

Мутации, гены, выравнивание

План

- Геном: Наследственность, изменчивость, отбор [, обмен информацией между организмами]
- Мутации и их отражение в выравнивании последовательностей :
 - Локальные, большие перестройки
 - Соматические и наследуемые (про мутации de novo)
 - Последствия мутаций: не совместимые с жизнью, слабо вредные, нейтральные, полезные (очень редкие 😊)
- Что написано в геномах: Гены белков, Гены РНК, Участки ДНК участвующие в регуляции, “мусорная ДНК”
- Мутации в белках:
 - Синонимические (silent)
 - Аминокислотные (missense)
 - (nonsense)

План

- Основа живых организмов – информация, содержащаяся на ДНК
 - Наследственность, изменчивость, отбор [, обмен информацией между организмами]
 - При дублировании ДНК (репликации) неизбежны ошибки (мутации)
 - Что происходит
 - Когда происходит
 - Как организм борется с мутациями
 - Ошибка ошибке рознь. Мутации
 - не совместимые с жизнью
 - слабо вредные
 - нейтральные
 - полезные (очень редкие☺)
 - Что написано в геномах?
 - Гены белков
 - Гены РНК
 - Участки ДНК участвующие в регуляции

Информация, заключенная в геноме, закодирована в молекуле (-ах) ДНК

- Судя по тому, что секвенирована ДНК из кости лошади, пролежавшей ок. 700 000 лет (найдена в Канаде), ДНК надежный носитель информации.
- Регулярное переписывание информации как способ сохранения имеет определенные преимущества. Рукопись «Слова о полку Игоревом» сгорела, но информация сохранилась.
- Наша ДНК получилась переписыванием ДНК организма, жившего 3,5 млрд лет тому назад (>>> 700 тыс лет). Этот организм называют **LUCA** (Last Universal Common Ancestor)
- Текст, конечно, очень сильно изменился, но кое-что осталось похожим на первоисточник

Информация, заключенная в геноме, закодирована в молекуле (-ах) ДНК

- При делении клетки (**митоз**) ДНК наследуется ее потомками; ДНК «переписывается» при каждом делении клетки (**репликация**)
- При переписывании большого «текста» неизбежны ошибки. Ошибки могут случиться и при хранении. Ошибки – **мутации** в ДНК – определяют **изменчивость** организма.
- **Отбор** элиминирует вредные и даже слабо вредные мутации из популяции организмов одного вида. Тем самым повышается приспособленность организмов к среде и к изменениям условий среды (**fitness**).

Наследственность, изменчивость, отбор
[, обмен информацией между организмами]

Мутации и их изображение (выравнивание)

- Локальные (точечные) мутации
 - Замена нуклеотида на другой
 - Потеря одного или нескольких нуклеотидов (**делеция**)
 - Вставка (**инсерция**)
- Такие мутации принято изображать выравниванием последовательностей с черточками на месте потерянных (или не вставленных) нуклеотидов. См. след. слайд

Выравнивание гомологичны последовательностей – отражение мутаций, произошедших на пути к ним от их общего предка

- Крупные перестройки ДНК, затрагивающие сотни, тысячи, десятки и сотни тысяч пар нуклеотидов

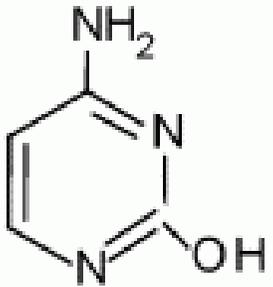
Выравнивание последовательностей геномов *M. capricolum* и *M. mycooides* (маленький фрагмент)

<i>1M. mycooides</i>	1091	t a a - - - t t a a t t a t a a a t t t a t a a a t a t t t t c a t t a a G T C T G A	1130
<i>1M. capricolum</i>	1116	T A A T T T T T A A T T A T A A A T T T A T A A A T A T T T T T C A T T A A G T C T A A	1158
<i>1M. mycooides</i>	1131	T G T A T T C A C C T T T T T T A A T A T A T A A A A C T C C A G A A A G A A A A T C	1173
<i>1M. capricolum</i>	1159	T A T A T T C A C C T T T T T T A A C A T A T A A A A C T C C A G A A A G A A A A T C	1201
<i>1M. mycooides</i>	1174	T T T A A A A C G T T T A G C T T T A T T A T C A T C T A A G T T T T T A A A A T C T	1216
<i>1M. capricolum</i>	1202	T T T A A A A C G T T T A G C T T T A T T A T C A T C T A A G T T T T T A A A A T C T	1244
<i>1M. mycooides</i>	1217	A C A A C A A C A A C T T T T T G A T C T A A T A A A G T A T C T A C A A T T G A T T	1259
<i>1M. capricolum</i>	1245	A T A A C A A C A A C A T T A T G T T C T A A T A A A G T A T C A A C A A T T G A T T	1287
<i>1M. mycooides</i>	1260	G A A C T T C A G A A A A T T T C A T A G G A C T A A A T A C A T A A G T G T T A A T	1302
<i>1M. capricolum</i>	1288	G A A T T T C A G A A A A T T T C A T A G G A C T A A A A A C A T A T G T A T T A A C	1330

В среднем, 92% совпадающих букв на гомологичных участках

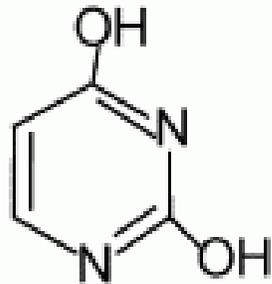
Замены пурин > пурин и пиримидин > пиримидин (**транзиции**)
встречаются чаще, чем
замены пурин > пиримидин и пиримидин > пурин (**трансверсии**)

Но не настолько, как на предыдущем слайде!



