)[Data Dictionary](#)[Downloads/FTP site](#)[FAQ](#)[Contact Us](#)[RSS feed/What's new?](#)[Factsheet](#)

Tools

[ACMG Recommendations for Reporting of Incidental Findings](#)[ClinVar Submission Portal](#)[Submissions](#)[Variation Viewer](#)[Clinical Remapping - Between assemblies and RefSeqGenes](#)[RefSeqGene/LRG](#)

Related Sites

[ClinGen](#)[GeneReviews®](#)[GTR®](#)[MedGen](#)[OMIM®](#)[Variation](#)

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- **Базы данных геном-фенотип**
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии



Authorized Access Portal

[Log In to dbGaP](#)

dbGaP Data Download

The management portal to request and download individual level data

Click [here](#) to login to the dbGaP controlled-access portal and to begin a project request. For guidance on the development of a data access request to complete project requests, please see

 [Tips for preparing a successful Data Access Request.](#)

[Who can apply for access?](#)

[How does one apply?](#)

[Why is Access Controlled?](#)

[Additional help.](#)

dbGaP Data Browser – View Only

With dbGaP Data Browser approval through the simplified controlled-access application, users may view the collection “Compilation of individual-level data from general research use (GRU).”

[What is the purpose of the dbGaP Data Browser; why is it useful?](#)

[How does one apply?](#)



Data Portal
Get Cancer Data

Data Access Compliance Office
Apply for Access to Controlled Data

Contact Us

[Log In](#) | [Create an Account](#)



International
Cancer Genome
Consortium

Enter keywords

Search

[Home](#)

[Cancer Genome Projects](#)

[Committees and Working Groups](#)

[Policies and Guidelines](#)

[Media](#)

[Publications](#)

ICGC Cancer Genome Projects

Committed projects to date: **90**

Sort by: Project

Biliary Tract Cancer
Japan

Biliary Tract Cancer
Singapore

Bladder Cancer
China

Bladder Cancer
United States

Blood Cancer
China

Blood Cancer
Singapore

ICGC Goal: To obtain a **comprehensive** description of **genomic, transcriptomic and epigenomic changes** in **50 different tumor types and/or subtypes** which are of clinical and societal importance across the globe.

[Read more »](#)

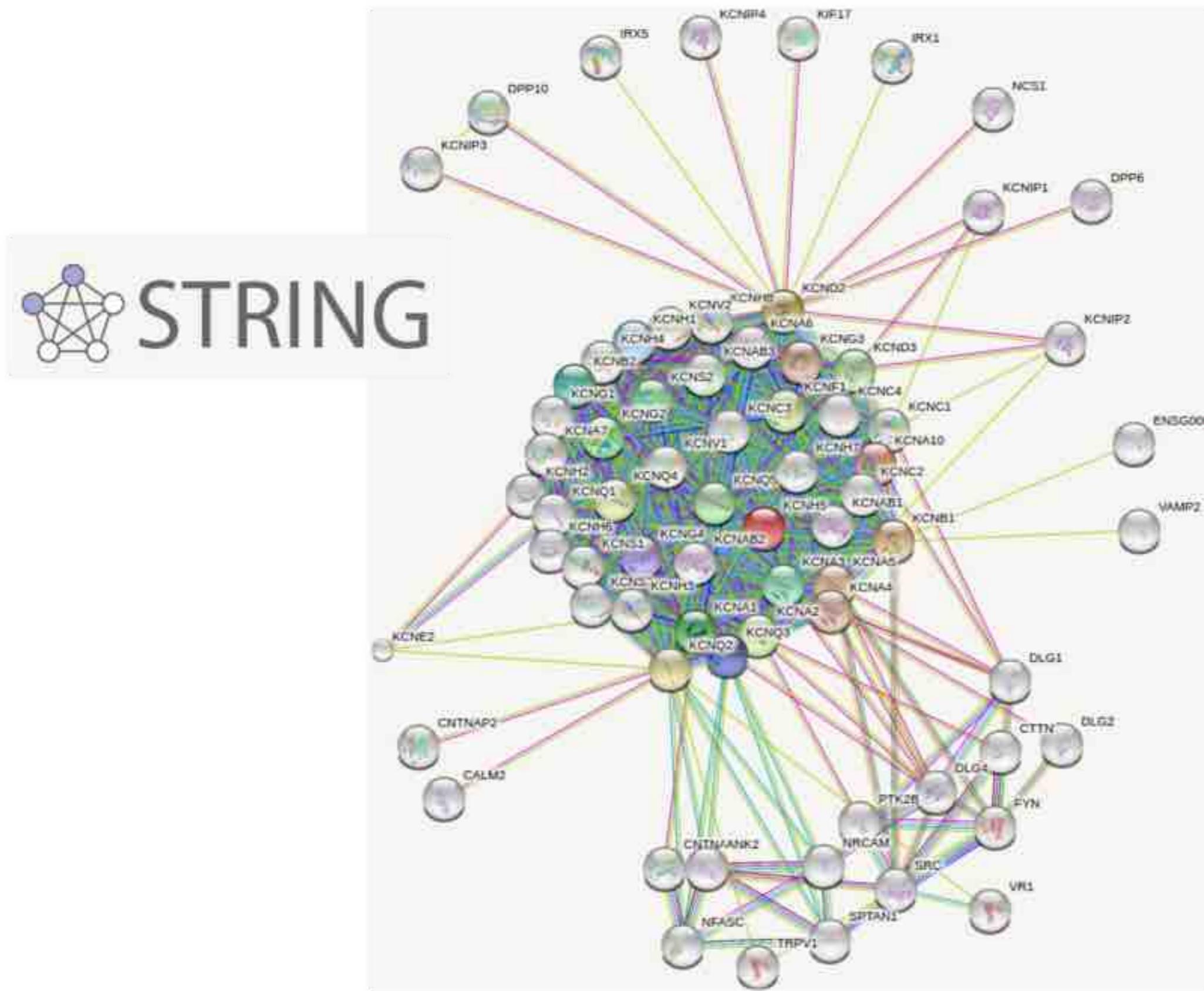
[Launch Data Portal »](#)

