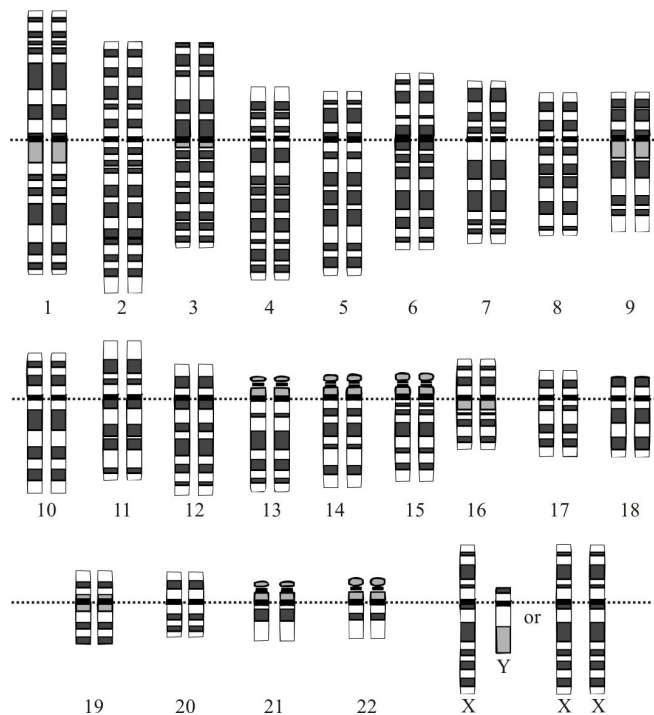


Геном человека



Wikimedia commons

Надежда Потапова

nadezhdalpotapova@gmail.com

26.02.2020

ФББ МГУ, ИППИ РАН

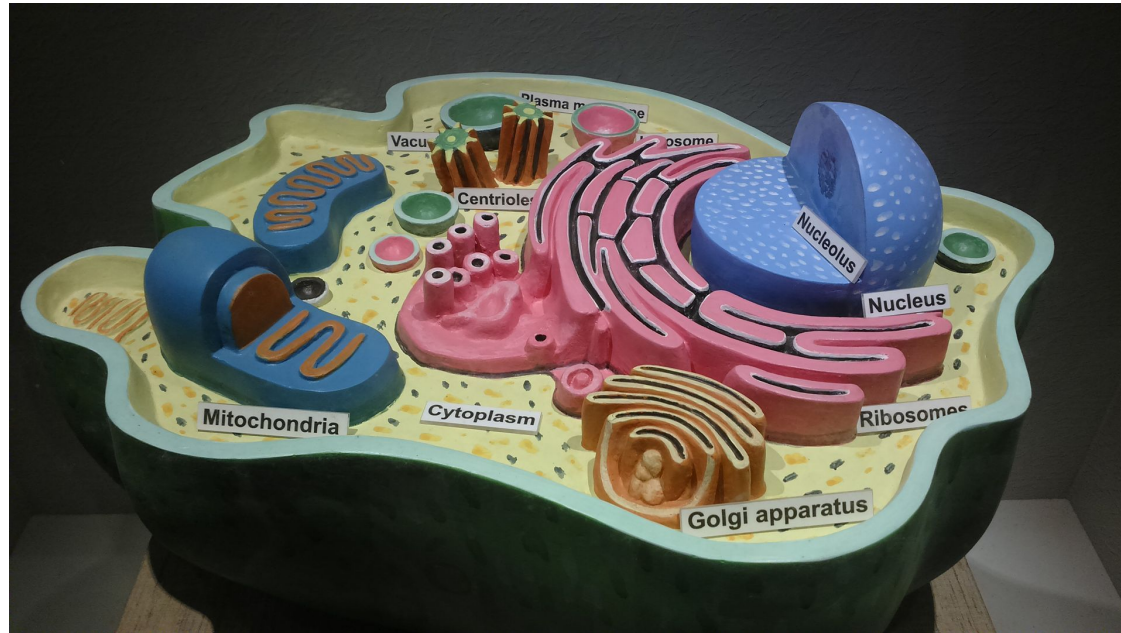
План нашей лекции

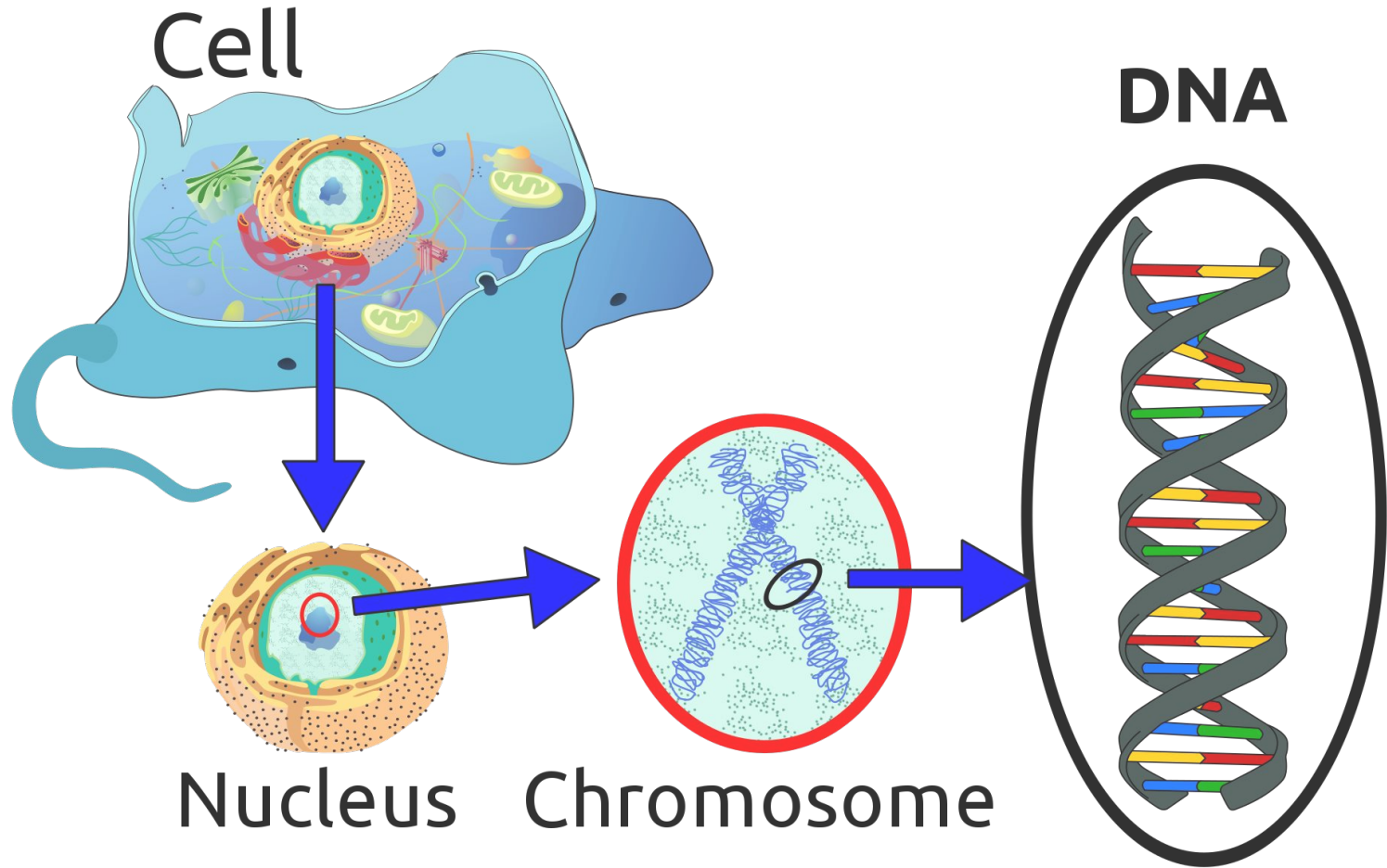
1. Что такое геном и где он находится?
2. Что было известно про геном человека до проекта “Геном Человека”?
3. Почему начали этот проект?
4. Как проект проходил и что могло (гипотетически) его ускорить?
5. Какие результаты проекта?
6. Что теперь?

Что такое геном и где он находится?

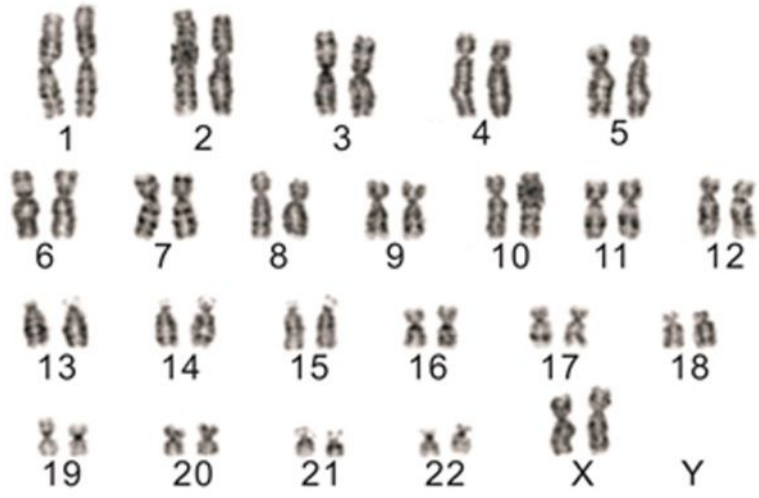
Геном

совокупность наследственного материала, заключенного в клетке организма



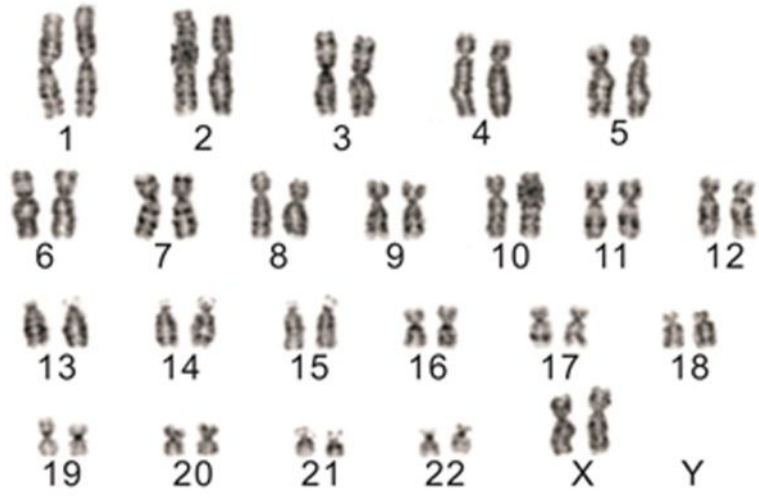


Что было известно про геном человека до проекта “Геном Человека”?



↑
идиограмма

doi: 10.1371/journal.pone.0081844



↑
идиограмма

doi: 10.1371/journal.pone.0081844



<https://president.utexas.edu/past-presidents/theophilus-shickel-painter>

Теофилус Пейнтер

Не был уверен в точном числе хромосом у человека:

“In my own materials the counts range from 45 to 48 apparent chromosomes, although in the clearest plates so far studied only 46 chromosomes have been found.”

(Painter, T. S. The Y-chromosome in mammals. *Science* 53, 503–504 (1921)).

Решил, что от 46 до 48 хромосом.

В другой своей статье тоже сомневается, но решает остановиться на 48.



Figure 2 | **Camera lucida drawing of a human spermatogonial metaphase made by Theophilus S. Painter.** The drawing presumably shows 48 chromosomes. The cytological methods used were state of the art for the time, but it is still difficult, if not impossible, to make an exact count. Reproduced with permission from REF. 25 © (1979) Springer Verlag.