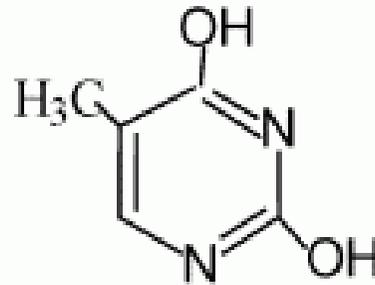
ЦИТОЗИН

и в ДНК и в РНК



урацил

в РНК



ТИМИН

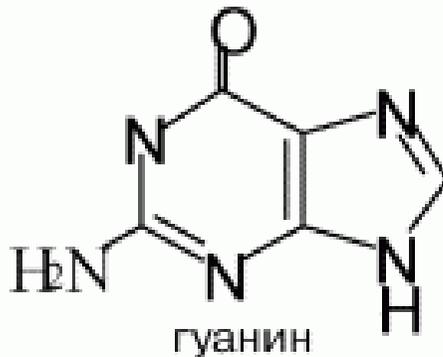
в ДНК

Пиримидины



аденин

и в ДНК и в РНК



гуанин

и в ДНК и в РНК

Пурины

Гомология – общность происхождения

- При репликации почти всегда каждый нуклеотид потомка происходит от определенного нуклеотида предка. В выравнивании гомологичных последовательностей у разных потомков одного и того же предка гомологичные нуклеотиды должны стоять в одной колонке.
- Как правило, нам известны геномы современных организмов и, потому, у нас нет способа проверить какие нуклеотиды гомологичны.

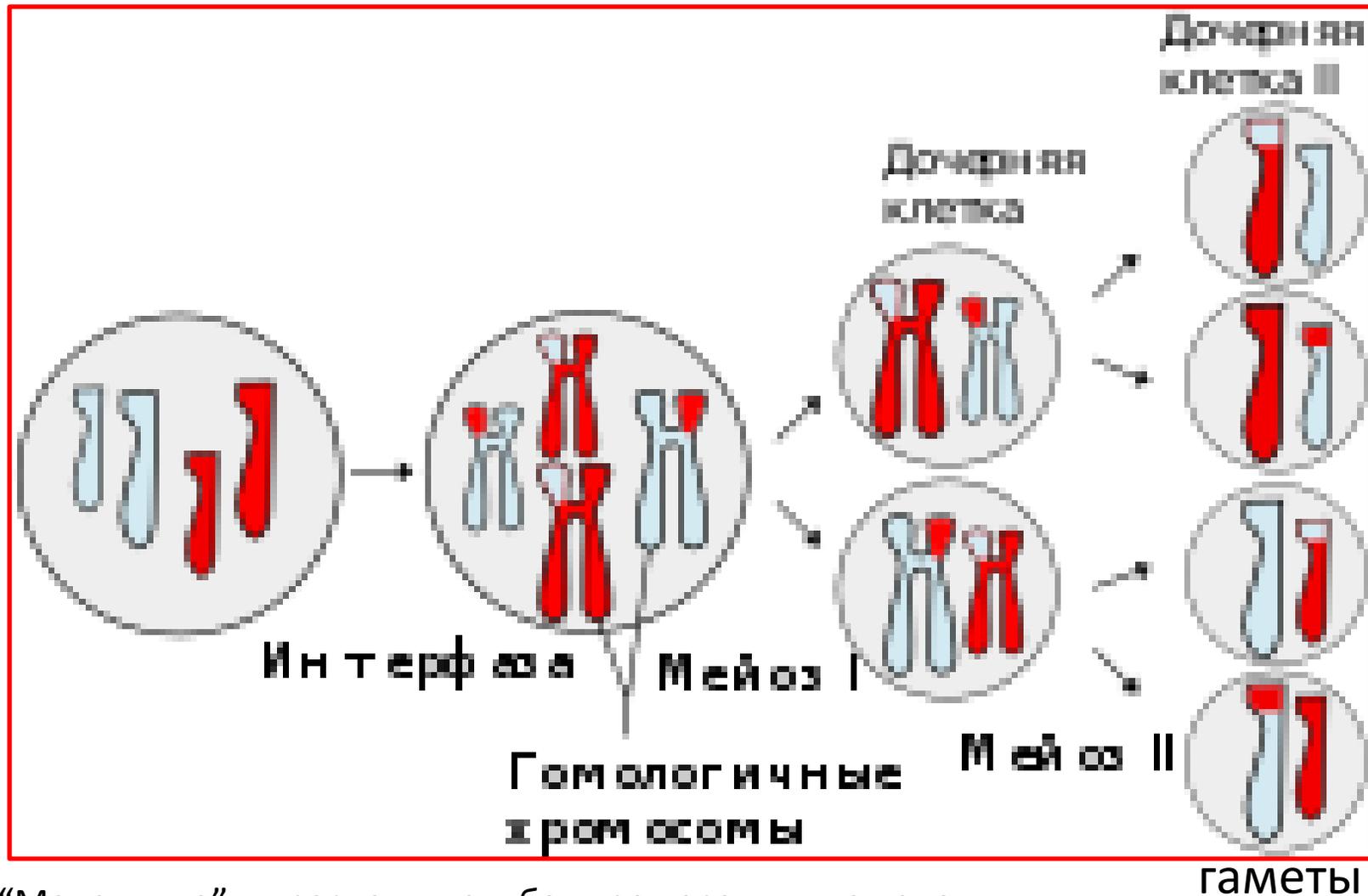
Проблема построения выравнивания будет обсуждаться ниже.