[Apply for Access to Controlled Data »](#)

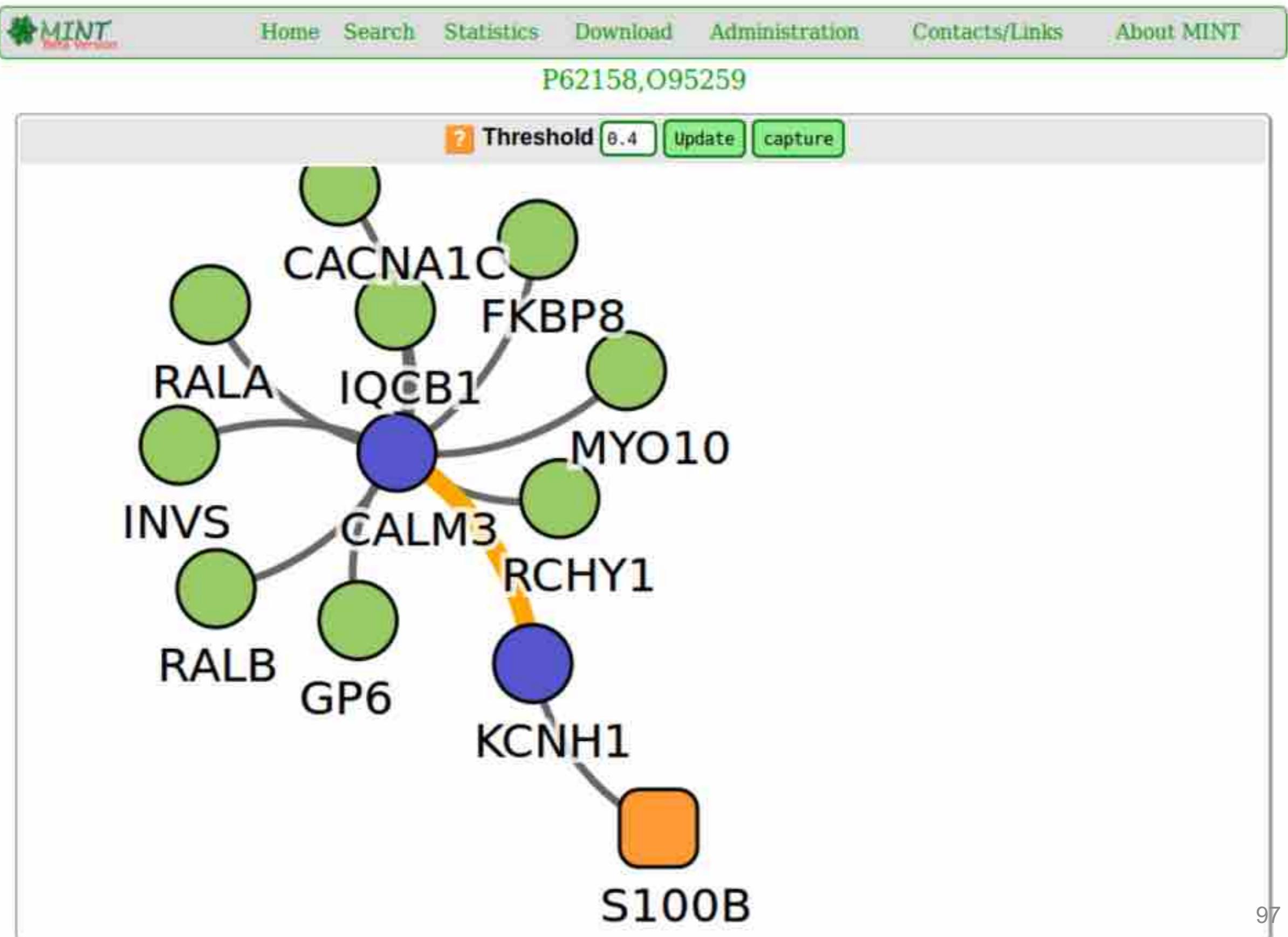
План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- **Базы данных взаимодействий**
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии

Базы данных взаимодействий



Базы данных взаимодействий



Базы данных взаимодействий

Kalium
Database of potassium channel
toxins from scorpion venom

Help & FAQs About & Contact

ИБХ

174 toxins shown										
	Organism		Name	Uniprot ID	Sequence	PDB	Mass	Pub. Date	Activity	
	<i>Leiurus quinquecinctus</i>		<i>α-KTx 1.1</i>	CNTX; CNTX-Lq...	P13487	ZFTNVS C TTSKE C WSV C QRLHNTSRGKCMNKKC R CYS	1BAH,1C1	4295.05	1988	Shaker,Kv1.1,Kv1.2,
	<i>Leiurus quinquecinctus</i>		<i>α-KTx 1.2</i>	ChtX-Lq2, ChTx...	P45628	ZFTQES C TASNQC C WSICKRLHNTNRGKCMNKKC R CYS	1LR	4335.08	1989	Kv1,KCa1.1
	<i>Mesobuthus tamulus</i>		<i>α-KTx 1.3</i>	IdTx; Iberiotoxin	P24663	ZFTD V DCSVSK C EWSV C KDLFGVDRGKCMGKKC R CYS		4230.02	1990	KCa1.1
	<i>Mesobuthus tamulus</i>		<i>α-KTx 1.5</i>	BmTX1, Neurotoxin	Q9NI06	ZFTDVK C TGSK C WPV C KQMF C GKPN C MNGK C R C YS	1B1G	4169.04	1997	Kv1,Kv1.3,KCa1.1
	<i>Mesobuthus tamulus</i>		<i>α-KTx 1.6</i>	BmTX2, Neurotoxin	Q9NI05	ZFTNV C SASSQ C WPV C KKLFGTYRGKCMNSK C R C YS	2BMT	4178.97	1997	Kv1,Kv1.3,Kv11.1,Kv1.2
	<i>Parabuthus transvaalicus</i>		<i>α-KTx 1.10</i>	PBTx3; Parabutoxin	P83112	EVDMRCKSSKE C LVK C K C QATGRPN C MNR C K C YPR		4274.23	2002	Kv1.1,Kv1.2,Kv1.3,Kv1.4
	<i>Centruroides noxa</i>		<i>α-KTx 1.11</i>	SloTx; Slotoxin	P0C182	TFID V D C TVSK C E C WAP C KA C AFGV C DRGKCMGKK C YV		4085.98	2001	KCa1.1
	<i>Mesobuthus eupeorus</i>		<i>α-KTx 1.16</i>	Toxin MeKTx11-1	C0HJ08	ZFTDVK C TGT C Q C WPV C KKM C F C RPNGKCMNGK C R C YP		4221.16	2015	Kv1.1
	<i>Mesobuthus eupeorus</i>		<i>α-KTx 1.17</i>	Toxin MeKTx11-3	C0HJ07	ZFTDVK C TVTK C Q C WPV C KKM C F C RPNGK C MNGK C R C YS		4253.21	2015	Kv1.1
	<i>Centruroides noxa</i>		<i>α-KTx 2.1</i>	NTx; Noxiustoxin	P08815	TIINVK C TSPK C Q C SKP C K C E C LYGSSAGAK C MNGK C K C YNN-NH ₂	1SXM	4195.06	1982	Shaker,Kv1.1,Kv1.2,
	<i>Centruroides marthae</i>		<i>α-KTx 2.2</i>	MgTx; Margatoxin	P40755	TIINVK C TSPK C Q C LP C PK C A C Q C F C Q C SAGAK C MNGK C K C YRH	1MTX	4179.11	1993	Shaker,Kv1.1,Kv1.1,
	<i>Centruroides limpidus</i>		<i>α-KTx 2.3</i>	CII Tx1; Toxin I.T.	P45629	ITINVK C TSPQ C Q C LRP C K C DRFG C Q C HAG C G C INGK C K C YP		4191.06	1994	Kv1
	<i>Centruroides noxa</i>		<i>α-KTx 2.4</i>	NTx-2, NTx2; Noxiustoxin	Q9TXD1	TIINEK C FAT C Q C WTP C KA C IGSL C Q C SK C MNGK C K C YNG		4183.05	1996	Kv1
	<i>Centruroides limpidus</i>		<i>α-KTx 2.5</i>	HgTx1; Hongotoxin	P59847	TVIDVK C TSPK C Q C LP C PK C A C Q C F C G C IRAGAK C MNGK C K C YRH	1HLY	4220.20	1998	Kv1.1,Kv1.2,Kv1.3,Kv1.4
	<i>Centruroides limpidus</i>		<i>α-KTx 2.7</i>	CII Tx2; Toxin II.T.	P45630	TVIDVK C TSPK C Q C LP C PK C E C I C Y C GRHAGAK C MNGK C K C EC		3905.82	1994	Kv1