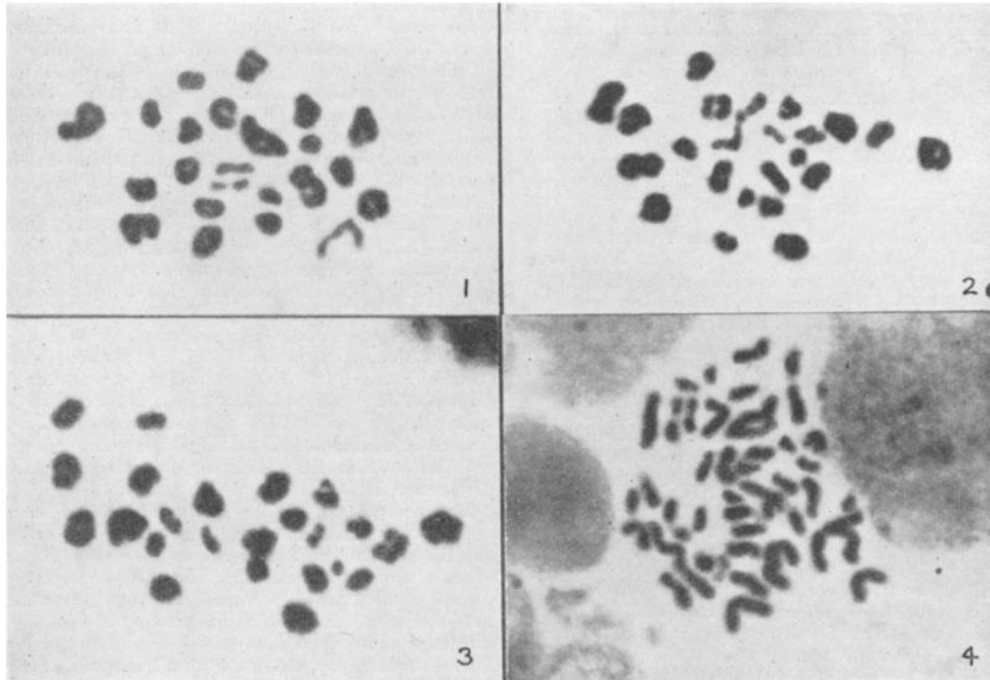


Fig. 1. First spermatocyte of patient 1 at diakinesis, 23 bivalents
 Fig. 2. First spermatocyte of patient 3 at diakinesis, 23 bivalents
 Fig. 3. First spermatocyte of patient 2, 22 bivalents plus univalent X- and Y-chromosomes
 Fig. 4. Spermatogonial metaphase of patient 1, 46 chromosomes
 All $\times 2,800$

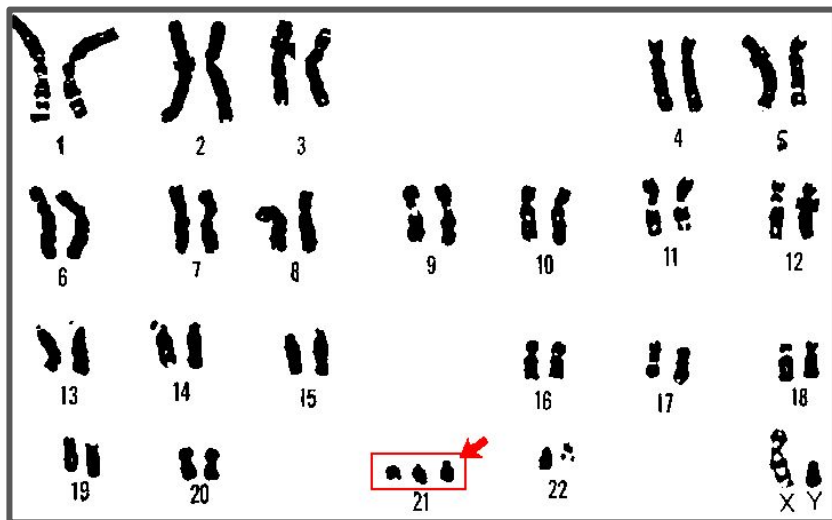
“We do not wish to generalize our present findings into a statement that the chromosome number of man is $2n=46$, but it is hard to avoid the conclusion that this would be the most natural explanation of our observations.” (Tijo, Levan, 1956)

Спустя 33 года, в 1956 году, Tijo and Levan (а вдобавок Форд и Хамертон, и др.) объявили — у человека 46 хромосом (doi:10.1038/1781020a0)

Что изучать по хромосомам?

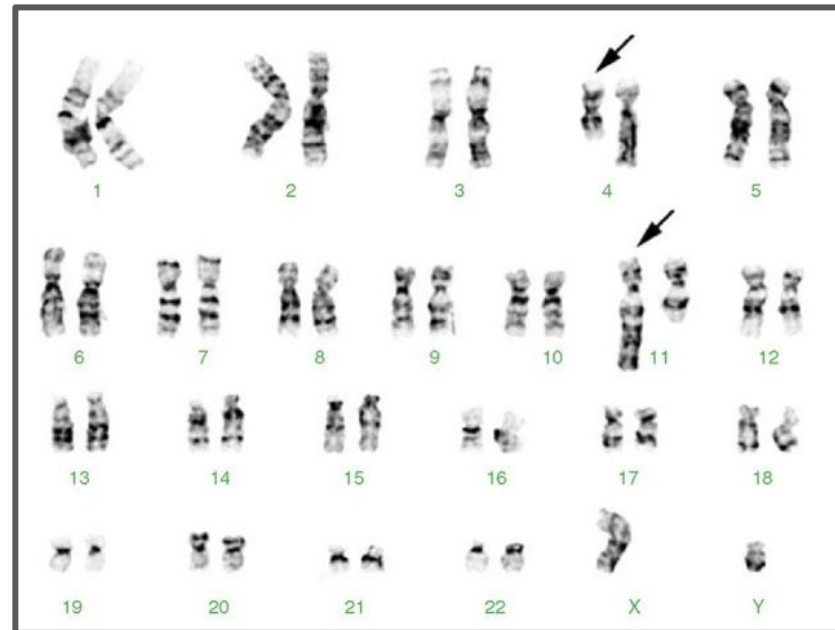
- Крупные мутации

но для этого вообще надо понимать с чем работаем и сколько хромосом. И если мы рассматриваем какой-то определённый участок, то где он вообще находится.



Wikimedia commons

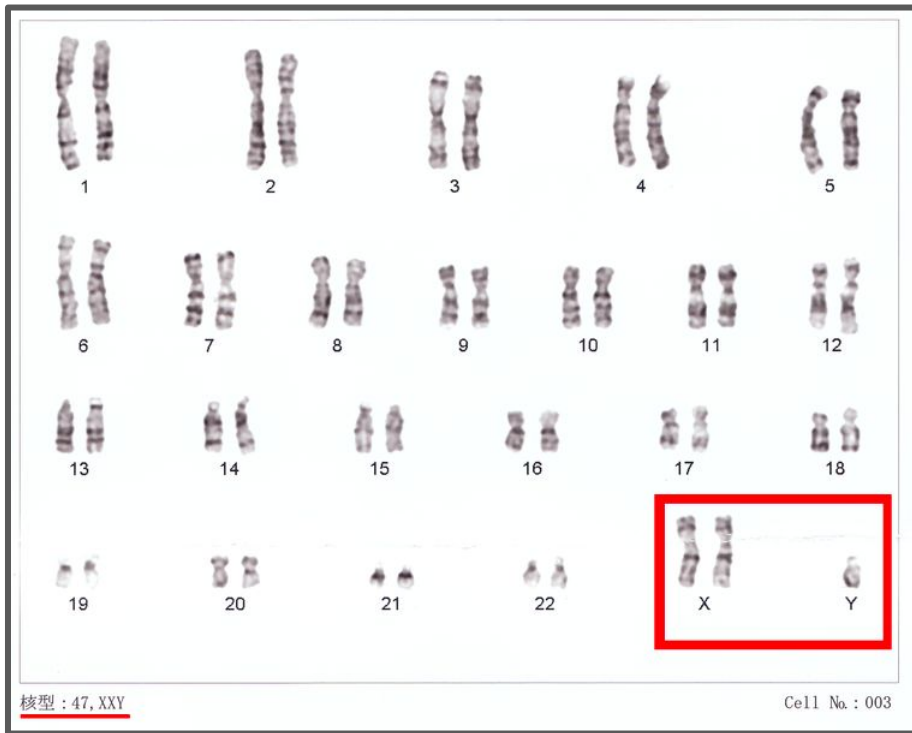
Трисомия по 21 хромосоме



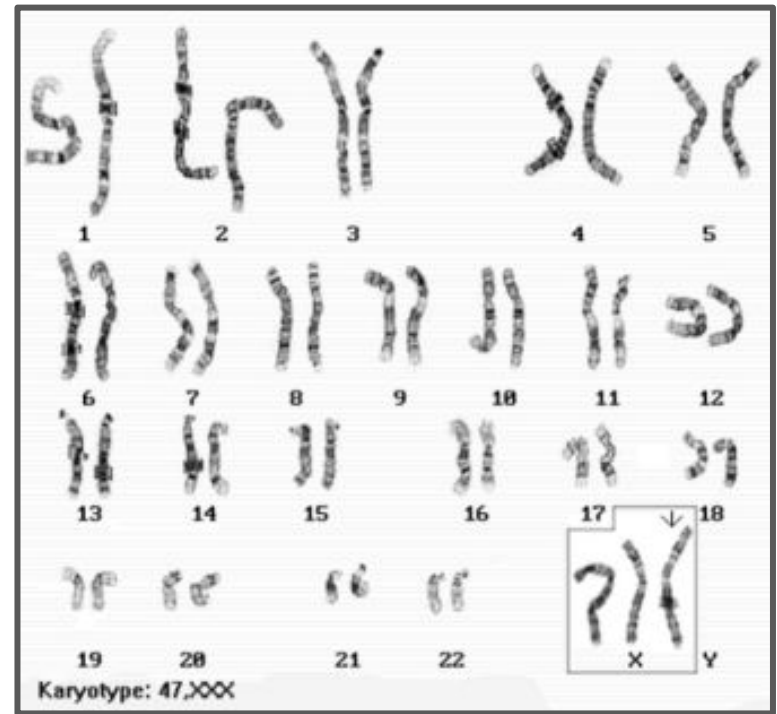
<http://what-when-how.com/>

Пациент с острым лимфобластным лейкозом.
Транслокация с 4 на 11 хромосому.