Наследование ДНК у прокариот (бактерии и археи) и эукариот (организмы, в клетках которых ДНК – кроме митохондриальной – содержится в ядре

- Бактерии – одноклеточные. ДНК в каждой клетке удваивается: расплетается и к каждой цепочке достраивается комплементарная; затем клетка делится на две (**МИТОЗ**)
- Соматические клетки эукариот (напр. человека) делятся так же. Отличие в только в том, что в каждой клетке – две копии каждой хромосомы: от папы и от мамы. Такие хромосомы наз. гомологичными

- Половые клетки человека и ему подобных (дрозофила напр.) – яйцеклетки у женщин и сперматозоиды у мужчин, - делятся иначе (**мейоз**).
 - Все хромосомы удваиваются, получается по 4е гомологичных хромосомы в клетке.
 - Гомологичные, но не идентичные хромосомы объединяются в пары. Т.о. в каждой паре одна хромосома от папы, вторая – от мамы.
 - Хромосомы в паре обмениваются несколькими гомологичными участками (**кроссинговер**)
 - Происходит первичное деление, а потом – вторичное. В итоге образуется четыре клетки, в каждой из которых одинарный набор хромосом (гаметы)
У мужчины – из них образуется четыре зрелых сперматозоида
У женщины – одна зрелая яйцеклетка, три других клетки уничтожаются.
 - При оплодотворении одинарные наборы хромосом объединяются в оплодотворенной яйцеклетке - зиготе. Т.о. она содержит правильный – двойной набор хромосом. Далее эмбрион развивается путем митоза этой клетки и ее потомков.

Мейоз



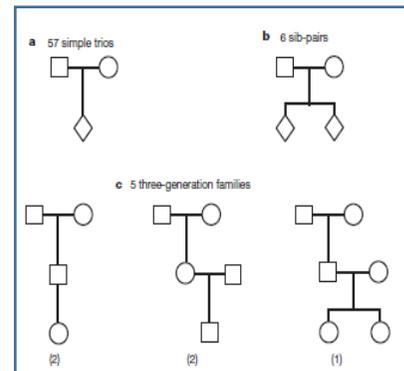
“Маленькие” – красная и голубая хромосомы – гомологичны.
Аналогично относительно “больших” хромосом.

В клетке до мейоза показаны только две пары гомологичных хромосом

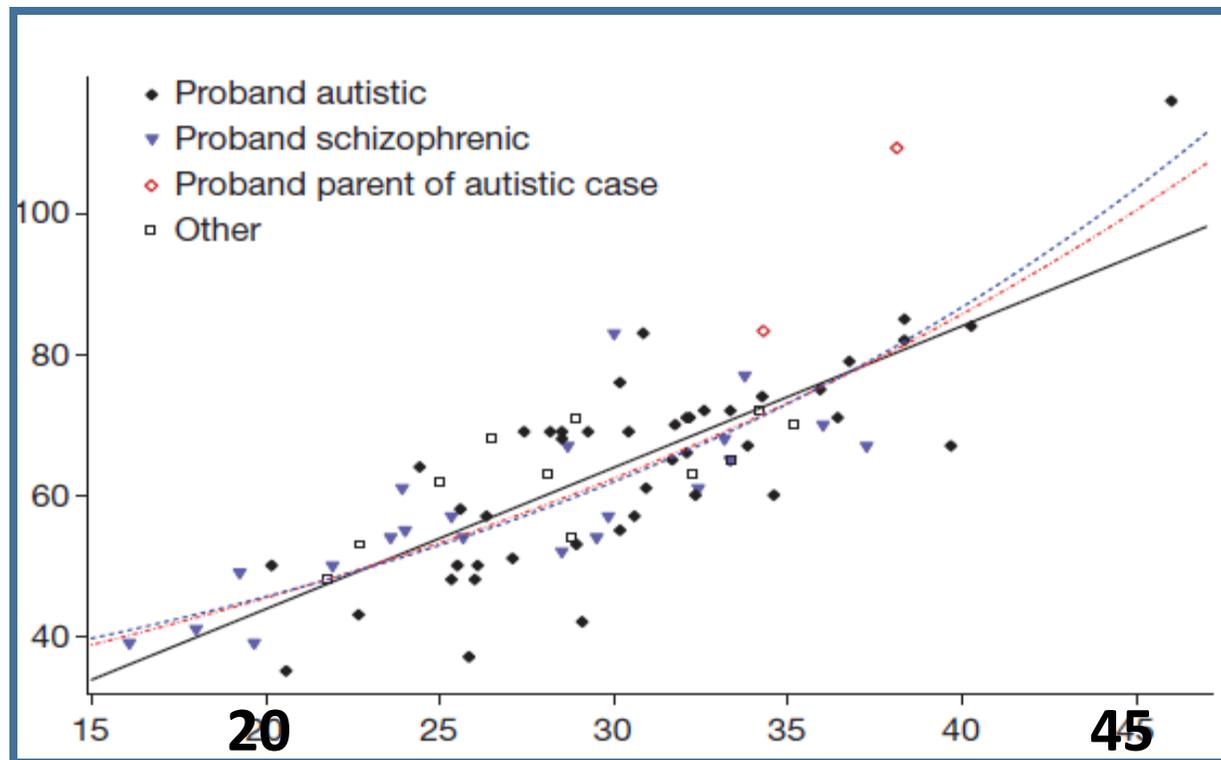
- Яйцеклетки закладываются у девочек на всю жизнь. ДНК в зрелой яйцеклетке перед оплодотворением проходит несколько десятков циклов деления. Это важно для числа мутаций: чем больше делений, тем больше мутаций накапливается в половой клетке.
- Сперматозоиды у мужчин образуются постоянно. Поэтому число делений до зрелого сперматозоида растет с возрастом. Оно составляет несколько сотен.
- С помощью секвенирования геномов троек папа-мама-ребенок (исландцы, голландцы) найдены все мутации *de novo* у детей. Это те нуклеотиды, которые отсутствовали и у папы, и у мамы (и у других людей в данной позиции генома)
- Показано, что число мутаций *de novo* растет в зависимости от возраста отца (но не матери!). Примерно одна мутация на каждый год жизни отца.