Базы данных химических соединений



WHAT ARE YOU LOOKING FOR?

Tylenol



Drugs



Targets



Pathways



Indications



План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- **Базы данных сигнальных/метаболических путей**
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии



KEGG PATHWAY Database

Wiring diagrams of molecular interactions, reactions and relations

Menu PATHWAY BRITE MODULE KO GENES LIGAND NETWORK DISEASE DRUG DBGET

Select prefix

map

Organism

Enter keywords

Go

Help

[[New pathway maps](#) | [Update history](#)]

Pathway Maps

KEGG PATHWAY is a collection of manually drawn pathway maps representing our knowledge on the molecular interaction, reaction and relation networks for:

1. Metabolism

Global/overview Carbohydrate Energy Lipid Nucleotide Amino acid Other amino Glycan
Cofactor/vitamin Terpenoid/PK Other secondary metabolite Xenobiotics Chemical structure

2. Genetic Information Processing

3. Environmental Information Processing

4. Cellular Processes

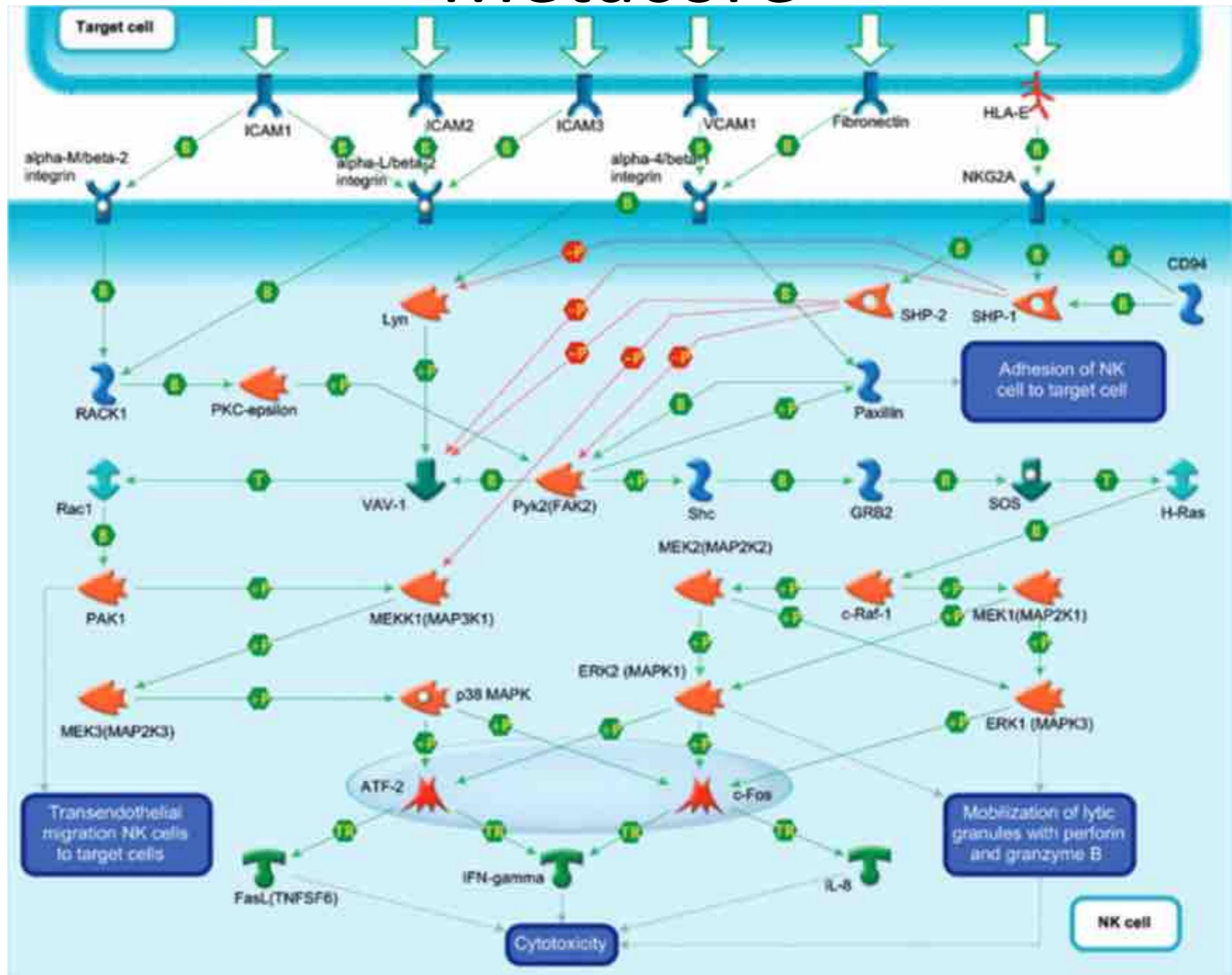
5. Organismal Systems

6. Human Diseases

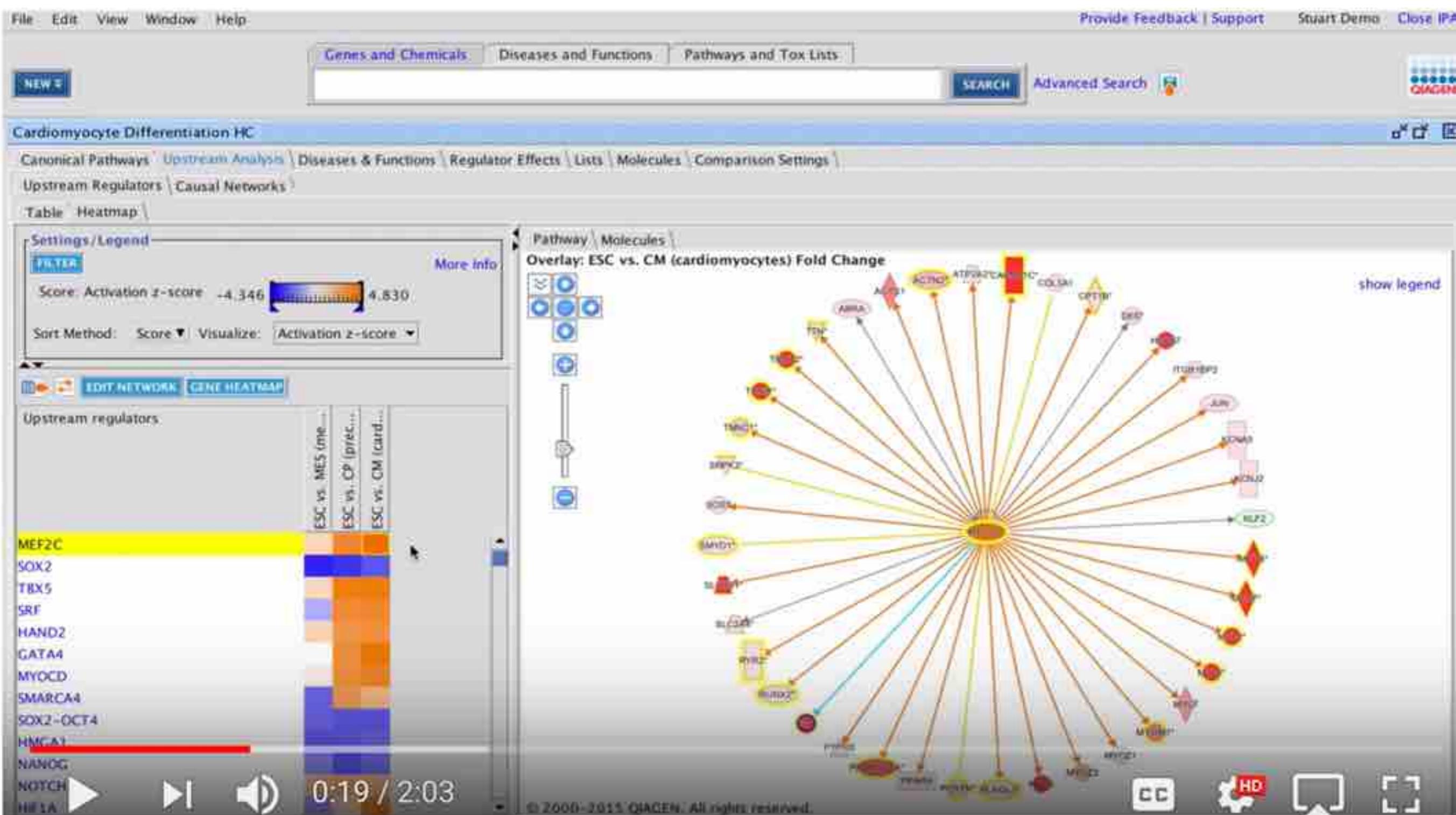
7. Drug Development

KEGG PATHWAY is a reference database for **Pathway Mapping**.