неконтролируемое деление незрелых лимфоидных
клеток (лимфобластов)

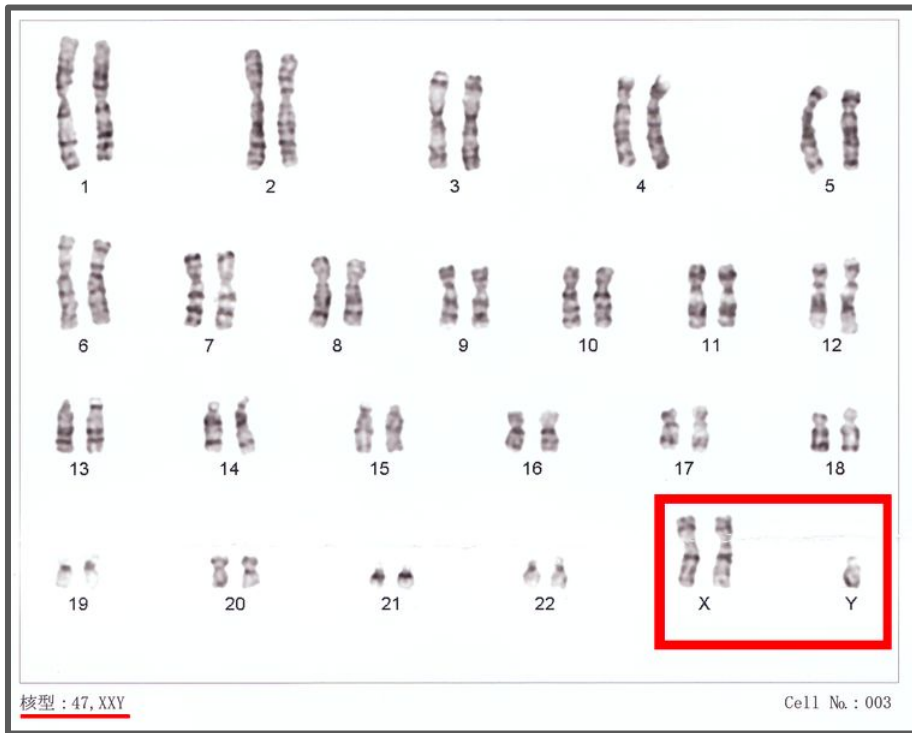


Wikimedia commons

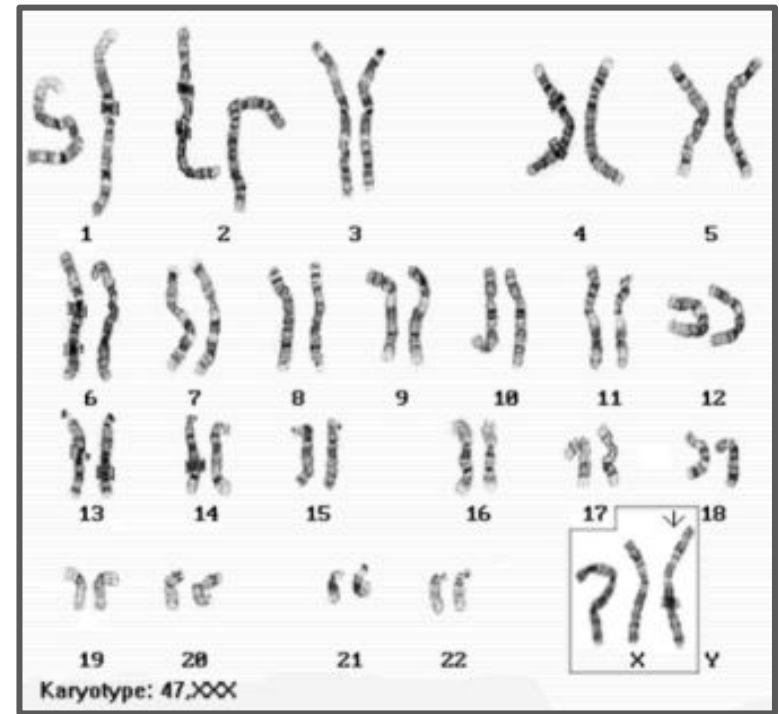


<https://triplexyndrome.wikispaces.com/>

Синдром Клайнфельтера



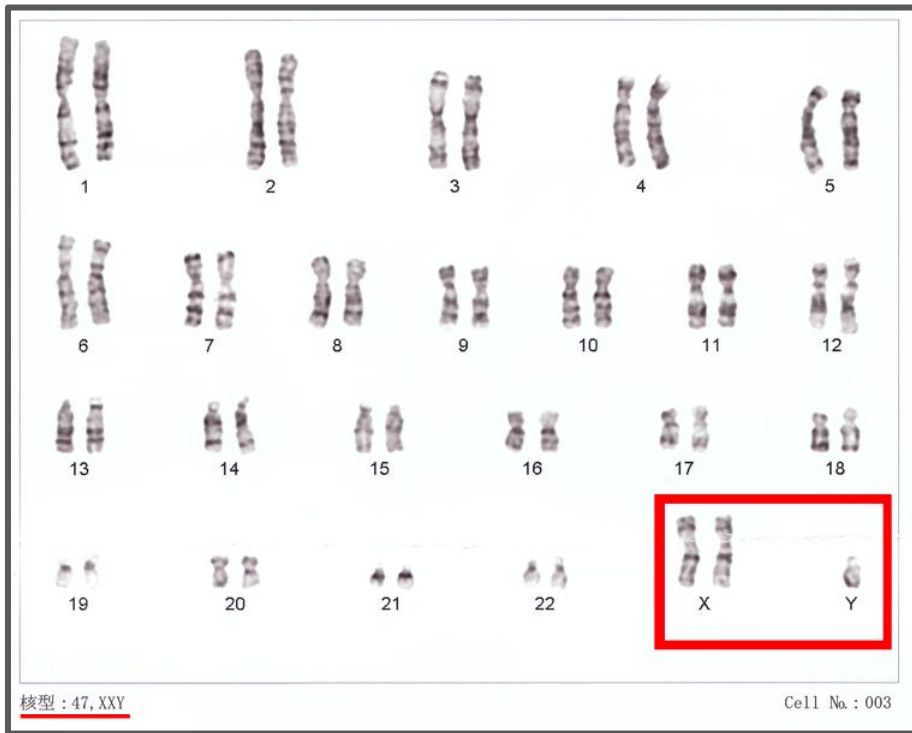
Wikimedia commons



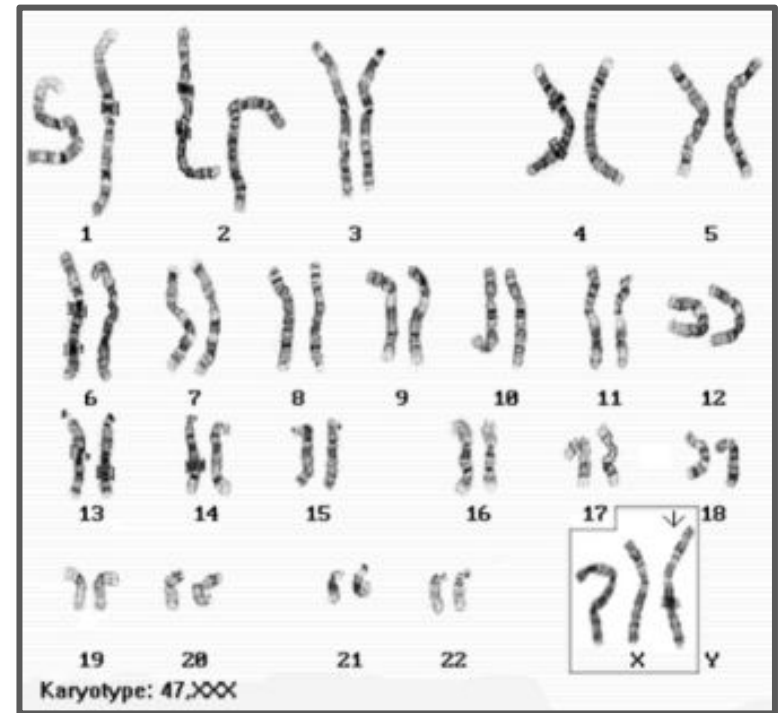
<https://triplexsyndrome.wikispaces.com/>