Геномы 78 исландских trios

Kong et al., Nature, 2012



число новых мутаций



возраст отца

От зиготы до зиготы:

- у матери 30 делений до яйцеклетки

- у отца -несколько сот делений до сперматозоида

Мутация мутации рознь!

- Мутации могут быть
 - Не совместимыми с жизнью
 - Вредными
 - Слабовредными
 - Нейтральными
 - Полезными (очень редко 😞)
- Чтобы оценивать последствия мутаций надо понимать какая информация записана в геноме
 - Гены белков
 - Гены РНК
 - Регуляторные участки

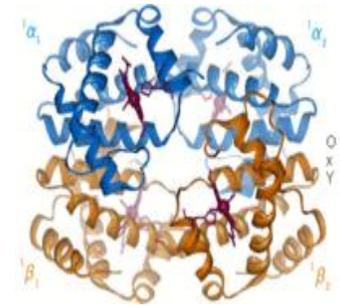
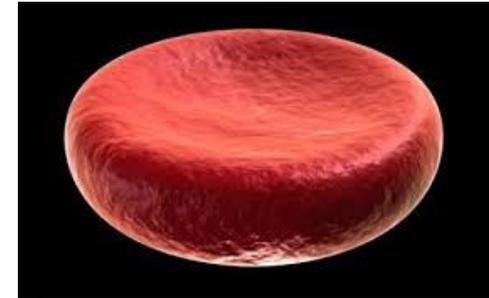
Точечные мутации в генах белков

- Изменяющие аминокислотный остаток (**missense**)
- Кардинально меняющие (убивающие) белок (**nonsense**)
- Молчащие (**silent**)

Примеры см. на след. слайдах

Серповидно-клеточная анемия

- Кислород к тканям переносят эритроциты.
(в эритроцитах нет ядра и ДНК! Они циркулируют в крови около 100—120 дней и затем поглощаются макрофагами.
Приблизительно 25% всех клеток в теле человека — эритроциты)
- Запасают кислород молекулы белка гемоглобина, состоящего из 2х альфа-цепей и двух бета-цепей
- Миссенс мутация одного нуклеотида в гене бета-цепи приводит к заболеванию.



	Thr	Pro	Glu	Glu	beta ^A chain
	...A C T	C C T	G A G	G A G...	beta ^A gene
Codon #	4	5	6	7	
	Thr	Pro	Val	Glu	beta ^S chain
	...A C T	C C T	G T G	G A G...	beta ^S gene

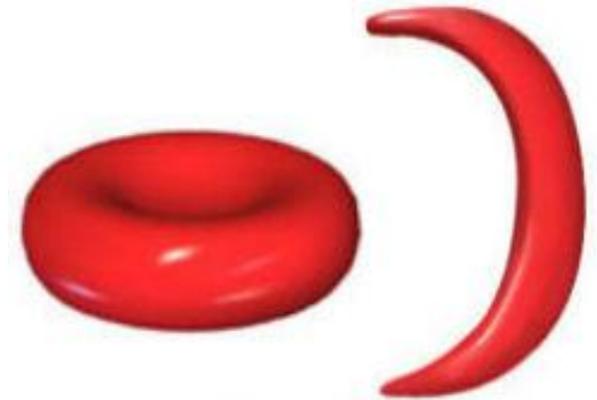
Последовательность правильного гемоглобина бета.

Последовательность мутантного гемоглобина бета.

- Мутация Glu > Val (глутамат > валин) меняет гидроФИЛЬНЫЙ аминокислотный остаток на гидроФОБНЫЙ.
Гидрофобные поверхности стремятся соединиться друг с другом чтобы избежать контакта с водой
- Из-за этого молекулы гемоглобинов в эритроците слипаются в длинную цепочку и эритроцит принимает форму серпа.
- Такие эритроциты живут меньше и хуже переносят кислород.

Не надо думать, что такая мутация окажется вредной в любом белке:

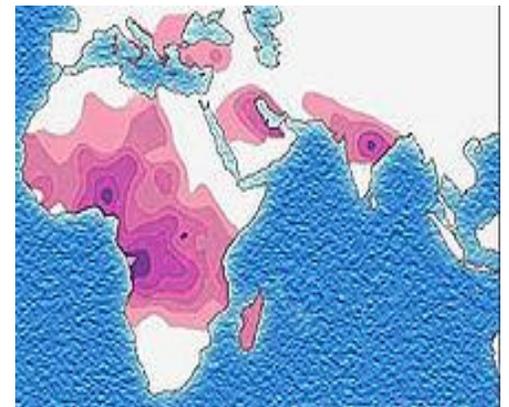
Всё очень индивидуально!



Особенности заболевания

- Если ген с мутацией (Hb-betaS) и в маминой, и в папиной хромосомах (**гомозигота по гену**), то человек тяжело болеет, т.к. ткани получают недостаточно кислорода
- Если ген Hb-betaS только в одной из двух гомологичных хромосом (**гетерозигота по гену**), то человек, в целом, здоров; хуже переносит большую нагрузку.
- Гетерозиготы значительно реже болеют малярией!
Этим объясняется распространенность **аллеля** Hb-betaS в местах где распространена малярия.