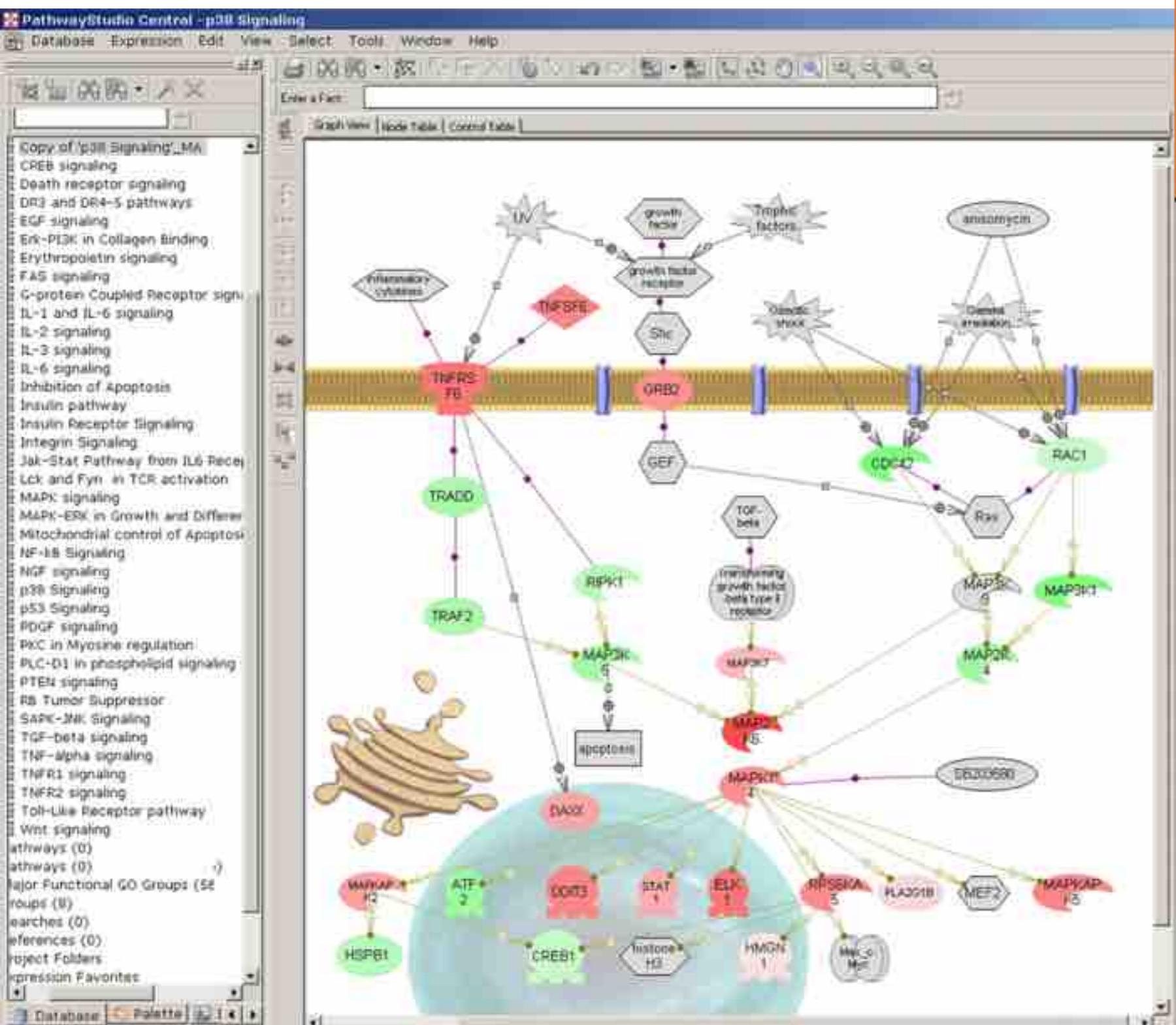
Metacore



Ingenuity Pathway Analysis



Pathway Studio®



План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- **Базы данных секвенирования**
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии



SRA

Sequence Read Archive (SRA) makes biological sequence data available to the research community to enhance reproducibility and allow for new discoveries by comparing data sets. The SRA stores raw sequencing data and alignment information from high-throughput sequencing platforms, including Roche 454 GS System®. Illumina Genome Analyzer®. Applied Biosystems SOLiD System®.