Синдром Клайнфельтера

47, XXY; 47, XYY; 48, XXXY; 48, XYYY; 48 XYY; 49 XXXXY; 49 XXXYY



Wikimedia commons

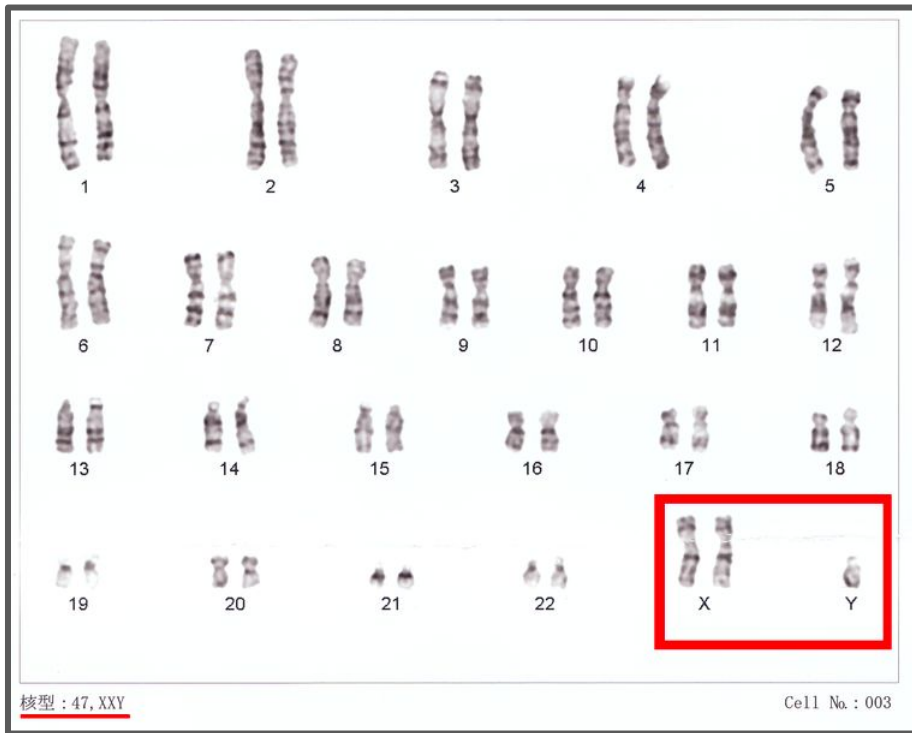


<https://triplexyndrome.wikispaces.com/>

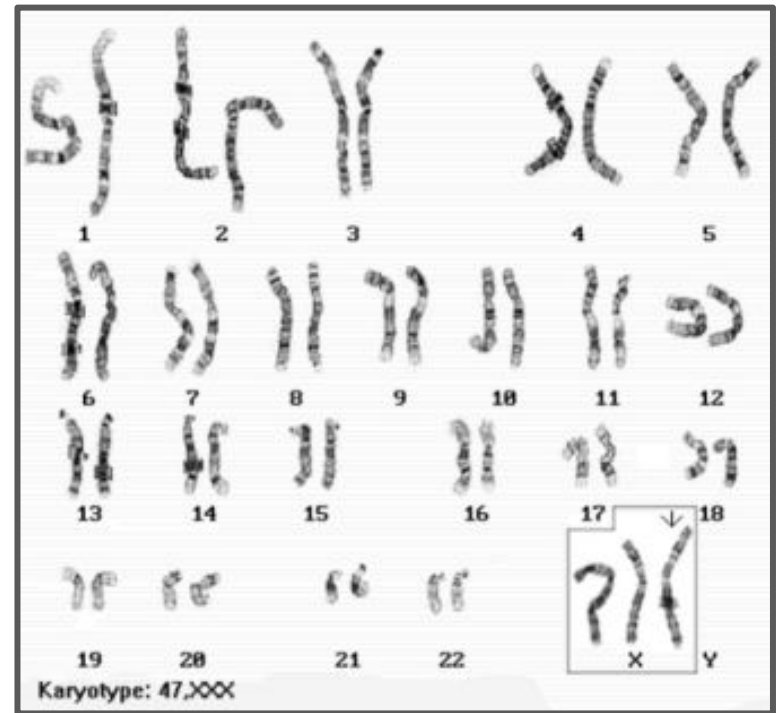
Синдром Клайнфельтера

1 на 500-700
новорождённых
мальчиков

47, XXY; 47, XYY; 48, XXXY; 48, XYYY; 48 XHYY;
49 XXXXY; 49 XXXYY

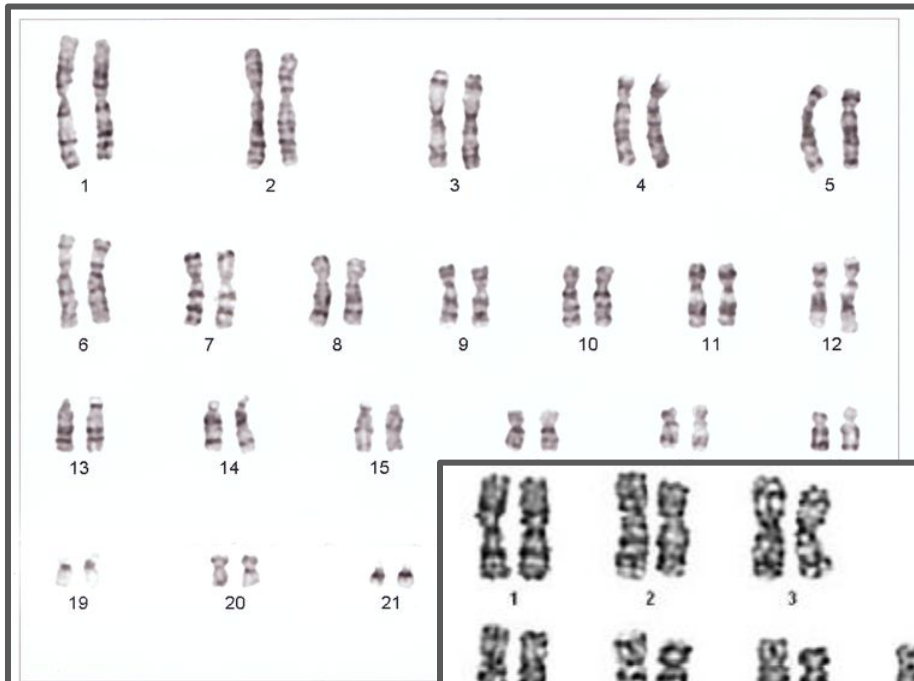


Wikimedia commons



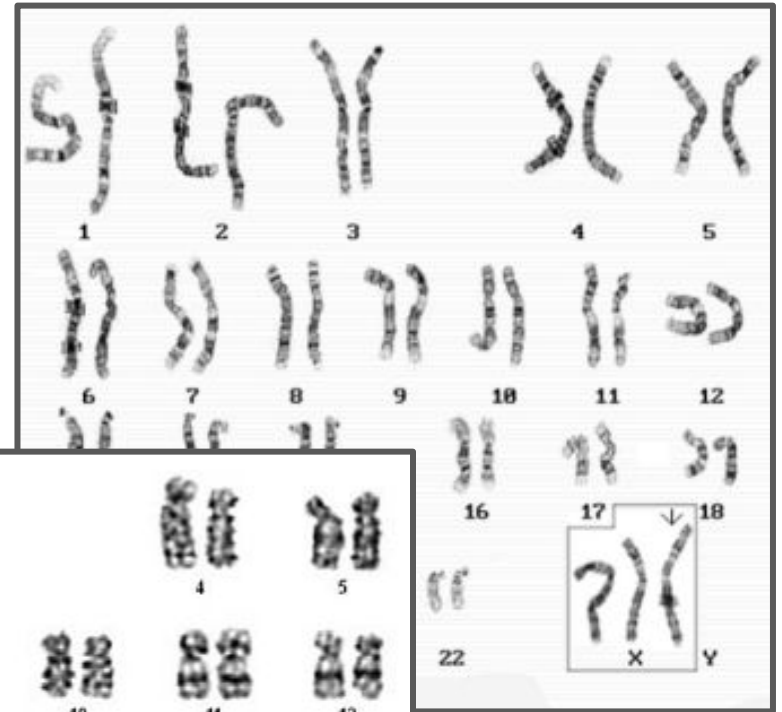
<https://triplexyndrome.wikispaces.com/>

Синдром Клайнфельтера

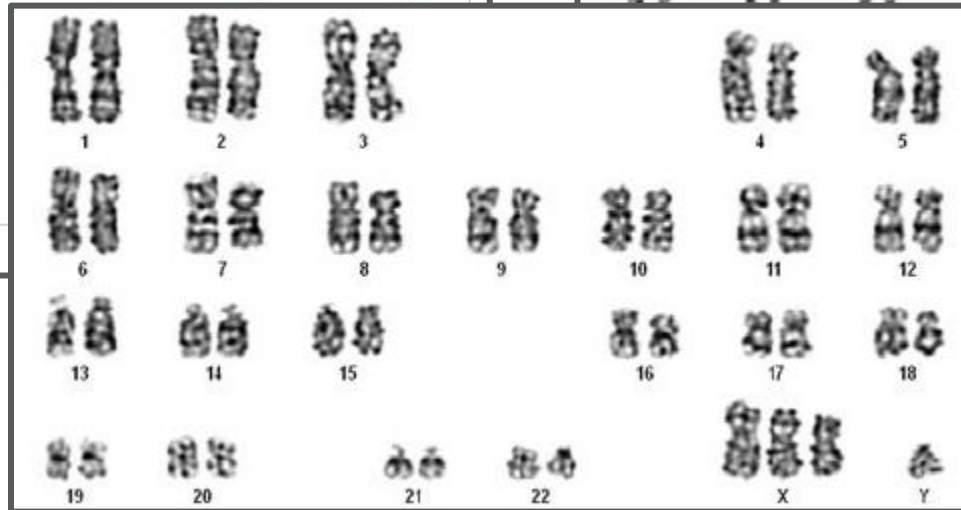


核型 : 47, XXY

Wikimedia commons



<https://triplexsyndrome.wikispaces.com/>



doi: 10.1007/s12098-009-0288-6.

Хромосомы – это хорошо!
Но давайте будем смотреть геномы!

Почему начали этот проект?



За и против проекта “Геном Человека”

За и против проекта “Геном Человека”

За:

1. Поможет улучшить диагностику и лечение заболеваний
2. Улучшить разработку лекарств
3. Поможет в криминалистике
4. Улучшит экономику (рабочие места)
5. Сможем изучать и геномы других организмов

За и против проекта “Геном Человека”

Против:

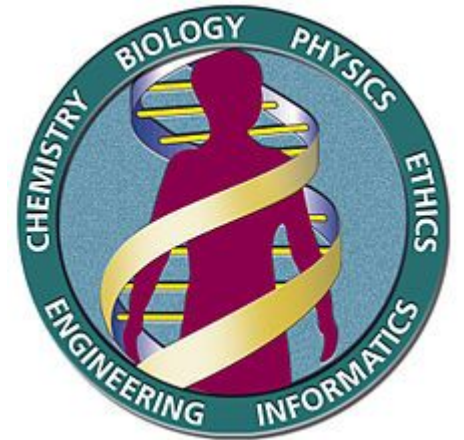
1. Начнётся мода на внешний вид, генетически опосредованный
2. Дискриминация по генотипу (посмотрите фильм “ГАТТАКА”)

The Human Genome Project

- 1990 год — начало проекта
- Скептицизм: зачем это надо и стоит ли оно того?

Теперь понятно, это новая веха в науке, медицине.

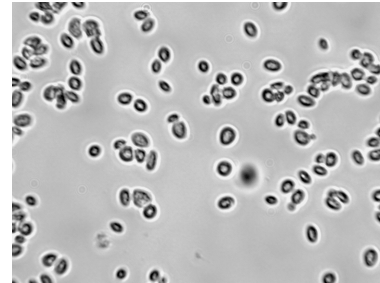
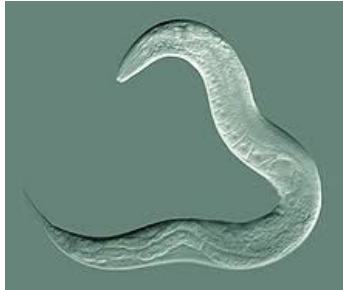
- Приняли участие 18 стран
- 2001 год — конец проекта (черновик)
- 2003 год — совсем конец проекта (чистовик)



The Human Genome Project

Первые цели:

1. Сделать физические и генетические карты геномов человека и мыши.
2. Секвенировать маленькие геномы дрожжи и червяка, чтобы потестировать, как будут секвенироваться большие геномы, и подготовиться.
3. Когда сделали червяка и дрожжу — время написать и показать их геномы



Принципы проекта

1. Принимали любых коллабораторов вне зависимости от страны и т.д. (чтобы все вместе и быстрее)
2. Информация о геноме человека будет в открытом доступе и бесплатна в течение 24 часов после её получения (стимулирует развитие науки, новые и быстрые открытия)

(International Human Genome Sequencing Consortium, 2001)

Фазы проекта

1. Shotgun phase — делить ДНК на мелкие фрагменты и секвенировать их.
2. Finishing phase — заполнять оставшиеся пробелы и разрешать сложные места в геноме.



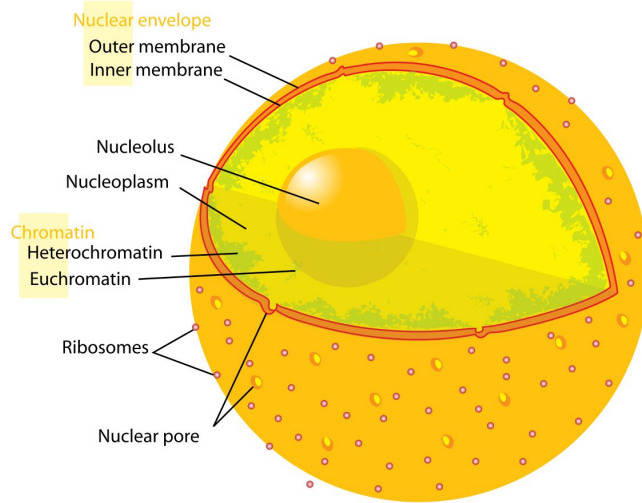
Hank Morgan/SPL

Early days: a DNA-sequencing lab in 1994.

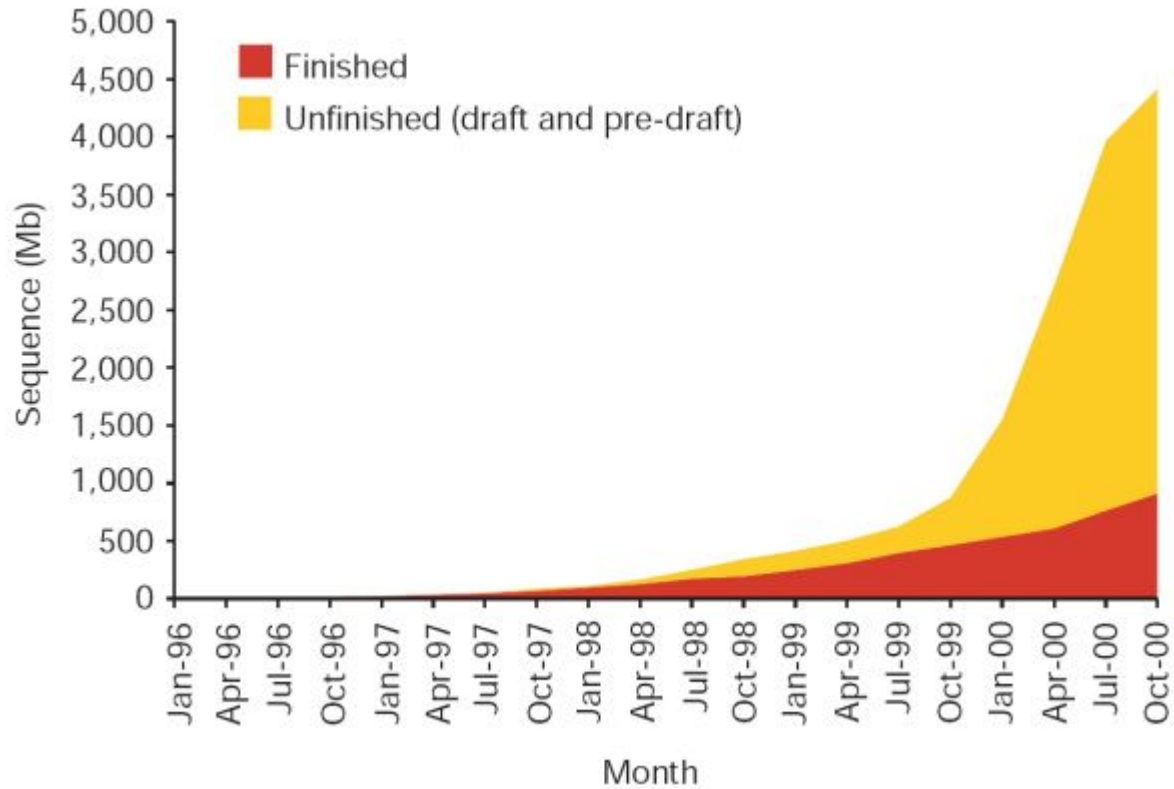
Что хотели посмотреть

1. Последовательности нуклеотидов — секвенирование
2. Где определённый ген расположен на какой хромосоме
3. Как гены и фрагменты генома располагаются друг относительно друга.