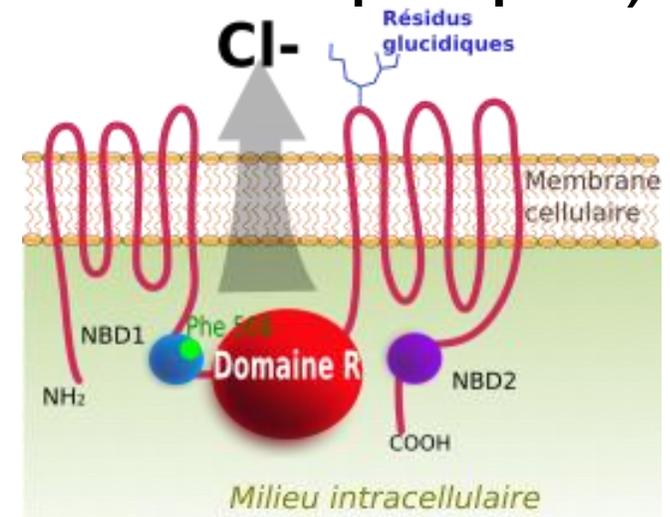
Эта миссенсе мутация:
вредная,
иногда не очень,
иногда даже полезная.



Муковисцедоз (он же – цистический фиброз)

В основе заболевания лежат мутации в гене CFTR:

Patient	Mutation	Result
A	482 C G C ↓ C A C	Arg-117 ↓ His-117
B	1609 C A G ↓ T A G	Gln-493 ↓ STOP
C	Insertion of 2 nucleotides (AT) at 2566	Frameshi ft
D	Deletion of one C at 3659	Frameshi ft
E	Deletion of 3 nucleotides at 1654-1656	Deletion of Phe-508



With a nonsense mutation, the new nucleotide changes a codon that specified an amino acid to one of the **STOP** codons (**TAA**, **TAG**, or **TGA**). Therefore, translation of the messenger RNA **transcribed** from this mutant gene will stop prematurely. The earlier in the gene that this occurs, the more truncated the protein product and the more likely that it will be unable to function.

TTT	Phe	TCT	Ser	TAT	Tyr	TGT	Cys
TTC	Phe	TCC	Ser	TAC	Tyr	TGC	Cys
TTA	Leu	TCA	Ser	TAA	STOP	TGA	STOP
TTG	Leu	TCG	Ser	TAG	STOP	TGG	Trp
CTT	Leu	CCT	Pro	CAT	His	CGT	Arg
CTC	Leu	CCC	Pro	CAC	His	CGC	Arg
CTA	Leu	CCA	Pro	CAA	Gln	CGA	Arg
CTG	Leu	CCG	Pro	CAG	Gln	CGG	Arg
ATT	Ile	ACT	Thr	AAT	Asn	AGT	Ser
ATC	Ile	ACC	Thr	AAC	Asn	AGC	Ser
ATA	Ile	ACA	Thr	AAA	Lys	AGA	Arg
ATG	Met*	ACG	Thr	AAG	Lys	AGG	Arg
GTT	Val	GCT	Ala	GAT	Asp	GGT	Gly
GTC	Val	GCC	Ala	GAC	Asp	GGC	Gly
GTA	Val	GCA	Ala	GAA	Glu	GGA	Gly
GTG	Val	GCG	Ala	GAG	Glu	GGG	Gly

Защита генома от мутаций

- Репарация ДНК
- система иммунного надзора элиминирует мутантные соматические клетки. (Страж генома – белок p53.)
- Онкологические заболевания почти? всегда связаны с мутациями в ДНК

