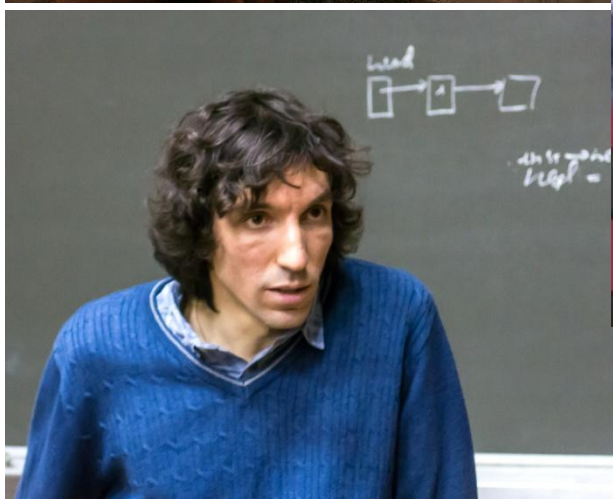
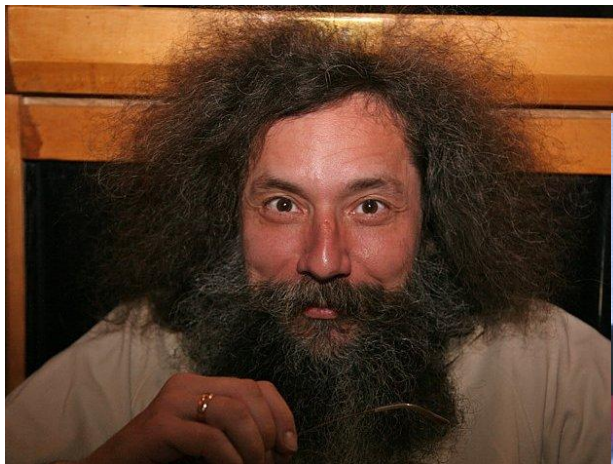


План занятия

- Представление преподадов
- Otzi
 - Про otzi
 - геном
 - секвенирование
- Что можно прочесть в геноме (и метагеноме) Otzi
 - Родство
 - Болезни
- **Мини-контрольная: компл цепочка; ген 3-пептида**
- Как получить зачет. Рейтинг
- Доменная система.

1 . ПОЗНАКОМЬТЕСЬ С НАМИ

Наши люди – важные и серьёзные



Задание по математике: установите взаимно-однозначное соответствие



Гельфанд
Миронов
Кондраше
в
Базыкин
Кунин
Фаворов
Первушин
Головин

Преподаватели в 1м семестре

Дарья
Владимировна
ДИБРОВА



Анастасия
ЖАРИКОВА



Артур
Олегович
ЗАЛЕВСКИЙ



Андрей
Владимирович
АЛЕКСЕЕВСКИЙ



Сергей
Александрович
СПИРИН



Иван
Сергеевич
РУСИНОВ



Анна
Степановна
ЕРШОВА



Администраторы комп. класса

Лидия
Куртовна
СМИРЕНИНА



Александр
Владимирович
БАРАНИН



Студенты старших курсов, участвующие в преподавании на младших курсах (не все)



Задание по жизни: узнайте как их зовут

Геном Ötzi

Чем занимаются
биоинформатики

1. Otzi



Бодрый старичок (45 лет), погибший 5 300 лет тому назад

Seidler et al. (1992). Some anthropological aspects of the prehistoric Tyrolean ice man. Science, 258(5081), 455–457

Эци был случайно обнаружен [19 сентября 1991 года](#) двумя немецкими туристами из [Нюрнберга](#), *Гельмутом и Эрикой Симонами*.

Мумия хорошо сохранилась благодаря тому, что была заморожена в лёд. При извлечении без археологических инструментов (с помощью [отбойного молотка](#) и [ледорубов](#)) было повреждено бедро тела; кроме того, присутствующие брали части его одежды в качестве сувениров.

В дальнейшем тело было передано в морг города [Инсбрук](#), где был установлен его истинный возраст.

On the 19th of September 1991, the naturally mummified body of the Iceman, also referred to as “Ötzi”, was discovered in the Ötztal Alps at an altitude of 3210 m.

The glacier mummy lived around 3300 BCE and was approximately 40–50 years old when he died in the high alpine area (Gaber and Künzle, 1998). The Iceman is now curated, along with his well-preserved clothing and equipment, at the Archaeological Museum in Bolzano.

Since his discovery in 1991, the body has been the focus of intensive study (e.g. Barfield et al., 1992; zur Nedden and Wicke, 1992; Seidler et al., 1992; Spindler, 2000). A variety of methods, including radiology and computer tomography (Murphy et al., 2003), histology (Nerlich et al., 2003), isotope analysis (Müller et al. 2003), botany (Oeggli et al., 2007), and genetic analyses (Ermini et al., 2008), have revealed important aspects of his life, health,



The Iceman (reconstruction by Adrie and Alfons Kennis). Credit: Reconstruction by Kennis © South Tyrol Museum of Archaeology, Foto Ochsenreiter

Read more

at: <https://phys.org/news/2016-01-scientists-helicobacter-pylori-contents-otzi.html#jCp>

Политика и наука

Закон есть закон.

Фердинанд роняет фразу о том, что он родился в этом трактире на кухне

Джузеппе провозглашает, что Фердинанд не имеет права задерживать его. Фердинанд - не француз: он родился в Италии, т.к. граница проходит через трактир и кухня – в Италии!