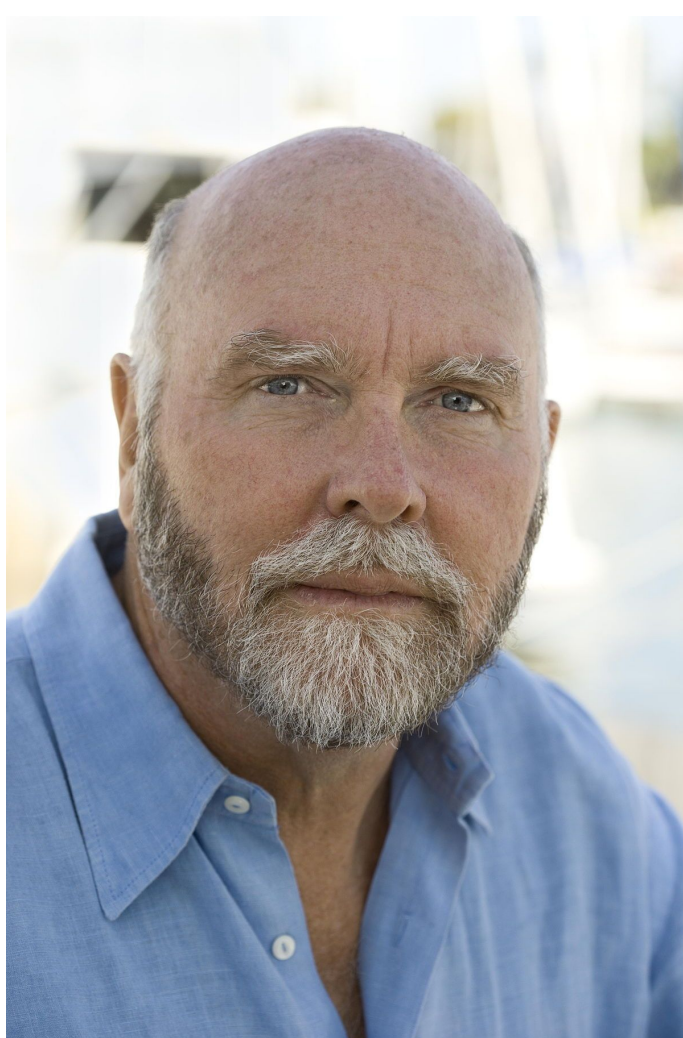
Смотрели только эухроматин — это ~92% генома. Гетерохроматин (центромеры, теломеры) не смотрели.



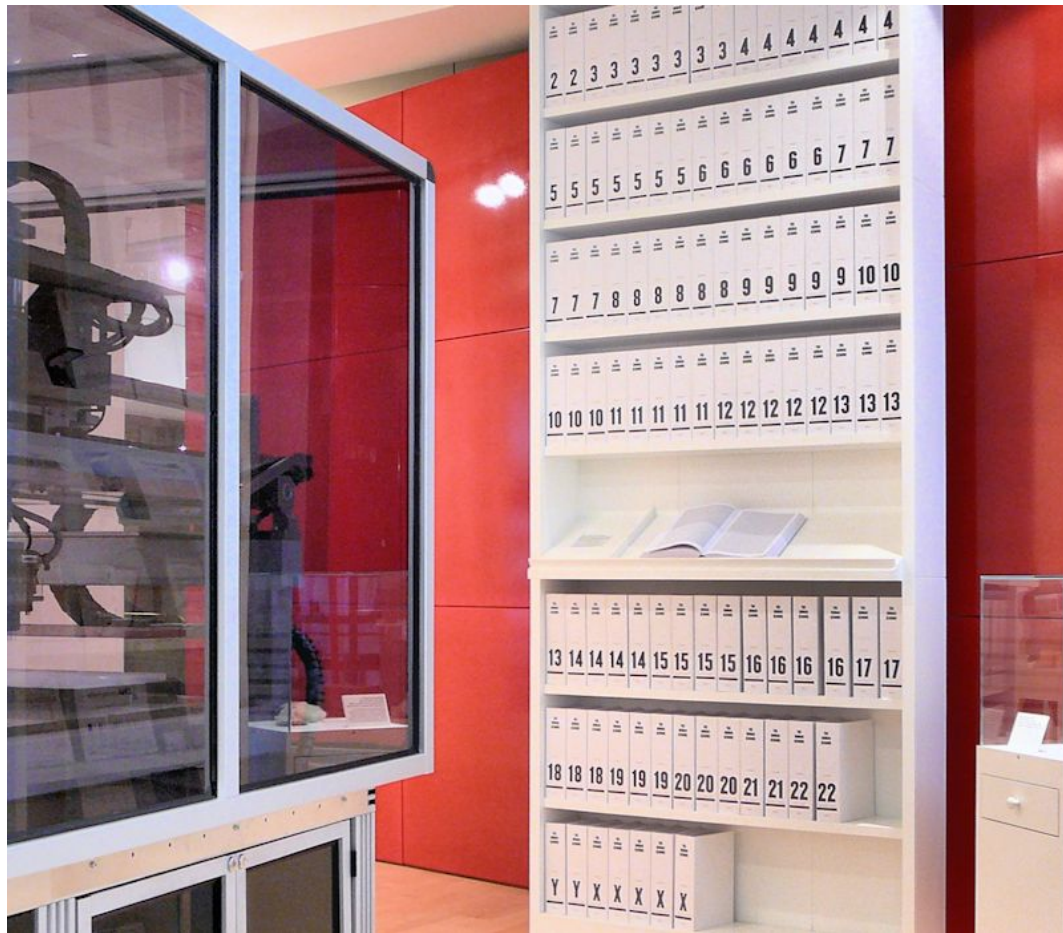
https://en.wikipedia.org/wiki/Heterochromatin#/media/File:Diagram_human_cell_nucleus.svg



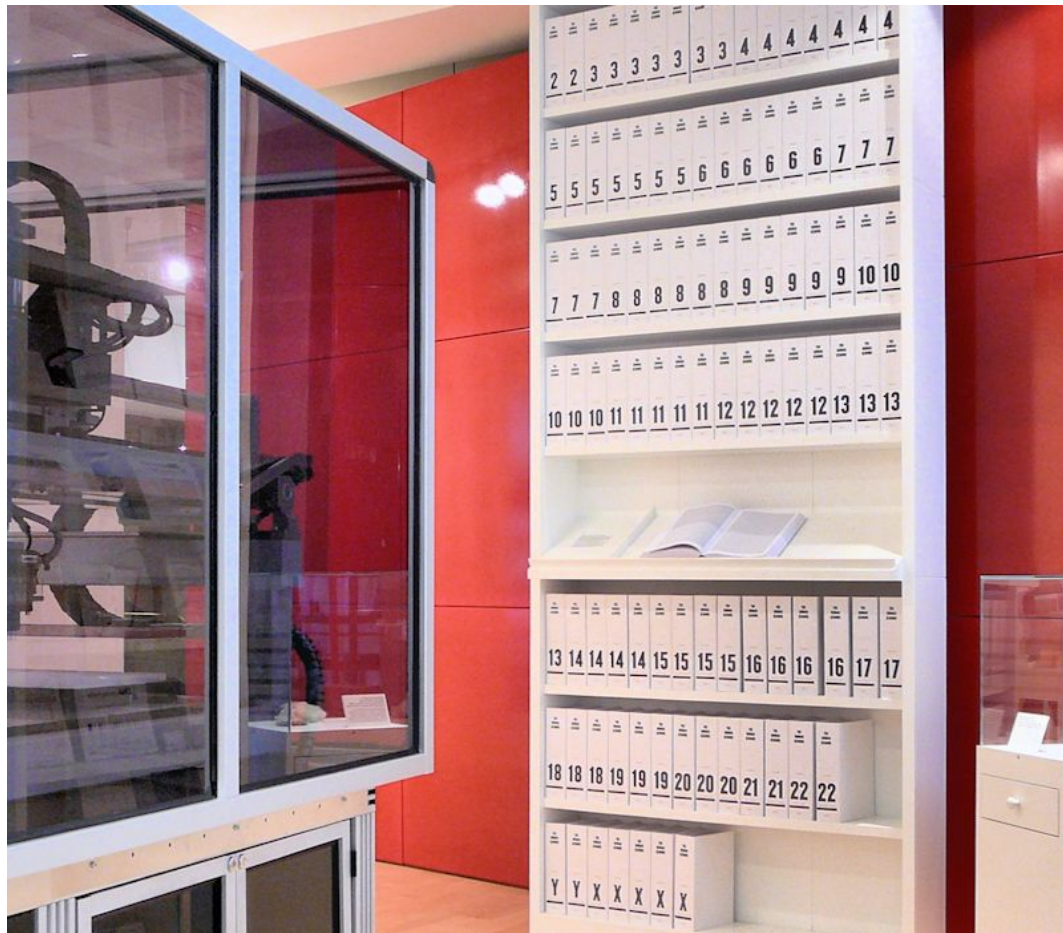
Nature **409**, 867 (2001)



Крейг Вентер



ЧЕЙ ГЕНОМ?