Крупные перестройки ДНК



Человек – шимпанзе: две хромосомы объединились в одну

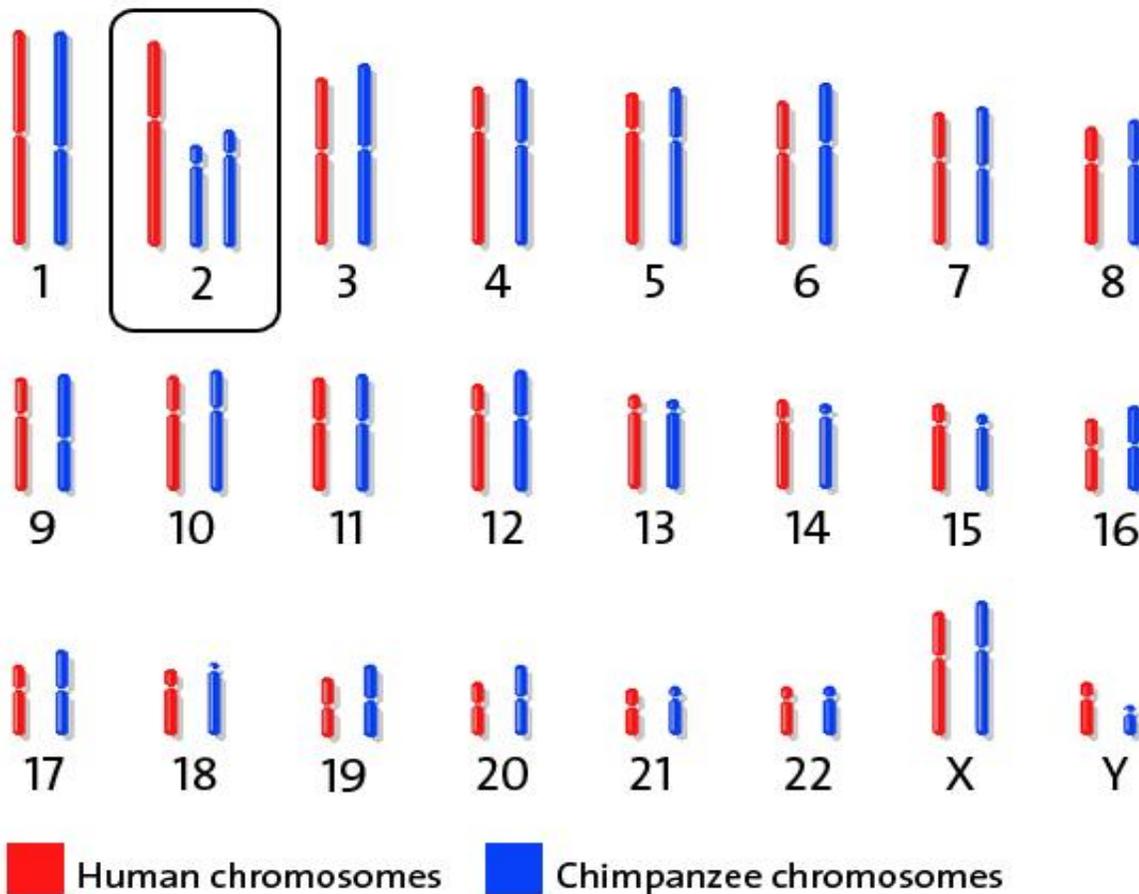
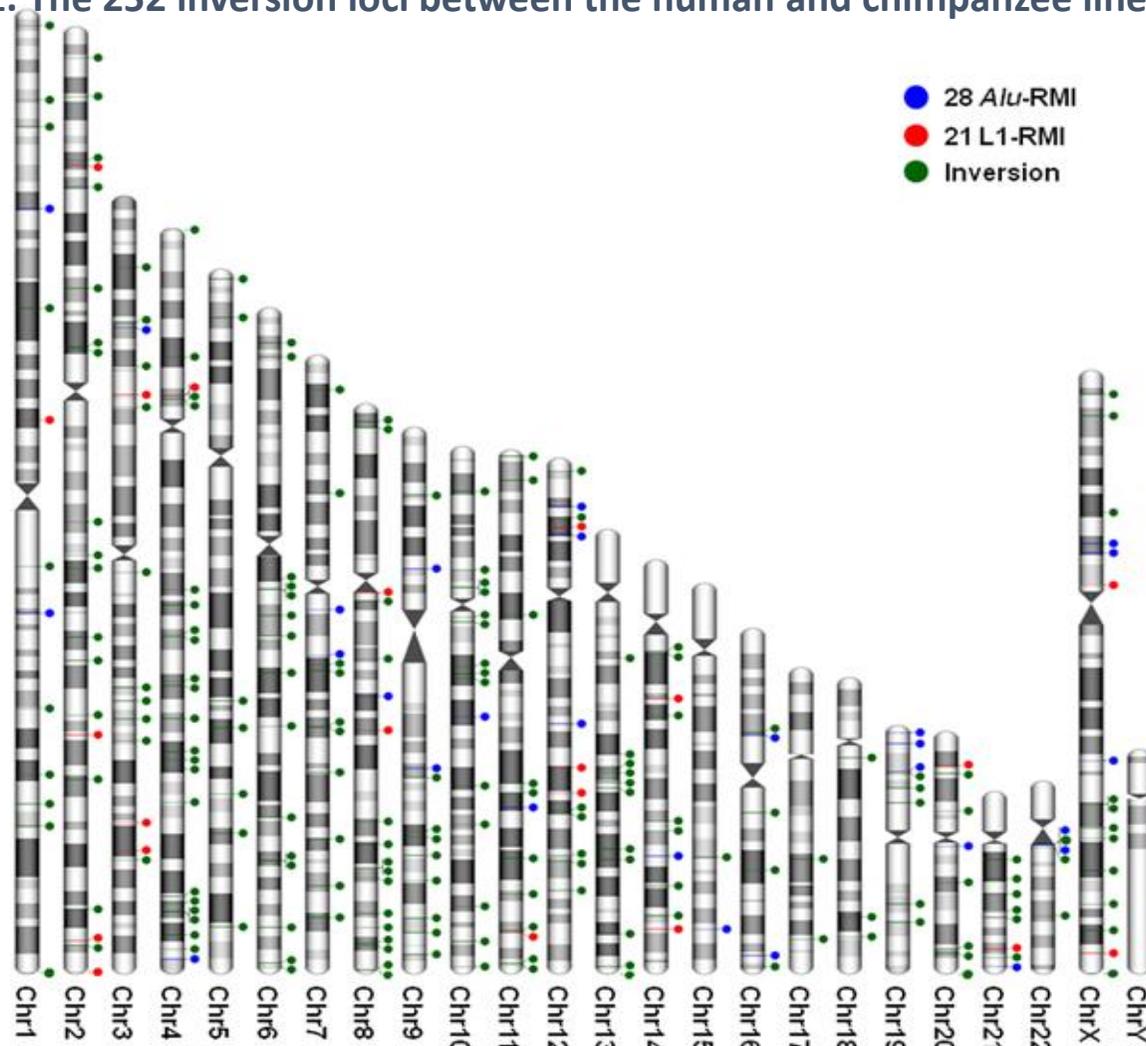


Figure 1. The 252 inversion loci between the human and chimpanzee lineages.