Во французской полиции. «Я родился в Италии от матери-итальянки и неизвестного отца. Работаю во французской полиции. Правда, что я француз?»

Джузеппе - итальянский контрабандист



Фердинанд – французский полицейский - таможенник



Хэппи энд:
Хозяин трактира: «Я перенес границу потому, что итальянцам для посещения сортира приходилось нарушать границу.

Граница между Италией и Австрией была установлена в 1919 году по [Сен-Жерменскому договору](#) между долинами Иннталь и Эчталь. В области долины Тизенйох из-за снежного покрова ледника границу было нелегко определить, поэтому

власти назначили на 2 октября 1991 года измерение прохождения границы, в результате чего было установлено, что тело было обнаружено в [Южном Тироле](#) на территории Италии, в **92,56 м** от границы с Австрией

Найден немецкими туристами.

На леднике неустановленной принадлежности.

Доставлен в Инсбрук (Австрия).

Изучен австрийскими учеными

Мумия принадлежит Австрии или Италии?

2. Геном

Гено́м — совокупность **наследуемой информации**, заключенной в клетке организма

Информация – это что-то нематериальное!

Должен быть носитель информации.

Возможно перекодирование с одного носителя на другой.

На каких носителях закодирована

наследуемая информация? Выслушайте смельчак информации в ответе Армянского радио:

Правда ли, что [шахматист] [Петросян](#) выиграл в лотерею машину?

— Правда, только не Петросян, а [футболист] [Иштоян](#) и не машину, а швейную машинку, и не в лотерею, а в карты, и не выиграл, а проиграл.» 13

Полный геном человека

aba@mouse.beTozersky.msu.ru:/data/genomes/human_Oct_2008

n	Name	Size	Date
..		Up	09/26/17
	shuffled	Folder	11/05/08
	chr1	fa 219 М	10/29/08
	chr10	fa 127 М	10/29/08
	chr11	fa 127 М	10/29/08
	chr12	fa 126 М	10/29/08
	chr13	fa 93 М	10/29/08
	chr14	fa 85 М	10/29/08
	chr15	fa 79 М	10/29/08
	chr16	fa 76 М	10/29/08
	chr17	fa 77 М	10/29/08
	chr18	fa 72 М	10/29/08
	chr19	fa 54 М	10/29/08
	chr2	fa 230 М	10/29/08
	chr20	fa 58 М	10/29/08
	chr21	fa 34 М	10/29/08
	chr22	fa 34 М	10/29/08
	chr3	fa 189 М	10/29/08
	chr4	fa 182 М	10/29/08
	chr5	fa 172 М	10/29/08
	chr6	fa 164 М	10/29/08
	chr7	fa 150 М	10/29/08
	chr8	fa 139 М	10/29/08
	chr9	fa 117 М	10/29/08
	chrMT	fa 16882	10/29/08
	chrX	fa 148 М	10/29/08
	chrY	fa 25 М	10/29/08

Информатика:

Что больше:
16882 или
148М?

Внимание к деталям:

троллейбусы
4, 4М, 4К

На будущее, к
вопросу о

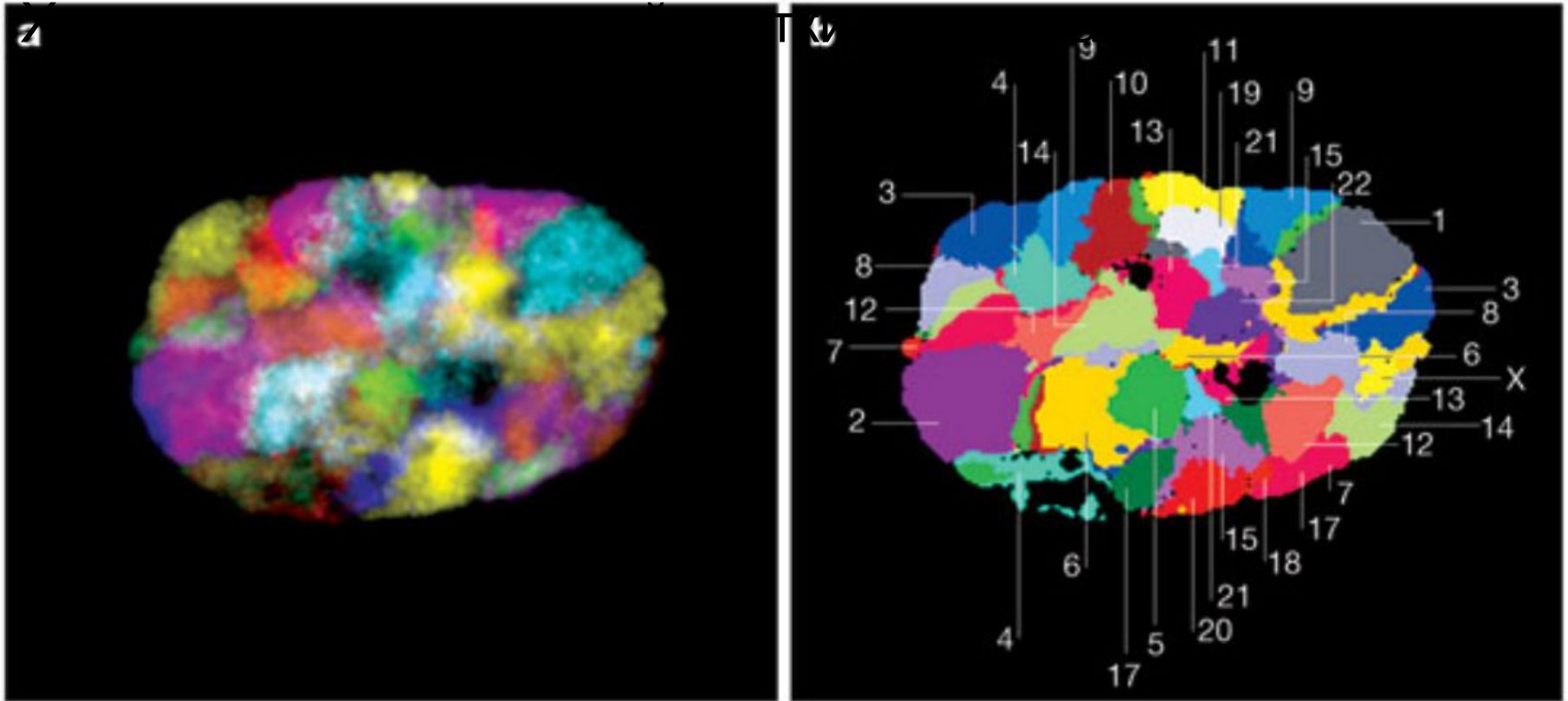
биоинформатик

e: fa – сокращение
от **fasta**;
fasta – формат
файла с
последователь-
ностью
нуклеиновой
кислоты или
последователь-
ностью белка.


```
>gi|17981852|ref|NC_001807.4| Homo sapiens mitochondrion, complete genome
GATCACAGGTCTATCACCTATTAACCACTCACGGGAGCTCTCCATGCATTTGGTATTTTCGTCTGGGG
GTGTGCACGCGATAGCATTGCGAGACGCTGGAGCCGGAGCACCTATGTCGCAGTATCTGTCTTTGATTC
CTGCCTCATTCTATTATTTATCGCACCTACGTTCAATATTACAGGCGAACATACTACTAAAGTGTGTTA
ATTAATTAATGCTTGTAGGACATAATAATAACAATTGAATGTCTGCACAGCCGCTTTCCACACAGACATC
ATAACAAAAAATTTCCACCAAACCCCCCTCCCCCGCTTCTGGCCACAGCACTTAAACACATCTCTGC
CAAACCCCAAAAACAAGAACCCTAACACCAGCCTAACAGATTTCAAATTTTATCTTTAGGCGGTATGC
ACTTTTAAACAGTCACCCCCCACTAACACATTATTTTCCCCTCCCCTACTCCATACTACTAATCTCATCAA
TACAACCCCGCCCATCTACCCAGCACACACACACCCGCTGCTAACCCCATACCCCGAACCAACCAACC
CCAAAGACACCCCCACAGTTTATGTAGCTTACCTCCTCAAAGCAATACACTGAAAATGTTTAGACGGGC
TCACATCACCCATAAACAATAGGTTGGTCTAGCCTTCTATTAGCTCTTAGTAAGATTACACATGC
AAGCATCCCCGTTCCAGTGAGTTCACCCTCTAAATCACCACGATCAAAGGGACAAGCATCAAGCACGCA
GCAATGCAGCTCAAACGCTTAGCCTAGCCACACCCCCACGGGAAACAGCAGTGATTAACCTTTAGCAAT
AAACGAAAGTTTAACTAAGCTATACTAACCCAGGGTTGGTCAATTTCTGTGCCAGCCACCGCGGTACAC
GATTAACCCAAGTCAATAGAAGCCGGCGTAAAGAGTGTTTTAGATCACCCCTCCCAATAAAGCTAAAA
CTCACCTGAGTTGTAAAAACTCCAGTTGACACAAAATAGACTACGAAAGTGGCTTTAACATATCTGAAC
ACACAATAGCTAAGACCCAACTGGGATTAGATACCCCACTATGCTTAGCCCTAAACCTCAACAGTTAAA
TCAACAAAATGCTCGCCAGAACACTACGAGCCACAGCTTAAAATCAAAGGACCTGGCGGTGCTTCATA
TCCCTCTAGAGGAGCCTGTTCTGTAATCGATAAACCCCGATCAACCTCACCACTCTTGCTCAGCCTATA
TACCGCCATCTTCAGCAAACCCTGATGAAGGCTACAAAGTAAGCGCAAGTACCCACGTAAAGACGTTAGG
TCAAGGTGTAGCCCATGAGGTGGCAAGAAATGGGCTACATTTTCTACCCCAAGAAAATACGATAGCCCTT
ATGAAACTTAAAGGTGCAAGGTGGATTTAGCAGTAACTGAGAGTAGAGTGTAGTTGAAACAGGGCCCT
GAAGCGGTACACACCCGCGTCAACCTCCTCAAGTATACTTCAAAGGACATTTAACTAAAACCCCTACG
CATTTATATAGAGGAGACAAGTCTGTAACATGGTAAGTGTACTGGAAAGTGCATTTGGACGAACCAGAGTG
TAGCTTAAACACAAAGCACCCAACTTACACTTAGGAGATTTCAACTTAACTTGACCGCTCTGAGCTAAACC
TAGCCCCAAACCCACTCCACCTTACTACCAGACAACCTTAGCCAAACCAATTTACCCAAATAAAGTATAGG
CGATAGAAATTTGAAACCTGGCGCAATAGATATAGTACCGCAAGGGAAAGATGAAAAATTATAACCAAGCA
TAATATAGCAAGGACTAACCCCTATACCTTCTGCATAATGAAATTAAGTAAATAAATTTGCAAGGAGAG
CCAAAGCTAAGACCCCGAAACCAGACGAGCTACCTAAGAACAGCTAAAAGAGCACACCCCGTCTATGTAG
CAAATAGTGGGAAGATTTATAGGTAGAGGCGACAAACCTACCGAGCCTGGTGATAGCTGGTTGTCCAAG
ATAGAATCTTAGTTCAACTTTAAATTTGCCACAGAACCCTCTAAATCCCCTTGTAATTTAACTGTTAG
TCCAAAGAGGAACAGCTCTTTGGACACTAGGAAAAACCTTGTAGAGAGAGTAAAAATTTAACACCCAT
AGTAGGCCTAAAAGCAGCCACCAATTAAGAAAGCGTTCAAGCTCAACCCACTACCTAAAAAATCCCAA
ACATATAACTGAACCTCTCACACCCAATTTGGACCAATCTATCACCTATAGAAGAACTAATGTTAGTATA
AGTAACATGAAAACATTCTCCTCCGATAAGCCTGCGTCAGATCAAACACTGAACTGACAATTAACAGC
CCAATATCTACAATCAACCAACAAGTCATTATTACCCTCACTGTCAACCCAACACAGGCATGCTCATAAG
GAAAGGTTAAAAAAGTAAAAGGAACTCGGCAACCTTACCCCGCTGTTTACCAAAACATCACCTCTA
GCATCACCAAGTATTAGAGGCACCGCCTGCCAGTGACACATGTTTAAACGGCCGCGGTACCCTAACCGTGC
AAAGGTAGCATAATCACTTGTTCCCTTAAATAGGGACCTGTATGAATGGCTCCACGAGGGTTTCAGCTGTCT
CTTACTTTTAAACAGTGAAATTTGACCTGCCCGTGAAGAGGCGGGCATGACACAGCAAGACGAGAAGACCC
TATGGAGCTTTAATTTATTAATGCAAACAGTACCTAACAAACCCACAGGTCCTAACTACCAAACCTGCA
TTAAAAATTTGGCTTGGGGCGACCTCGGAGCAGAACCCTCGGAGCAGTACATGCTAAGACTTCACC
```