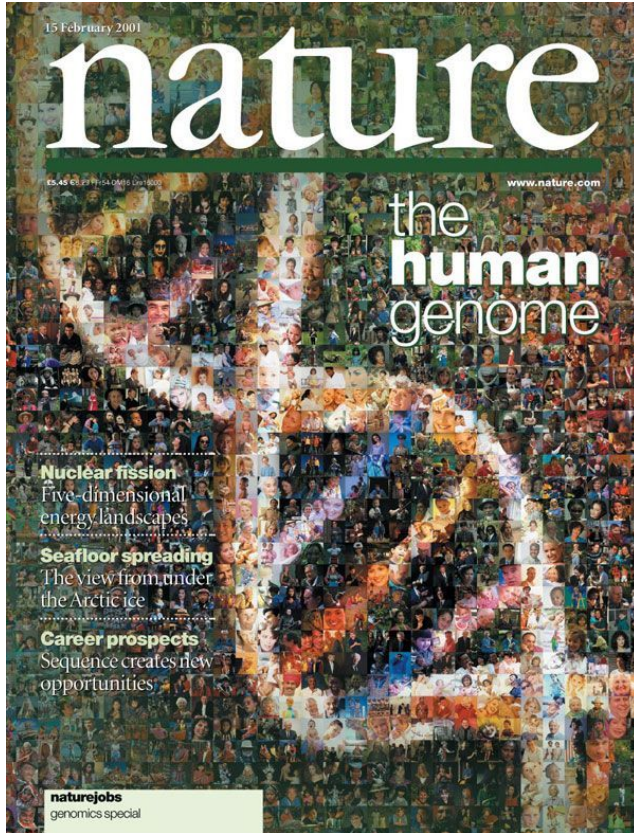
Проект Вентера — из 21 имевшихся образцов ДНК добровольцев взяли для 5 человек, их и секвенировали

Проект Коллинза — из 40 взяли 4 (2 мужчины и 2 женщины) и отсекувенировали (есть мнение, что большая часть отсекувенированной ДНК — от так называемого “Мужчины из Буффало”)



Джеймс Уотсон

Same result



Результаты The Human Genome Project

- Получили гаплоидный геном
- Длина генома : 3.2 млрд пар нуклеотидов с ошибкой 1 на 100 000 нуклеотидов
-

Результаты The Human Genome Project

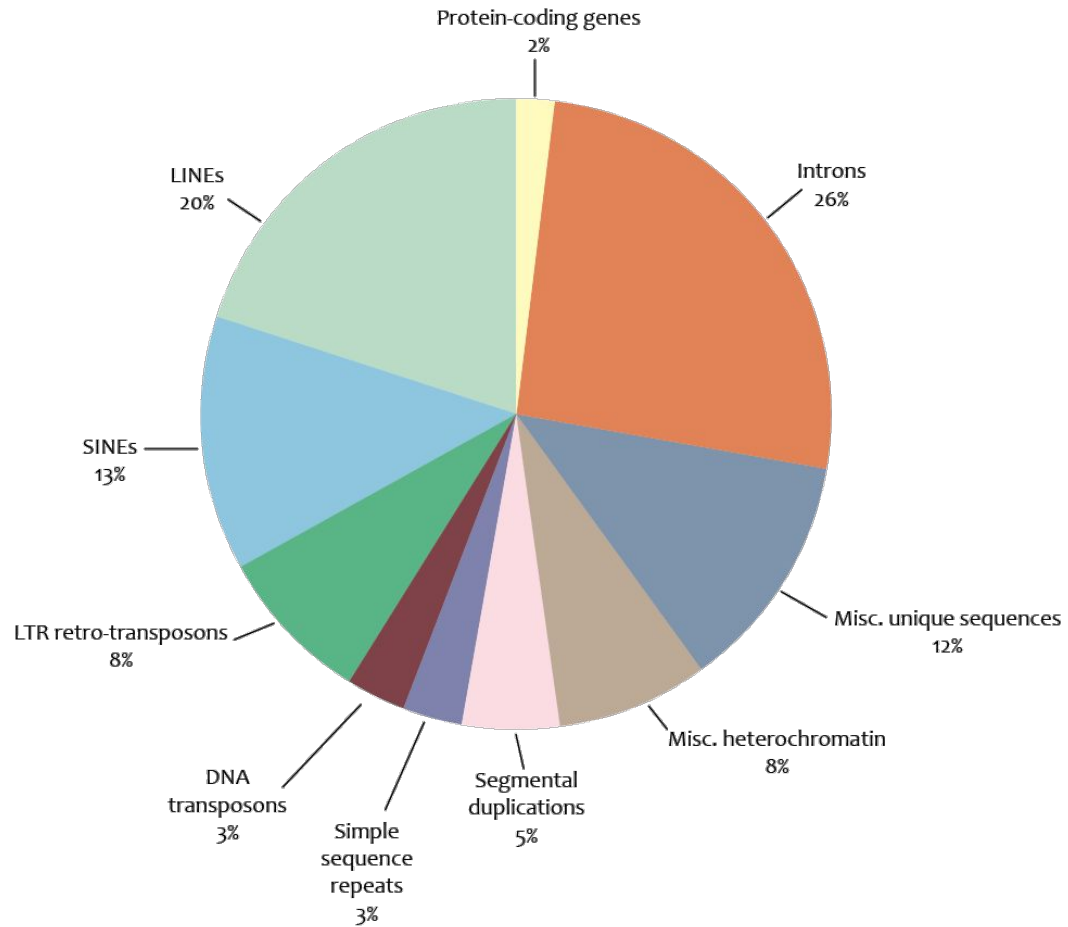
- Получили гаплоидный геном
- Длина генома : 3.2 млрд пар нуклеотидов с ошибкой 1 на 100 000 нуклеотидов
- Около 341 “пробела” в геноме. Это сложные для секвенирования места.
- Черновая версия: генов около 40 000 штук,
чистовая версия: генов около 20 000 — 25 000 штук
*думали — 100 000 штук
-

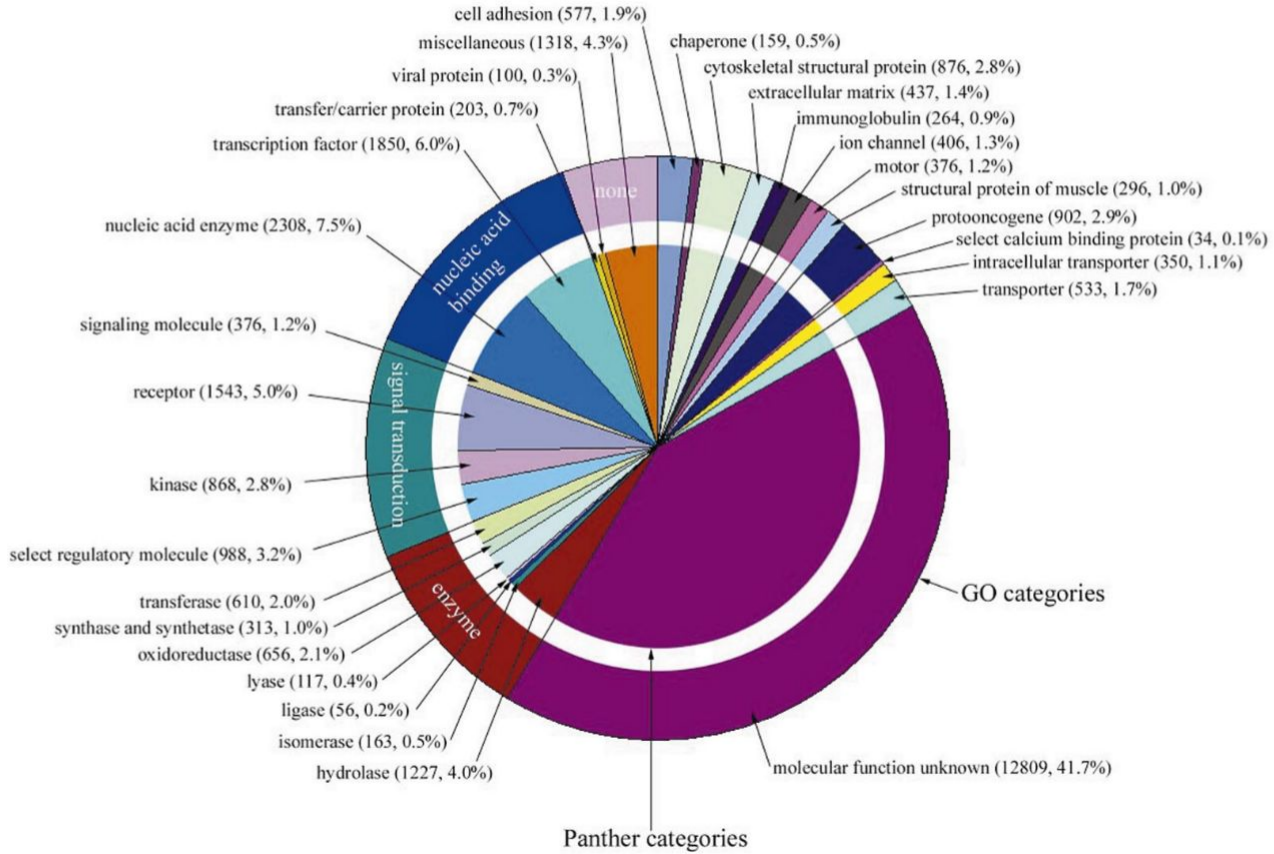
Результаты The Human Genome Project

- Получили гаплоидный геном
- Длина генома : 3.2 млрд пар нуклеотидов с ошибкой 1 на 100 000 нуклеотидов
- Около 341 “пробела” в геноме. Это сложные для секвенирования места.
- Черновая версия: генов около 40 000 штук,
чистовая версия: генов около 20 000 — 25 000 штук
*думали — 100 000 штук
- Последовательность геномов почти идентична у двух любых людей (99.9%)
- Стоимость проекта: 3 млрд долларов (23 000 квартир в Москве)