Lee J, Han K, Meyer TJ, Kim H-S, et al. (2008) Chromosomal Inversions between Human and Chimpanzee Lineages Caused by Retrotransposons.

PLoS ONE 3(12): e4047. doi:10.1371/journal.pone.0004047

<http://www.plosone.org/article/info:doi/10.1371/journal.pone.0004047>

Длинные повторы в геноме человека

- (Почти) одинаковые последовательности в геноме человека – повторы.
- В геноме человека содержится около 1 млн копий Alu-повтора длина которого около 300 п.н.. Это составляет около 10,7 % от всего генома.
- Alu повторы могут “размножаться” в геноме. В зависимости от места вставки могут приводить к заболеваниям
- Ретровирусы в геноме

Размеры геномов

Сколько информации содержится в инструкции по созданию живого организма?

Типичные размеры генома

- палочки Коха – возбудитель туберкулеза (*Mycobacterium tuberculosis*)
 - Одна кольцевая хромосома, примерно 4,4 миллиона букв - пар нуклеотидов (п.н.)
 - 3 906 генов белков
 - 102 гена РНК
- человека
 - 24 разных хромосомы 3,3 млрд п.н.
 - 20 – 25 тыс. генов белков
 - ??? генов РНК
 - > 10 000 псевдогенов
- ладанной сосны (*Pinus taeda*)
 - 22 млрд п.н.
 - Аннотировано 50,172 генов

Хромосома – одна молекула геномной ДНК вместе с белками, служащими для поддержания ее структуры (у эукариот)

Минимальный геном свободноживущей бактерии

- Вентер, 2016. На основе ранее созданной “синтетической” бактерии *M. Mycoides* JCVI-syn1.0. новую, названную *M. Mycoides* JCVI-syn3.0
- Эта бактерия размножается (в лабораторных условиях) и все с ней в порядке
- Её геном состоит из 531 пар нуклеотидов и содержит 473 генов.
- Эта бактерия имеет самый маленький геном среди свободно живущих (не паразитов) организмов!
- Цель работы состояла в отборе генов, необходимых для жизни, и удалении всех не необходимых, основываясь на современном понимании.

Hutchison, ..., Venter, “Design and synthesis of a minimal bacterial genome”, *Science*, 2016

Упражнения

- Сколько томов займет публикация генома
 - палочки Коха?
 - человека?
 - ладанной сосны?

Выберите правдоподобное число букв на странице и страниц в томе

Интересно было бы сравнить с объёмом полной технической документации по созданию космического корабля!

Геномы вирусов

Вирусы содержат инструкции на языке, понятном клеткам их хозяина. Инструкции записаны на ДНК или РНК

По этим инструкциям белки клетки хозяина создают новые вирусные частицы

Вне клеток хозяина вирусы не могут размножаться – у них нет белков, способных это сделать

Размеры геномов

- герпесвируса человека 1-го типа – 154 746 пар нукл ДНК
- вируса гриппа – 13 158 нуклеотидов РНК
- вируса Эбола – 18 959 нуклеотидов РНК
- вириода веретеновидности клубней картофеля – 359 (!) нуклеотидов РНК



Вот полная инструкция вириода.

Он даже не вирус – только кольцевая молекула РНК, оболочки нет!

>NC_002030 Potato spindle tuber viroid, complete genome

```
CGGAACТАААСТСГТGGТTCCTGTGGТTCACACCTGACCTCCTGAGCAGA  
AAAGAAAAAAGAAGGCGGCTCGGAGGAGCGCTTCAGGGATCCCCGGGGAA  
ACCTGGAGCGAACTGGCAAAAAAGGACGGTGGGGAGTGCCCAGCGGCCGA  
CAGGAGTAATTCCCGCCGAAACAGGGTTTTACCCCTTCCTTTCTTCGGGT  
GTCCTTCCTCGCGCCCGCAGGACCACCCCTCGCCCCCTTTGCGCTGTCGC  
TTCGGCTACTACCCGGTGGAAACAАСТGAAGCTCCCGAGAACCGCTTTTT  
CTCTATCTTACTTGCTTCGGGGCGAGGGTGTTTAGCCCTTGGAACCGCAG  
TTGGТTCCT
```

Алгоритм выравнивания последовательностей ДНК и белков

Парное выравнивание: динамическое программирование

Число операций в алгоритме

- Пусть мы применяем алгоритм к последовательности из N букв. Для примера, $N = 10\,000$
 - если число операций $10 * N$, то любой компьютер посчитает за секунду
 - если число операций $5 * N^2$, то в примере имеет более миллиарда операций; часы или дни счета
 - если N^3 , то уже совсем плохо, но если кластер, то за месяц посчитает
 - если 2^N то посчитать невозможно
 - Алгоритм сортировки (упорядочить данные по алфавиту) ?за $N * \log N$;
 $\log 10\,000 = 4$, значит $4 * 10\,000$ – ерунда!