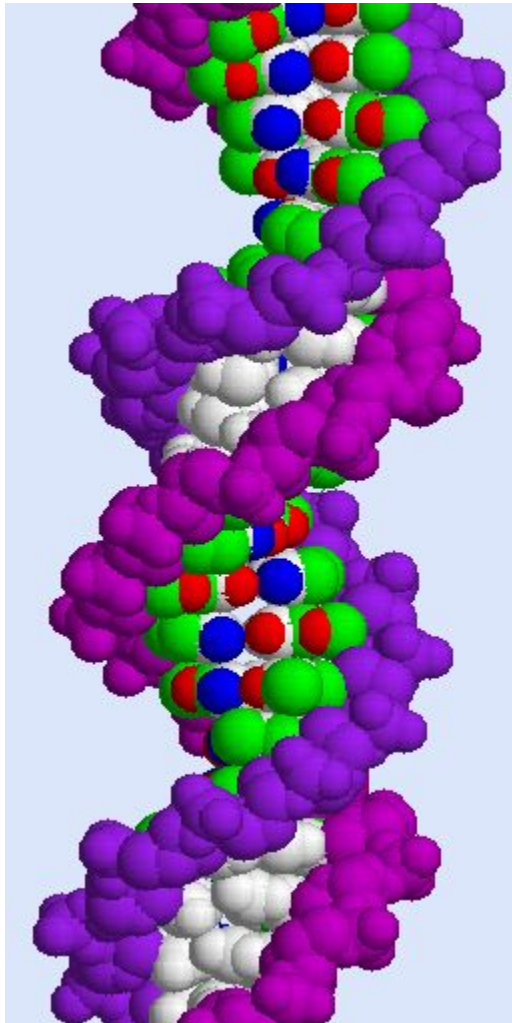
А вот что
внутри
файла
chrMT.fa
(фрагмент)

Полный геном человека. Другой носитель



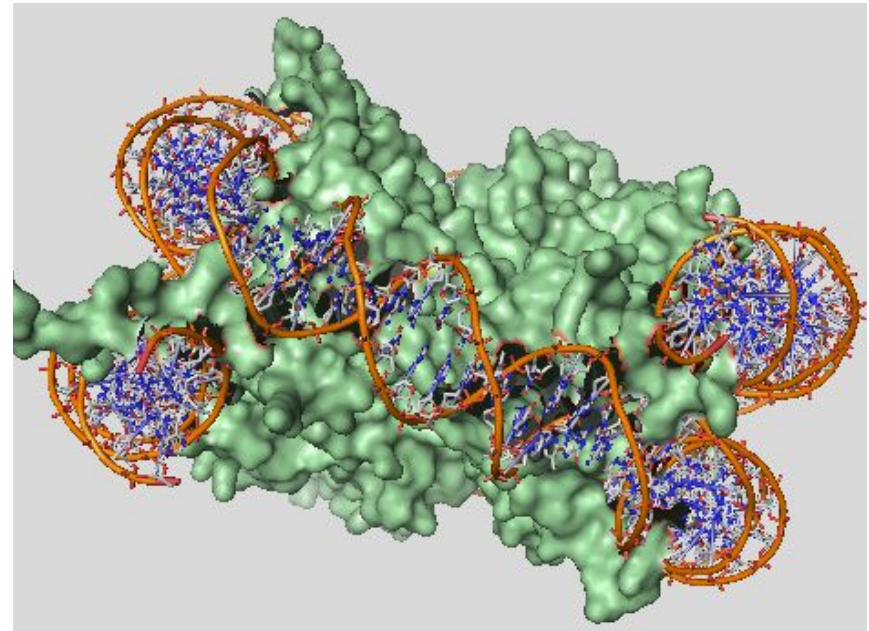
Bolzer A et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 2005 May;3(5):e157

Перекодировка. ДНК (фрагмент)



Двойная
спираль
ДНК.