Getting Started

[How to Submit](#)[Log in to SRA \(for updating and troubleshooting submissions\)](#)[Log in to Submission Portal \(for submitting sequence data\)](#)

Tools and Software

[Download SRA Toolkit](#)[SRA Toolkit Documentation](#)[SRA-BLAST](#)[SRA Run Browser](#)

Related Resources

[Submission Portal](#)[Trace Archive](#)[dbGaP Home](#)[BioProject](#)

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- **Базы данных клинических исследований и лекарств**
- Базы данных по экспрессии генов/гистологии
- Базы данных по таксономии



U.S. National Library of Medicine

ClinicalTrials.gov

[Find Studies ▾](#)

[About Studies ▾](#)

[Submit Studies ▾](#)

[Resources ▾](#)

[About Site ▾](#)

ClinicalTrials.gov is a database of privately and publicly funded clinical studies conducted around the world.

Pharmaprojects: track pharma R&D

End-to-end tracking of the global pharma R&D pipeline from bench to patient, including company development trends, global development status, and therapeutic class status.

The industry's most trusted drug development database, CiteLine's Pharmaprojects has been covering pharma R&D across global markets for 35+ years. With 68,000+ drug profiles including 15,000 drugs in active development, it's the go-to resource for preclinical, clinical, and pipeline coverage, and lifecycle management tracking.

-  Curated by experts
-  Drug R&D landscape
-  Historical trends

[Request a free demo ➤](#)

[Product login ➤](#)

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- **Базы данных по экспрессии генов/гистологии**
- Базы данные по таксономии

Базы данных экспрессии генов

Expression Atlas

Enter gene query...
Examples: ASPM, REACT_200624, ENSMUSG00000021789, "zinc finger"

Home Release notes FAQs Download Help About Feedback

Strand-specific RNA-seq of nine chicken tissues
Organism(s): *Gallus gallus*

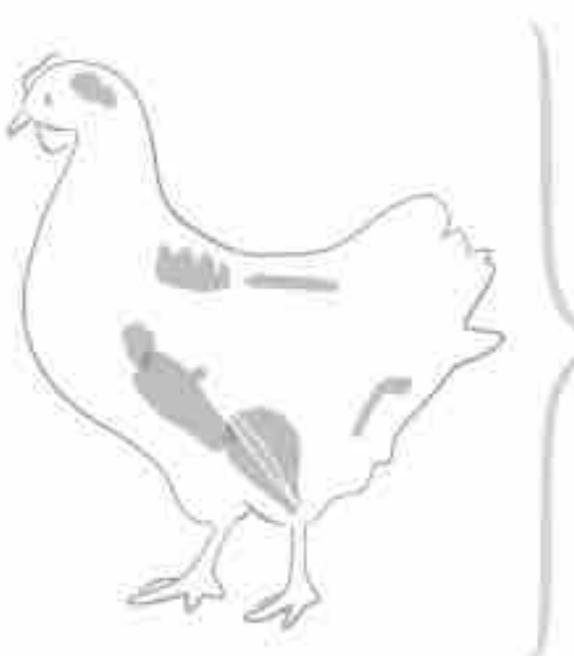
Gene query: KCNA1 | Exact match

Organism part: (any organism parts)
 Specific

Expression level cutoff: 0.4

Showing 1 of 1 genes found:

Gene	brain	colon	heart	kidney	liver	lung	skeletal muscle	spleen	testis
KCNA1									



Базы данных экспрессии генов

NCBI Resources How To Sign in to NCBI

GEO Home Documentation Query & Browse Email GEO

Gene Expression Omnibus



GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.