CAAAATATACAAACAAATTTCTAATAACAAAAAGAAAAACAACCCAAATTAAGATAGACACCTCATCAA
AAAGATATACAGATGACATGCCAGGTGCAGTGGCTCATGTCTGTAATCTCAGCACTTTGGGAAGAGGAG
GCTGGCAATCACTTGAGGCCAGGAATTTGAGACCAGCCTGGCCAAACATGGCAAATCCTGTATCTACTAA
AAATACAAAAATTAGCTGGGAGTGGTGGCTCACACCTGTAATCCACGCTACTTGGGAAGTGGAGGCAGGA
GAATCACTTGAACCGGGGAGGCAGAGGTTGCAGAAAGTTGAGATCATGCCACTGCACCTTCAGCCTGGGCG
ACAGACCAAGACACCAATCTTAAAAAAGGAGTTGACTCTATCACCCACTTTGTCACAGCCAAAGAACTAACAG
AAGCTGAAGACAGGCCAAAAACGGATACTTCCCTAGTGCCTTCAGAGGGAGCATGGCCCTGCTGAGACAG
AGTAGAAATAGGACTCGGCCACCCCACTAACGTGTTGTTCCATACATGCCCGCTGACCACCCAGACCTT
GCCGCACCCCTCCAGGTGCTATACTCGCTGGCCAGACCTTGCCAACTGTATGAAATAAACTAAGATAA
GCAGCATCTGGCCGTAAGTCTTACTCAAGGGAGTTGACTCTATCACCCACTTTGTCACAGGCCAGGAGA
ATGACTGATCCTTACCCCTGGCCTCATTATAACTAAAGTCCCACCCAGGTATAGGCTTATCTGCTAT
TTTTGATCATGCAACGTATGTGTTAGCACGATTCTTACTGCATCTGAGCACCCCTTCACTCCACCCGT
AGGTGTAACGACACTCACCTAACTATAAAGATGTCATGTCACCCCTTAAGACCCACAATGCACTCC
CCTTGGTGAGCCAGCCAAGAATCCTTCCCTGGCCAACCTTAAGGGCATAAGACTTAATACGTCTCTCT
TCTGCTCAACCTTTAGGGCATAAAGCTTAATAAAGCCTTGCTGGGAACTTGGCTTGGCCTTGTATCAA
CTTCTGTTTCATGGGAGCCTAAGAACATGTAGTCAGGAATGCTGCCAACATCTTTATTTCCAACTTCTGG
ACTTCAGAAGTGAAGCAAACTCTGTGGTCTAAGCTACCCATTTTGTGGCACTTTGTACGGCAGCCCT
AGGAAACAACGCAGGCACGTTCTGTCCATCTGCCGCTCTCTCAGACACAGCTCTTCTTCAAACCCA
CCCCATGAGAGGCTGAGGACAGAAGCAGGACCCCTTACAGGCCACCATCACTGCTGGATGGTCATTTCCA
GCACTGACTGGGAGCTCCAGCCTCATCTCTCAGTTTCTAGGAGCAAGTGGGAGGATGAGGACAAGGAGGA
ATGAGGGCTCTGGCCTCCAGGAAGTGTCCCTGAACCTCCACCTGCTGTCCATGAAGCTGCCCTGAGC
CTCCACCCGACCTCAGCCAGTGTCTCCCTGAGCCTCCACCTGCTGTCCCGAGAGTGCCTTGAAC
CTCCACCCGACCTCAGCCAGAGCTGTCCCTGAGCCTCCACTACTGCTGTCCCGAGCTGCCCTGAG
TCTCCACCCCTTACCTTAGCCAGAGCTGTCCCTGAACCTCCACCTGCTGTCCCGGAGCTGCCCTGAG
CCTCCACCCCTTACCTTAGCCAGAGCTGTCCCTGAACCTCCACCTGCTGTCCCGGAGCTGCCCTGAG
CCTCCACCCCTTACCTTAGCCAGAGCTGTCCCTGAACCTCCACCTGCTGTCCCGGAGCTGCCCTGAG
CCTCCACCCCTGCTGTCCCGAGCTGCCCTGAGGCTCCACCTGCTGTCCCGGAGCTGCCCTGAGGC
TCCACCCCTGCTGTCCCGGAGCTGCCCTGAGGCTCCACCTGCTGTCCCTGGAGTTGCCCTGAGCCTC
CACCTGCTGTCCCGAAACTGTCAATTGAGCATCCACCGGGCTGTCCCTAGAGCCAGAAAGCTTAGGGC
TGGCAAACCTCACCCCTCACTCCTCTGGCCCTTCCAGCCATCAGCACTTTGGAACAGCCATG
AAGCCCTTTAATCTCTAGAAAGGTGCCCTCAGGAAGGCACAGAGAGGTACACACAGGTGGTATGGTGC
CTTACCTGTGTACACTGGGCCAGGCTGGCCCTTAAAGGGATGATGGTGGAAACAGCTGAGACCACCC
CTCTTCTCAGAGAGGCAGGGATAAAGAAAAGGGACAGCGGAAGGAAGAACTGTGGGAGGATGCTGAG
GGTAGAGGCTGCTCCACAGGCCTGAGCGGAGGGATCCTTTCTCCCTGAAGCTGCCCTGAGCCTCCA
TCCAGACAGGACCTCCAACCACTGCTAGGGGCCATTCTCAGGAGGTTCACTCCAGACAAAGGGCTCCGG
AGACTTAAGTCCACTTTCCAAAATGGACATGATTCATCTGGCAAGTCAAGGACCCAGGACATCCAGGAC
CAAGCCTTGTAAAGGACGGAAGGGAGTAACATGCCCTGGTGGACGTGGGGCCCAAGTCTCTCTGTGGCAG
ATAGGCAAGGGGCTGTGCCCTCACTCCGTACACCTGCTGGCAGCACATGCTCTGAGCTCAGGTCCCAG
TGCCCTTCTGCTCCTCCTCAGGCCACAGATTTCTAAGGACCCGCTGGGCCACAGGTGAACAGTGCAGT
GGCCTCCTGGACGCAGGGCTCCAAGGCTCCTCATCGGGTGCAGCACAGTCTGGCTCAGCAGCACCA
GCCCTGTGGGCCCCATCCGAGTCTGGAACAGAGTGCAGGAGGAGGGAAGGGCTCAGGTAGCATGTTCA
AACTCATCTGGGAGCAGCTGCCTCCTTTCTGGCAAGCCAGCCTCCCTCCTCCTCAGGAAGCCCTCCT
GAGGTCCACACCCAGAGGCCAAGCAGGCTGTGTCTTGGACATTCCTGCTTTCTCTCATGTGCAAGCC

Homo sapiens chromosome 18, GRCh38 reference primary assembly





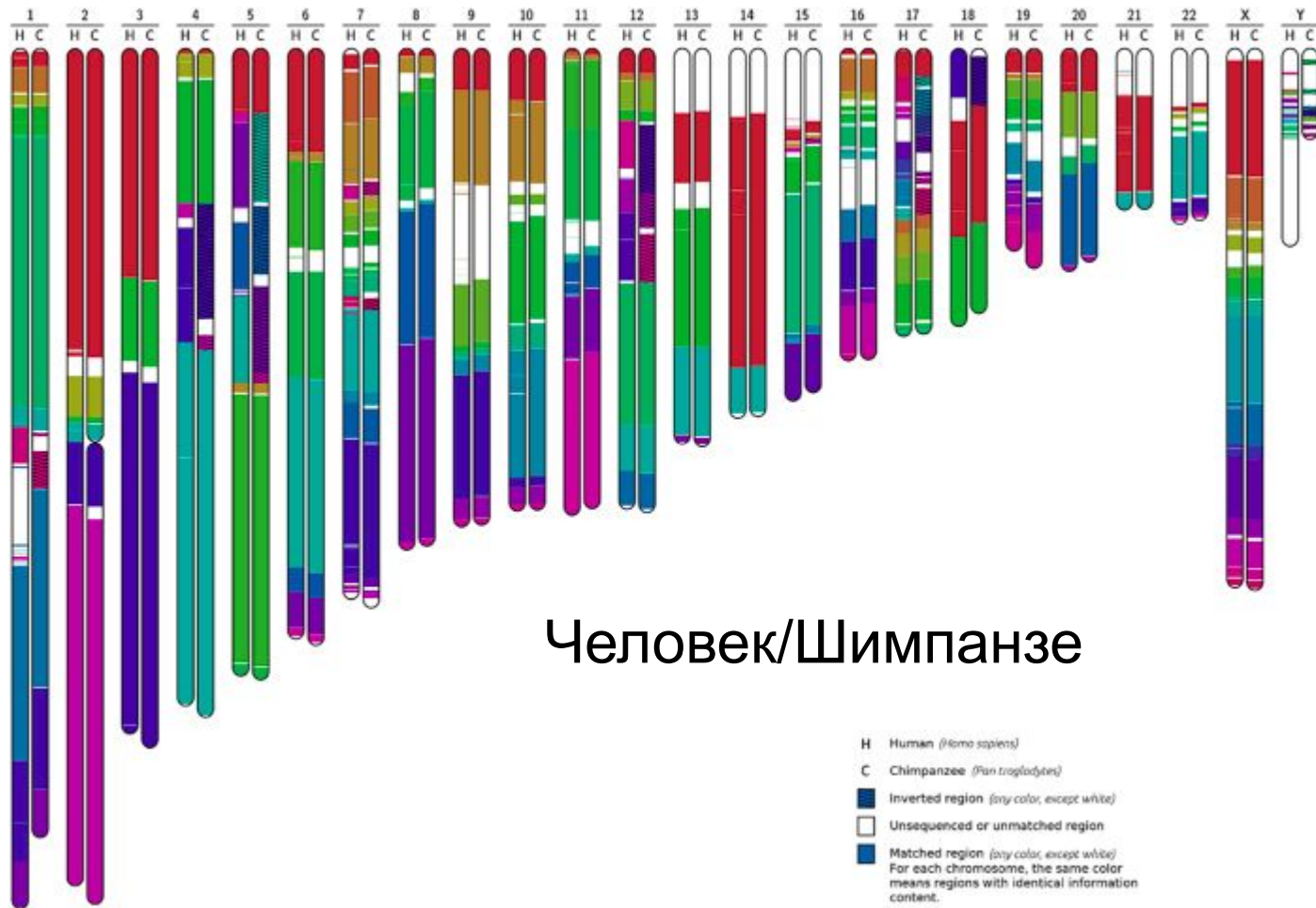
Результаты The Human Genome Project

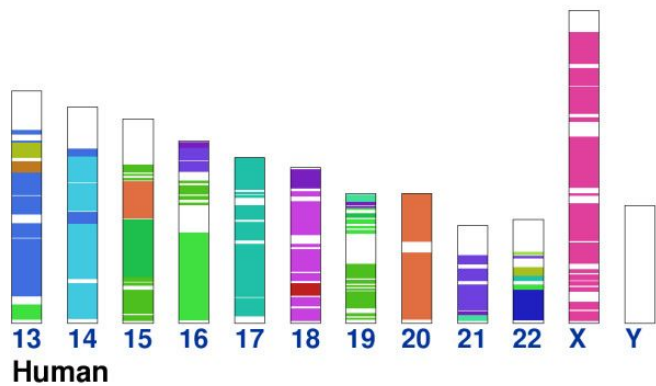
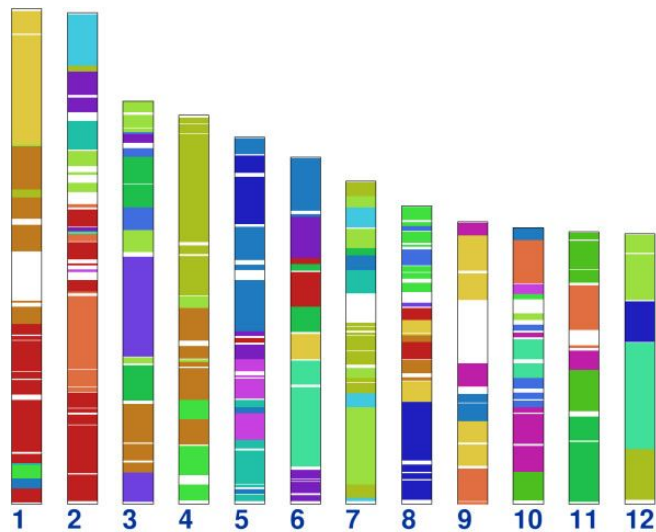
- Получили гаплоидный геном
- Длина генома : 3.2 млрд пар нуклеотидов с ошибкой 1/100 000 нуклеотидов
- Около 341 “пробела” в геноме. Это сложные для секвенирования места.
- Черновая версия: генов около 40 000 штук,
чистовая версия: генов около 20 000 -- 25 000 штук
- Последовательность геномов почти идентична у двух любых людей (99.9%)
- Стоимость проекта: 3 млрд долларов

Что делать теперь?

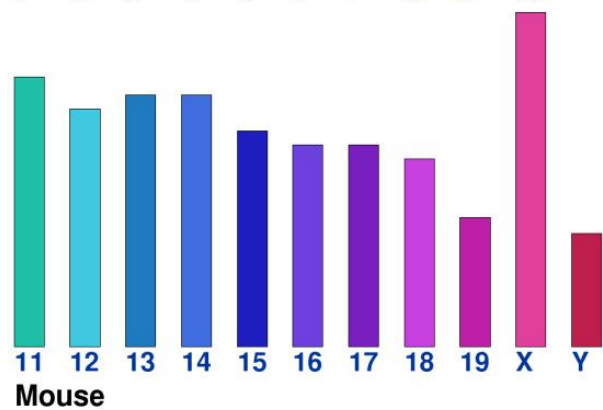
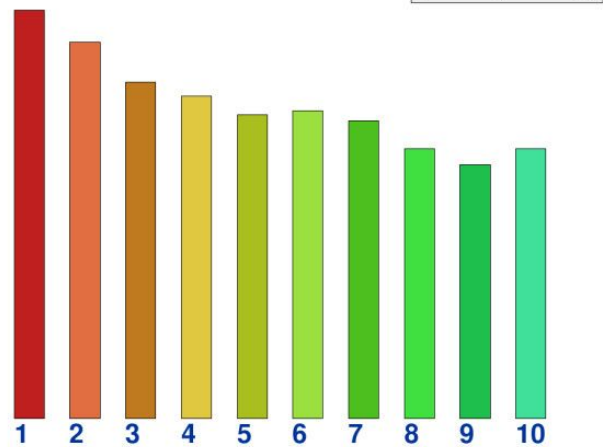
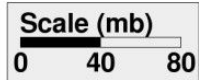
Где применять эти результаты?