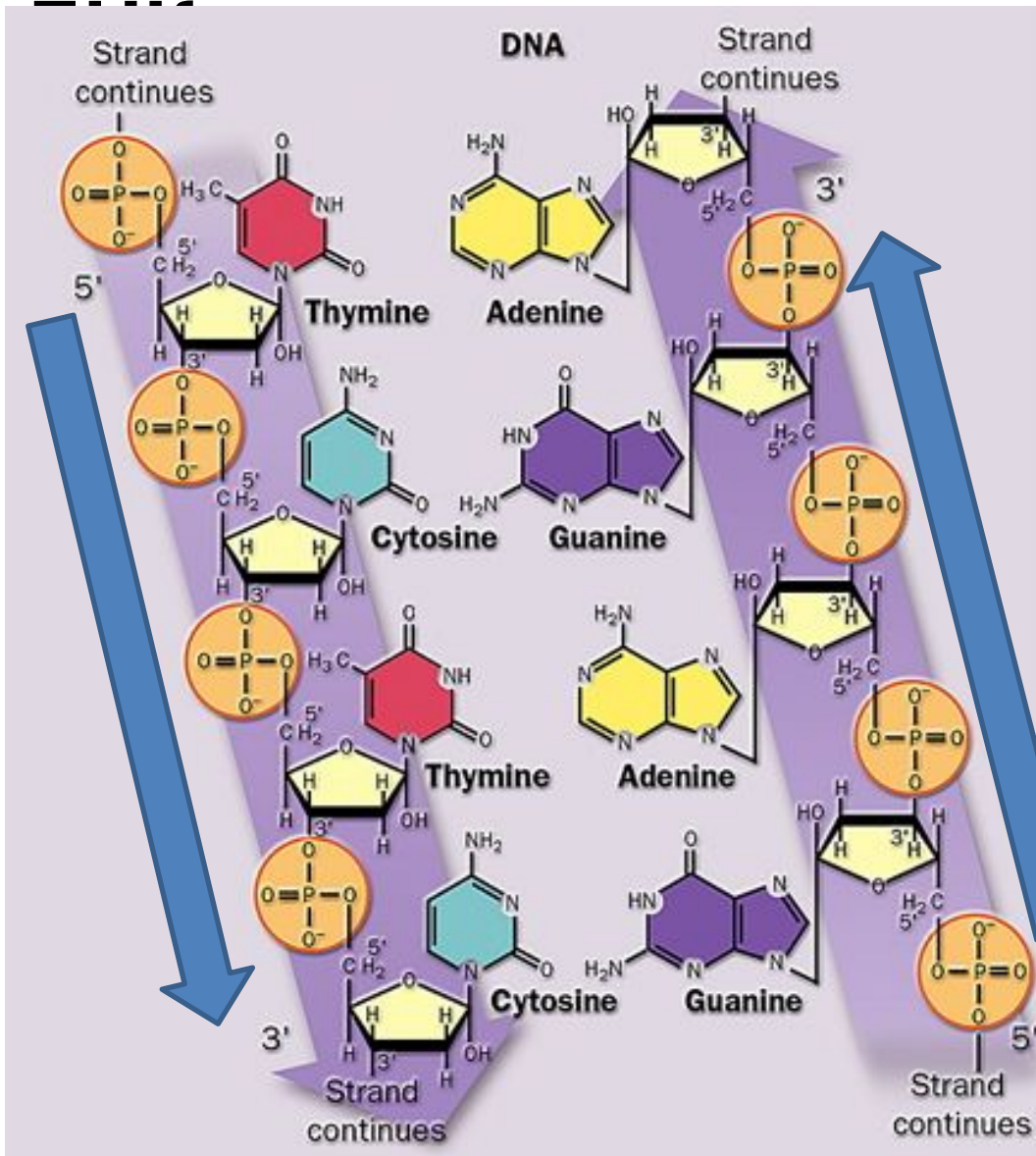
Раскраска
моя ААл 😊



ДНК человека на
"катушке" из гистонов:
вид сбоку
(гистоны – такие белки)

Обе структуры расшифрованы с
помощью рентгеноструктурного
анализа

Химическая формула



Последовательно
сть этой ДНК:

TCTC или **GAGA**?