Keyword or GEO Accession

Search

Getting Started

[Overview](#)

[FAQ](#)

[About GEO DataSets](#)

[About GEO Profiles](#)

Tools

[Search for Studies at GEO DataSets](#)

[Search for Gene Expression at GEO Profiles](#)

[Search GEO Documentation](#)

Browse Content

[Repository Browser](#)

DataSets: 4348

Series: 94968

Platforms: 18150

THE HUMAN PROTEIN ATLAS



≡ MENU

HELP

NEWS

SEARCH¹

Search

Fields »

e.g. RBM3, insulin, CD36



TISSUE ATLAS



CELL ATLAS



PATHOLOGY ATLAS

План

- Библиографические/реферативные базы данных литературных источников (статьи, тезисы, патенты, материалы конференций и т.д.)
- Базы данных последовательностей ДНК
- Базы данных последовательностей белков
- Базы данных 3D структур
- Базы данных хим. соединений
- Базы данных геномов и аннотаций
- Базы данных вариаций генома
- Базы данных геном-фенотип
- Базы данных взаимодействий
- Базы данных сигнальных путей
- Базы данных секвенирования
- Базы данных заболеваний и медицинской информации
- Базы данных по экспрессии генов/гистологии
- **Базы данных по таксономии**

Таксономические базы данных

Таксоному Browser – знаменитая таксономическая БД, имеющая иерархическую структуру, основанную на анализе последовательностей и призванная упорядочить классификацию организмов, для которых известна хотя бы одна последовательность ДНК или белка.

The screenshot shows the NCBI Taxonomy Browser interface. At the top, there's a navigation bar with links to PubMed, Entrez, BLAST, OMIM, Taxonomy (which is highlighted in blue), and Structure. Below the bar is a search input field with the placeholder "Search for" and a dropdown menu set to "complete name". There are also "lock", "Go", and "Clear" buttons. To the left, a sidebar contains links to "Taxonomy browser", "Archaea", "Bacteria", "Eukaryota", "Virroids", "Viruses", "Taxonomy common tree", "Taxonomy information", "Taxonomy resources", "Taxonomic advisors", and "Genetic codes". The main content area features a banner with illustrations of a mushroom, a gear, a fish, a flower, and a DNA helix, followed by the text "The NCBI Taxonomy Homepage". Below this, a section titled "These are direct links to some of the organisms commonly used in molecular research projects:" lists several organisms with their names underlined for linking:

Arabidopsis thaliana	Escherichia coli	Pneumocystis carinii
Bos taurus	Hepatitis C virus	Rattus norvegicus
Caenorhabditis elegans	Homo sapiens	Saccharomyces cerevisiae
Chlamydomonas reinhardtii	Mus musculus	Schizosaccharomyces pombe
Danio rerio (zebrafish)	Mycoplasma pneumoniae	Takifugu rubripes
Dictyostelium discoideum	Oryza sativa	Xenopus laevis
Drosophila melanogaster	Plasmodium falciparum	Zea mays

Видовые базы данных

Содержат таксономическую, библиографическую, географическую, визуальную и прочую информацию

 AlgaeBase content about team notulae algarum links contact search
genus · species · literature · journals · images · common names · distribution · glossary · taxonomy browser · higher taxonomy

143,152 species and infraspecific names are in the database, 18,634 images, 54,056 bibliographic items, 314,920 distributional records.

Haematococcus pluvialis Flotow

Publication details
Haematococcus pluvialis Flotow 1844: 415, 537, pls XXIV, XXV
Published In: Flotow, J. von (1844). Beobachtungen über *Haematococcus pluvialis*. *Verhandlungen der Kaiserlichen Leopoldinisch-Carolinischen Deutschen Akademie der Naturforscher* 12(Abt. 2): 413-606, 3 pls.
[Download PDF](#)

Type species
This is the type species (lectotype) of the genus *Haematococcus*.

Status of name
This name is of an entity that is currently accepted taxonomically.



algaebASE

Haematococcus pluvialis Flotow India

Classification:
Empire Eukaryota
Kingdom Plantae
Subkingdom Viriplantae
Infrakingdom Chlorophyta
Phylum Chlorophyta
Subphylum Chlorophytina
Class Chlorophyceae
Order Chlamydomonadales
Family Haematococcaceae
Genus Haematococcus

Taxonomy

References

Submit Feedback

Submit Reference

Links

Видовые базы данных

<https://plant.depo.msu.ru>

Микроорганизмы грибы Растения

RU EN Справка
Вход в систему

Сейчас в базе данных:
Объектов: 501892 Изображений: 501892 Видов: 9141

Национальный банк-депозитарий живых систем

Проект Московского университета "Ноев ковчег" посвящен созданию многофункционального сетевого хранилища биологического материала.

Планируется работа с материалом всех возможных типов - от отдельных биологических молекул до целых живых организмов.

Создание депозитария позволит сохранить биоразнообразие нашей планеты и создать новые способы полезного использования биологического материала.

Фото дня

Организм недели

Geum coccineum

Juncus compressus

2015-2017

Подать образец