1. Эволюция человека
2. Медицина — мутации, приводящие к развитию заболеваний
3. Дизайн лекарств
4. Криминалистика
5. etc.





Human

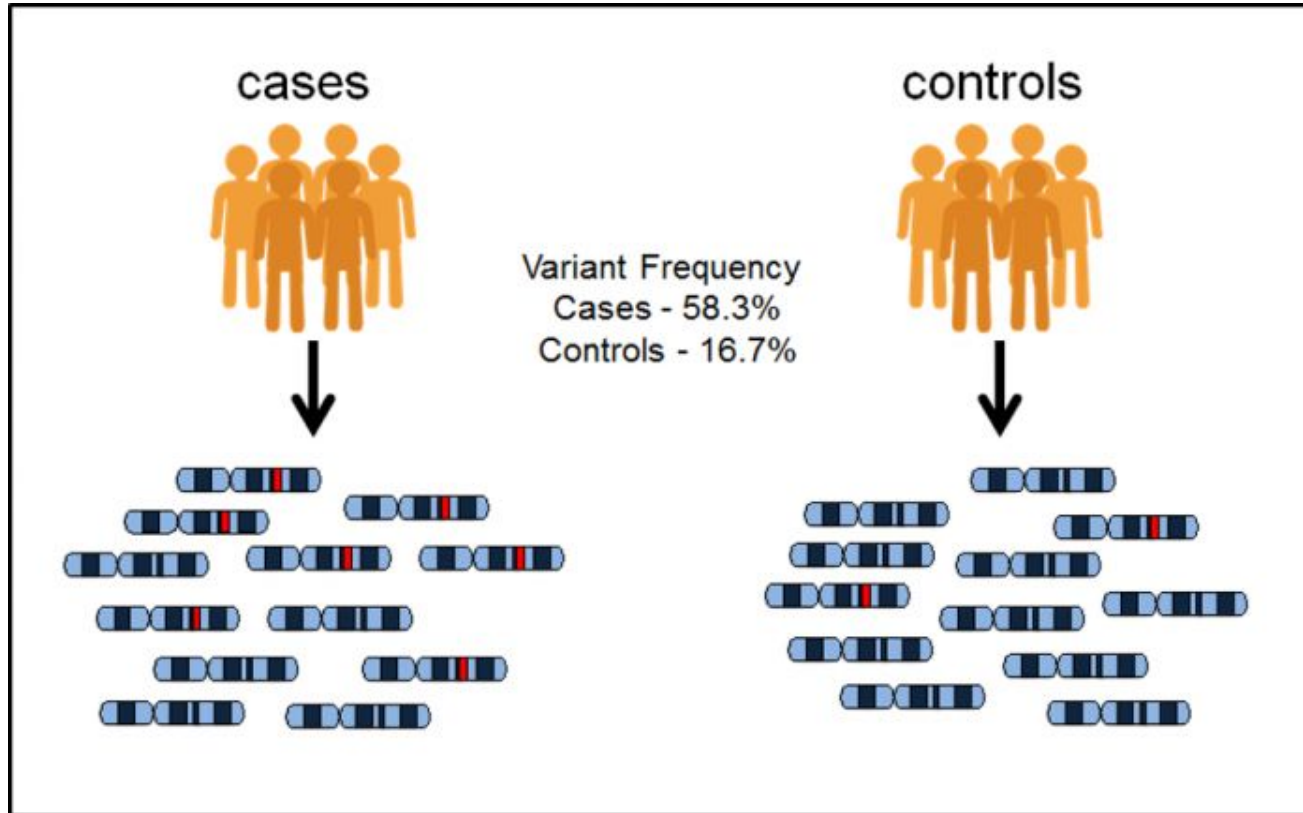


Mouse

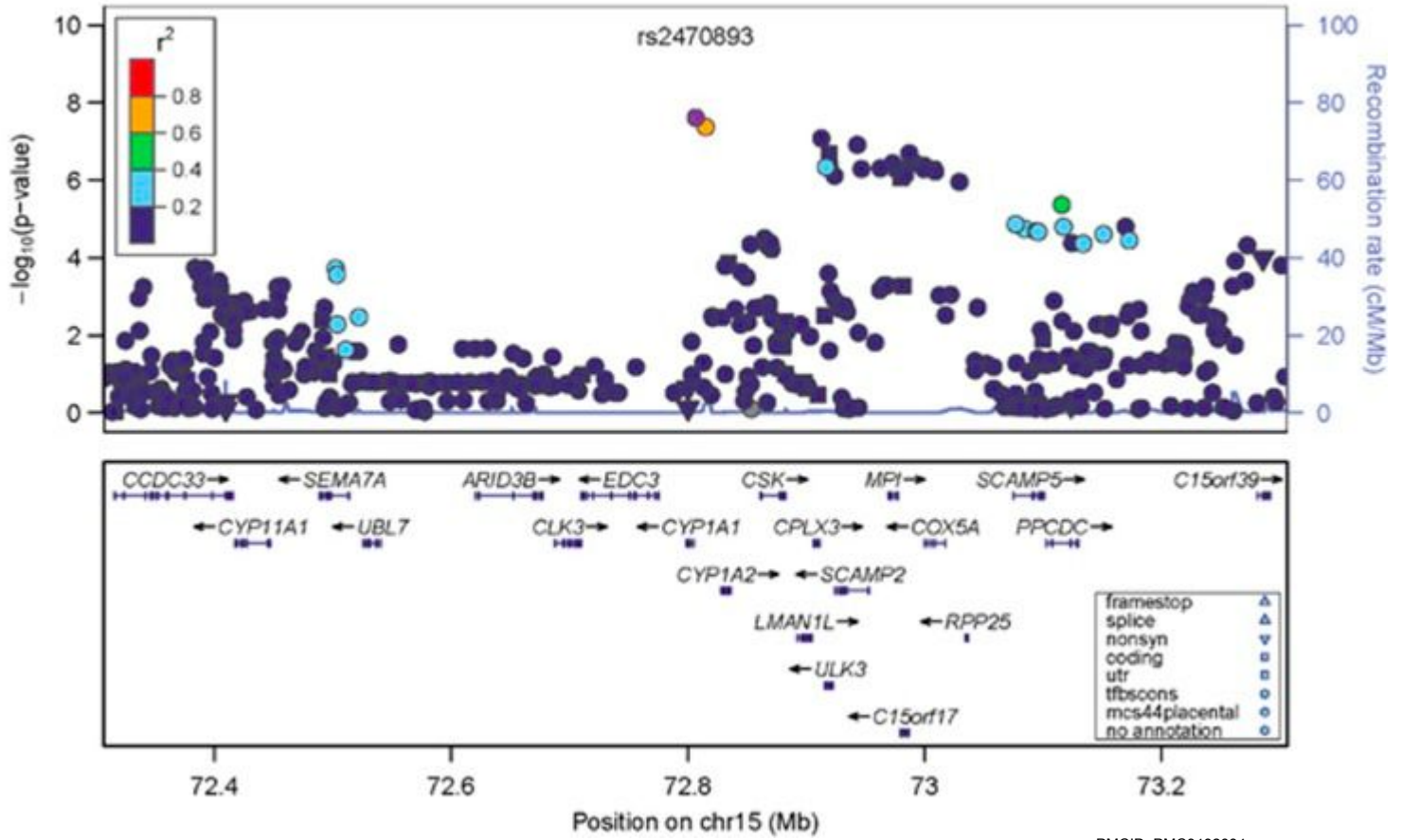


96 Mya

GWAS — полногеномные ассоциированные исследования



Plotted SNPs

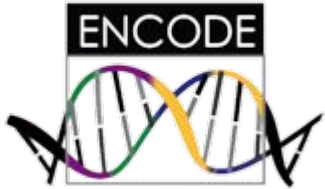


PMCID: PMC3482684

Проекты после HGP

HapMap — паттерны нуклеотидного разнообразия у человека (в разных популяциях)

ENCODE — функциональные элементы в человеческом геноме (роль “не_белок-кодирующей” ДНК, которая занимает 98% генома)



Популяционные выборки


1000 Genomes

IGSR: The International Genome Sample Resource
Supporting open human variation data

Home About Data Portal Analysis Contact Browser FAQ

Search IGSR

IGSR and the 1000 Genomes Project



Populations: ● - African; ● - American; ● - East Asian; ● - European; ● - South Asian;

The International Genome Sample Resource (IGSR) was established to ensure the ongoing usability of data generated by the 1000

Links

- [Announcements](#)
- [IGSR Sample Collection Principles](#)
- [1000 Genomes Project Publications](#)
- [File formats](#)
- [Software tools](#)
- [Download data](#)
- [Twitter](#)
- [IGSR NAR publication](#)

Популяционные выборки

1000 Genomes

UK Biobank

The screenshot shows the UK Biobank website homepage. At the top, the logo 'biobank UK' is displayed with the tagline 'Improving the health of future generations'. To the right, contact information is provided: 'Call us on: 0800 0276 276' and 'Mon-Fri 8am-6pm (Sat 8am-4pm)'. A search bar is located below the contact info. The main navigation menu includes: 'About | Key documents | Participants | Researchers | Data Showcase | Researcher login | Approved studies | Publications | Careers'. The central text block states: 'UK Biobank is a national and international health resource with unparalleled research opportunities, open to all bona fide health researchers. UK Biobank aims to improve the prevention, diagnosis and treatment of a wide range of serious and life-threatening illnesses – including cancer, heart diseases, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and forms of dementia. It is following the health and well-being of 500,000 volunteer participants and provides health information, which does not identify them, to approved researchers in the UK and overseas, from academia and industry. Scientists, please ensure you read the [background materials](#) before registering. To our participants, we say thank you for supporting this important resource to improve health. Without you, none of the research featured on this website would be possible.' Below this is a button 'Read more about Biobank UK'. The lower section features a grid of six featured articles: 'UK Biobank leads the way in genetics research to tackle chronic diseases', 'How is UK Biobank using your diet questionnaires?', 'Save the date: UK Biobank Scientific Conference 2020', 'New data available: exome sequence data on 50,000 participants', 'Two studies, one goal – to improve health', and 'How is UK Biobank tackling Dementia'. Below the grid are two sections: 'Latest News' with an article 'Sepsis and cholesterol – the story unfolds' dated February 24th, 2020, and 'Tweets by @uk_biobank' showing a tweet from UK Biobank about Mrs Rita Clements attending the Bristol imaging centre.

Знание генома человека можно применять для изучения

1. Эволюции человека
2. Медицины — мутации, приводящие к развитию заболеваний; (ранняя) диагностика
3. Дизайна лекарств (знаем, где мутация, как меняется из-за этого белок)
4. Криминалистики
5. etc.

План нашей лекции

1. Что такое геном и где он находится?
2. Что было известно про геном человека до проекта “Геном Человека”?
3. Почему начали этот проект?
4. Как проект проходил и что могло (гипотетически) его ускорить?
5. Какие результаты проекта?
6. Что теперь?



Задание

Узнайте сколько стоит секвенирование вашего персонального генома в Москве в одном из вариантов. Приемлемым покрытием считать в среднем 20 прочтений, покрывающих каждый нуклеотид. С биоинформатической работой или без - как фирма предлагает.

- экзом (все последовательности, кодирующие белки)
- транскриптом (последовательности всех мРНК клеток ткани с указанием процента каждой мРНК)
- полный геном

В ответе укажите: фирма, вариант секвенирования, покрытие, цена, как долго ждать.