Обратите внимание
на стрелки:

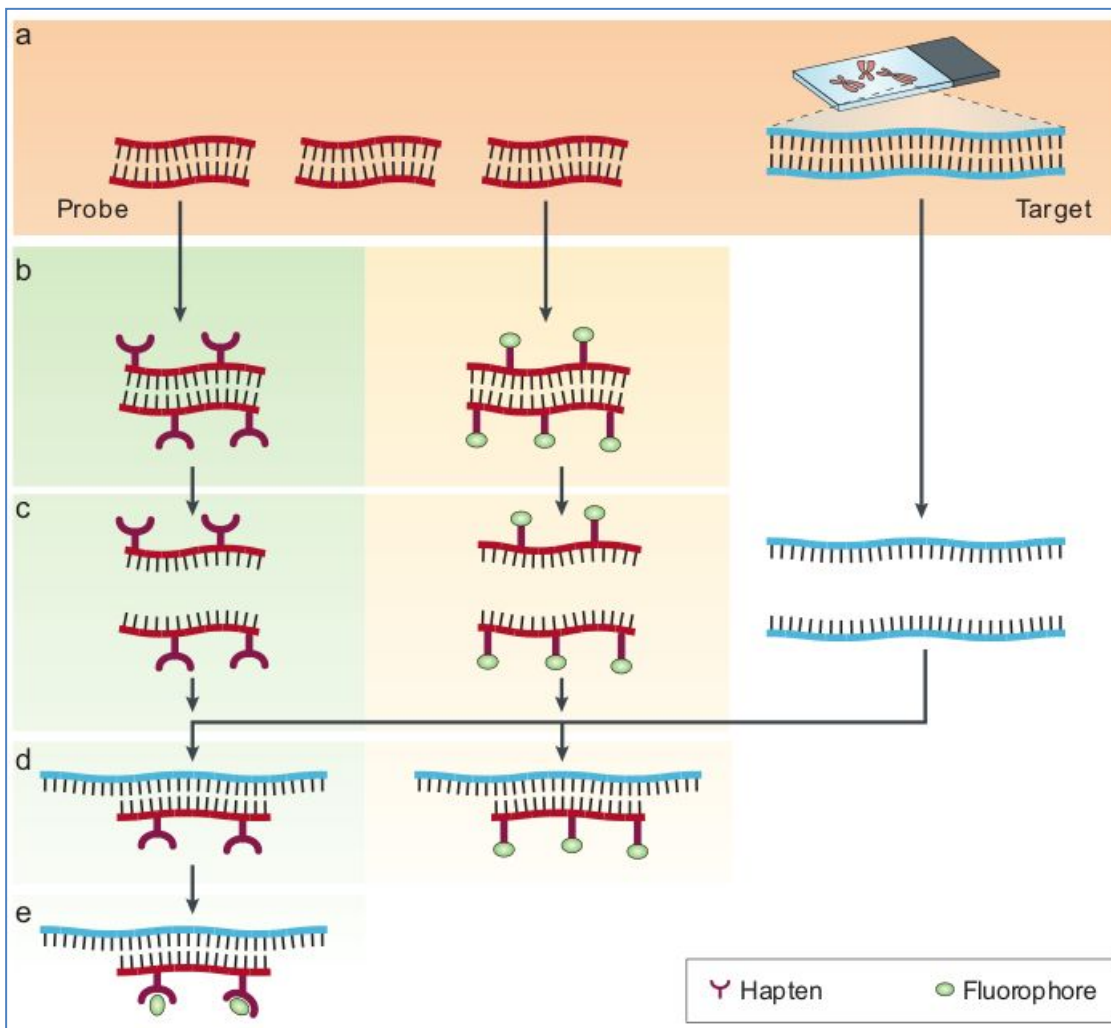
каждая цепочка ДНК
имеет ориентацию
от 5'- к 3'-концу.

Цепочки
ориентированы
противоположно.

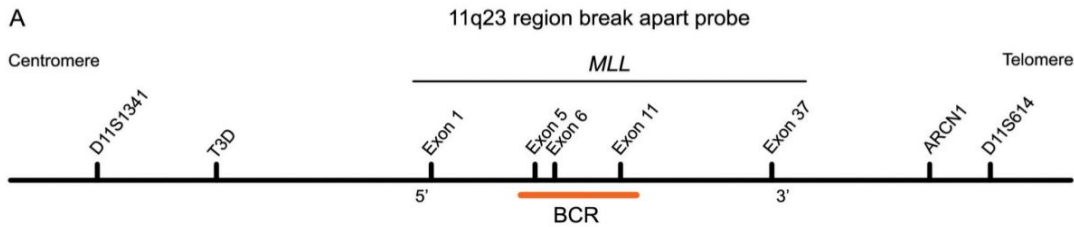
Информация
считывается от 5' к 3'

Как сфотографировали хромосомы.

Гибридизация ДНК
с меченой пробой
(например,
флюорофором)

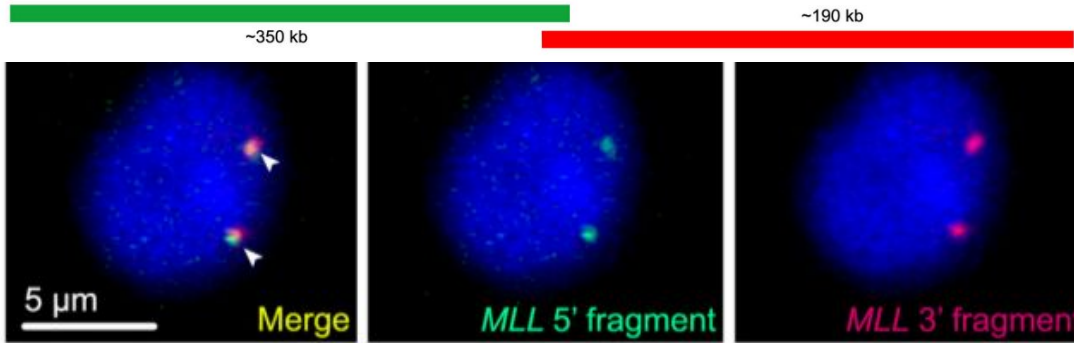


Speicher MR, Carter NP. The new cytogenetics: blurring the boundaries with molecular biology. Nat Rev Genet. 2005

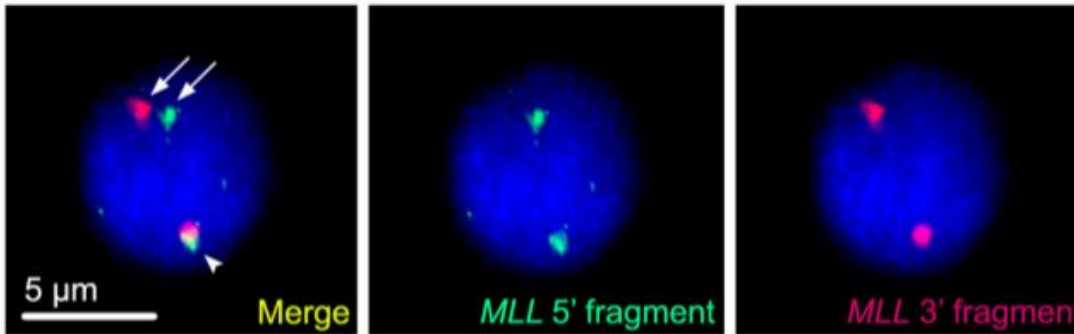


FISH (fluorescence in situ hybridisation)

B



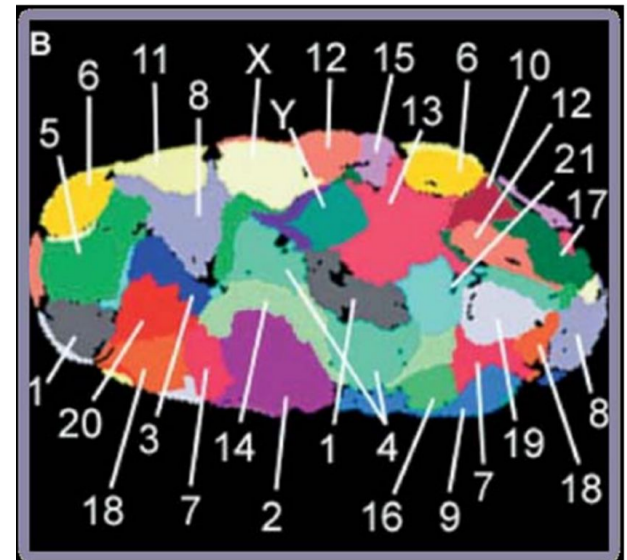
C



Glukhov et al. The Broken MLL Gene Is Frequently Located Outside the Inherent Chromosome Territory in Human Lymphoid Cells Treated with DNA Topoisomerase II Poison Etoposide. PLoS One. 2013 8(9):e75871

Chromosome painting.

Boltzer et al., PLOS Biol. 2005



Предложите сами как покрасить хромосомы целиком в разные цвета

Информатика: RGB – red green blue

Состав генома человека. Наследование ДНК

Как определить последовательность ДНК – секвенировать ДНК (на примере ДНК Otzi)

- Митохондриальный геном.
Ermini, L. et al. Complete mitochondrial genome sequence of the Tyrolean Iceman. *Curr. Biol.* 18, 1687–1693 (2008).
- Полный ядерный геном.
Keller et al., New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *Nat Commun.*, 2012, 28;3:698.

Этапы секвенирования

- Взяли образец из кости Otzi
- Выделили из него всю ДНК. (Важно: не есть в это время бутерброды! Почему?)
- Подготовили библиотеку для секвенирования
 - Порубили ДНК на фрагменты размером примерно [в статье не указано]
 - Присоединили к обеим концам (всех) фрагментов короткие одноцепочечные ДНК, называемые адаптерами, со специально разработанной последовательностью

Продолжение

- Полученный раствор ДНК загрузили в робот, называемый секвенатором, подготовленный к работе. В случае Otzi робот SOLiD™ 4
- Нажали большую красную (?) кнопку “Start”
- Подождали несколько часов или недель – зависит от секвенатора – и получили результат
- Без робота секвенирование генома человека заняло больше 10 лет (проект Human genome)
- В каком виде результат?

- We obtained a total of 2.93×10^9 sequencing reads out of which 1.11×10^9 reads (37.9%) mapped uniquely to the human reference genome (hg18). Paired-end reads covered 96% of the reference genome with small variation among chromosomes (max = 97.74% for chr22, min = 89.61% for chrY)
- DNA extraction from the bone sample whose DNA was subjected to next-generation sequencing was performed at the Division of Molecular Genetics, Institute of Human Genetics at the Eberhard-Karls-University of Tübingen, Germany.
- The library was prepared following the SOLiD™ protocol for low-input fragment library preparation. The library was mainly prepared following the manufacturer's instructions. Variations from the protocol are listed in the following.
- A 2 ng μl^{-1} dilution was sequenced using a paired-end sequencing protocol on a SOLiD™ 4 at the Applied Biosystems facility in Beverly, USA.

В каком виде результат?

от 5 до 10

- 25 файлов с последовательностями хромосом и МТ-ДНК?
 - Увы, нет. До этого технология пока не доросла.
- 2.93×10^9 последовательностей – прочтений (также называют ридами) длины 50 с двух концов фрагментов из библиотеки (парные прочтения) (этот секвенатор может секвенировать 25-75 нуклеотидов)
- Размер генома человека: **3.2 млрд пар нуклеотидов**

Упражнение по арифметике. Сколько прочтений, в среднем, содержат один и тот же нуклеотид из ДНК Otzi?

Автомат неплохо соображает ☺

Вот как выглядят два прочтения из огромного выходного файла в формате **fastq**

```
edit sample.fastq - Far 3.0.4242 x86
P:\Education\FBB\Year-18\term1\lecture1\otzi\sample.fastq
@DRR048282.1.1 1 length=255
CGTTTCGTTCTTAGCGTCGTATTGCGTGCTTGGGTGTTAACCTTTGGGCCCCCATTGTTGCT
+DRR048282.1.1 1 length=255
%%./3,,3.,+) '++ +++))*)*..)+)++.)(+)-15**,&((- '*O (**,0/-.-,.12
@DRR048282.2.1 2 length=375
GTTTTCGTCTGCTCTTCTGTGCTGCACCATTCTTTATGCTCTGTTAACCTGCATATTGACGC
+DRR048282.2.1 2 length=375
%%-1-)*)+))+( '(+***)*-*++) )(,.,.&++' (*--')*//0*,.,)+,+.,./..-7+)(
```

Вторая строка – последовательность прочтения

Четвертая строка – качество прочтения каждого нуклеотида; это качество равно

вероятности ошибочного прочтения соответствующей буквы, зашифрованное

Для дальнейшего анализа «одни блокнот программирование на python»: для данного файла в формате fastq расшифровать вероятности в понятном человеку числовом виде и вывести в файл.

Сборка генома

- Из кусочков надо было собрать полные последовательности хромосом.
- Для Otzi это было сделано путем картирования на «референсный» геном человека (hg18).
 - Картирование – «гибридизация в компьютере». Для каждого прочтения находится комплементарный (или совпадающий – он «гибридизуется» с комплементарной цепочкой) участок в референсном геноме.

Результат картирования для Otzi

- Из всех прочтений 1.11×10^9 прочтений (37.9%) картировались однозначно на референсный геном
- Прочтения покрыли 96% референсного генома с небольшой разницей по отдельным хромосомам (max = 97.74% for chr22, min = 89.61% for chrY)
- **Вопрос:** откуда взялись прочтения ДНК, не картирующиеся на референсный геном человека?
- **Вопрос:** почему прочтения покрывают не весь референсный геном?

Что люди узнали про Otzi, прочитав геном

- У него была 0 группа крови
- Глаза карие
- У него была лактозная недостаточность – он не переваривал молоко
- У него был высокий риск коронарной болезни (инфаркта миокарда)
- У него обнаружили ДНК *Borrelia burgdorferi*; предположительно он болел боррелиозом (болезнью Лайма, передаваемой иксодовыми клещами)

Предки и потомки Otzi

- Прямое наследование по материнской линии. По митохондриальной ДНК Otzi выделен в отдельную подгруппу K1o
- Вывод авторов прямых потомков не обнаружено.
- The Iceman's **mtDNA**, therefore, seems to belong to a novel branch of K1, defined by transitions at nucleotide positions 3513 and 8137.
- The transition at position 8137 has been found in samples from Pakistan, India, and the Iberian Peninsula. However, all of these samples belong to the major branch of U7 [22, 27, 28]. Therefore, again, the transition is known to have occurred during the evolution of the human mtDNA, and it is highly improbable that its presence in the Iceman's lineage is due to an artifact. As before, the Iceman's sequence does not show any of the other mutations that define U7, ruling out artificial recombination.
- The mummy's intestinal specimens utilized for the present mtDNA study have been shown to contain animal DNA from red deer (*Cervus elaphus*) and ibex (*Capra ibex*), as well as DNA from different kinds of land plants, such as cereals, pines, and ferns. In fact, phylogenetically reliable sequences of mtDNA for *C. elaphus* and *C. ibex* and for plant nuclear 18S ribosomal RNA (18S rRNA), as well as plastidial ribulose biphosphate carboxylase large subunit (*rbcl*) gene sequences, were previously obtained [8].

Предки и потомки Otzi

- Прямое наследование по материнской линии
- По Y-хромосоме принадлежит группе Y-haplogroup G2a4 has hitherto only been found at appreciable frequencies in Mediterranean islands of the Tyrrhenian Sea (**Sardinia and Corsica**).
- Этот вывод подтвержден анализом последовательностей аутосом (не половых хромосом)

Метагеномы Otzi

The results showed that the Iceman made considered choices of clothing material from both wild and domestic animals, including cattle sheep, goat, brown bear and roe deer (O'Sullivan et al., 2016).

It can be assumed that his clothing was very functional and allowed him to walk up in the mountains even during cold periods.

Zink et al., Possible evidence for care and treatment in the Tyrolean Iceman, Int J Paleopathol. 2018 Aug 8. pii: S1879-9817(18)30088-3.

Предки и потомки Otzi

- Later, a metagenomic approach coupled with targeted genome capture generated 15 350 reads of *H. pylori* DNA throughout the gastrointestinal tract with an abundance decreasing from the stomach towards the lower intestinal tract, and with negative results for the muscle tissue used as control.

THE END

Словарик

- Геном
 - Состав генома человека
 - Гибридизация, зонды
- Секвенирование, NGS
 - Fasta format
 - Прочтения, риды
 - Парные риды
 - Картирование
 - Сборка
 - Математические задачи