Алгоритм выравнивания решает математическую задачу, а не биологическую

- Задача:
 - Дано две последовательности
 - Построить выравнивание с наибольшим весом

РИСУНОК вес выравнивания белков

РИСУНОК Матрица весов

Вес выравнивания

Sequence 1	L	E	L	-	-	S	H	V	Q
Sequence 2	L	I	L	E	P	S	H	F	Q
Вес позиции	4	-3	4	-4	-4	4	8	-1	5

- Штраф за “-” (гэп) = -4
- Веса за замены берутся из матрицы; они отражают взаимозаменяемость аминокислот
- Вес выравнивания = 25

Матрица весов аминокислотных замен BLOSUM 62

Треугольная (симметричная)
матрица

	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W
C	9																			
S	-1	4																		
T	-1	1	5																	
P	-3	-1	-1	7																
A	0	1	0	-1	4															
G	-3	0	-2	-2	0	6														
N	-3	1	0	-2	-2	0	6													
D	-3	0	-1	-1	-2	-1	1	6												
E	-4	0	-1	-1	-1	-2	0	2	5											
Q	-3	0	-1	-1	-1	-2	0	0	2	5										
H	-3	-1	-2	-2	-2	-2	1	-1	0	0	8									
R	-3	-1	-1	-2	-1	-2	0	-2	0	1	0	5								
K	-3	0	-1	-1	-1	-2	0	-1	1	1	-1	2	5							
M	-1	-1	-1	-2	-1	-3	-2	-3	-2	0	-2	-1	-1	5						
I	-1	-2	-1	-3	-1	-4	-3	-3	-3	-3	-3	-3	-3	1	4					
L	-1	-2	-1	-3	-1	-4	-3	-4	-3	-2	-3	-2	-2	2	2	4				
V	-1	-2	0	-2	0	-3	-3	-3	-2	-2	-3	-3	-2	1	3	1	4			
F	-2	-2	-2	-4	-2	-3	-3	-3	-3	-3	-1	-3	-3	0	0	0	-1	6		
Y	-2	-2	-2	-3	-2	-3	-2	-3	-2	-1	2	-2	-2	-1	-1	-1	-1	3	7	
W	-2	-3	-2	-4	-3	-2	-4	-4	-3	-2	-2	-3	-3	-1	-3	-2	-3	1	2	11
	C	S	T	P	A	G	N	D	E	Q	H	R	K	M	I	L	V	F	Y	W

Из работы (Henikoff&Henikoff, **1992**, PNAS)

Три понимания «правильного» выравнивания

- 1** **Оптимальное выравнивание:** наилучшее совпадение последовательностей белков
Его ищут программы, иногда вместе с чуть менее хорошими, следующими за оптимальным выравниваниями (про оценку качества выравнивания будет сказано немного после)
- 2** **Эволюционное выравнивание:** запись, отражающая ход эволюции
Не поддается достоверной реконструкции в большинстве реальных случаев; может отличаться от оптимального выравнивания. Эволюция обеспечивает совпадение между белками, за счет этого можно ожидать, что это выравнивание будет среди нескольких оптимальных при корректном выборе алгоритма
- 3** **Функциональное выравнивание:** сопоставление функционально идентичных частей белков
 Объясняет сохранение в эволюции одних частей белка и варьирование других; поскольку функция и 3D-структура очень тесно связаны, должно отображать сопоставление структур белков

Идея алгоритма выравнивания

	♦	L	I	L	E	P	S	H	F	Q
♦	0	-4	-8							
L	-4	4	0							
E	-8	0	1							
L										
S										
H										
V										
Q										

Вес = max (4-3, #match

0-4,

0-4) = 1



LI

LE

Локальное выравнивание

- Основное отличие в том, что оно может начаться и закончиться в любой клеточке

	♦	L	I	L	E	P	S	H	F	Q
♦										
G					-2	-2				
P					-1					
S										
H										
W										

Вес = $\max(-2+7, \text{match}$

$-2-4,$

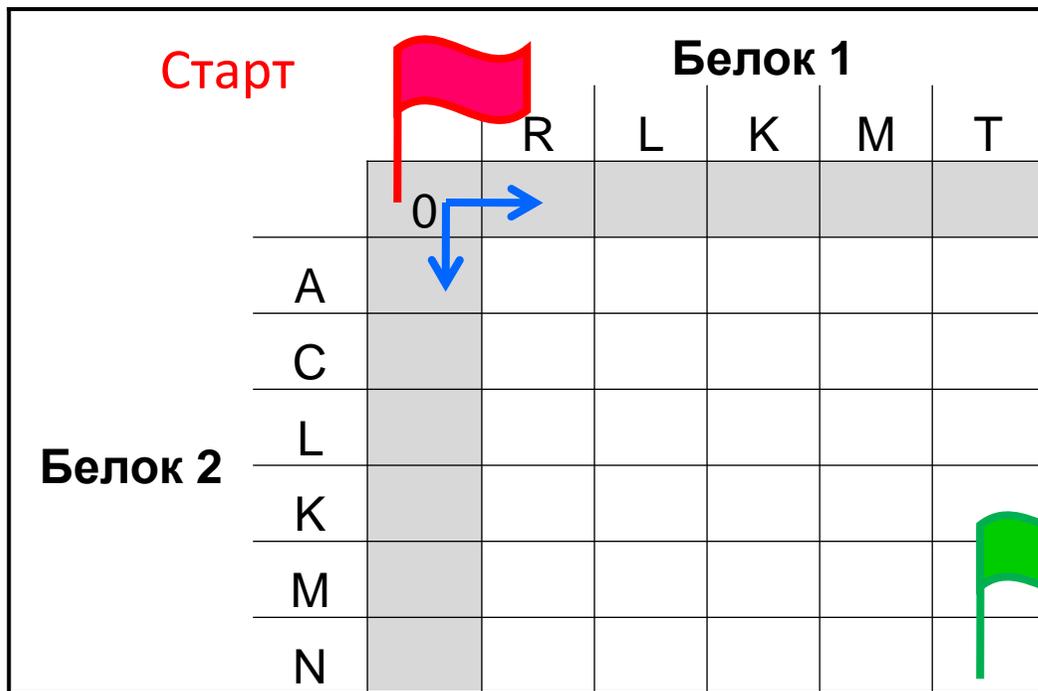
$-1-4,$

$0+7$

$) = 7$

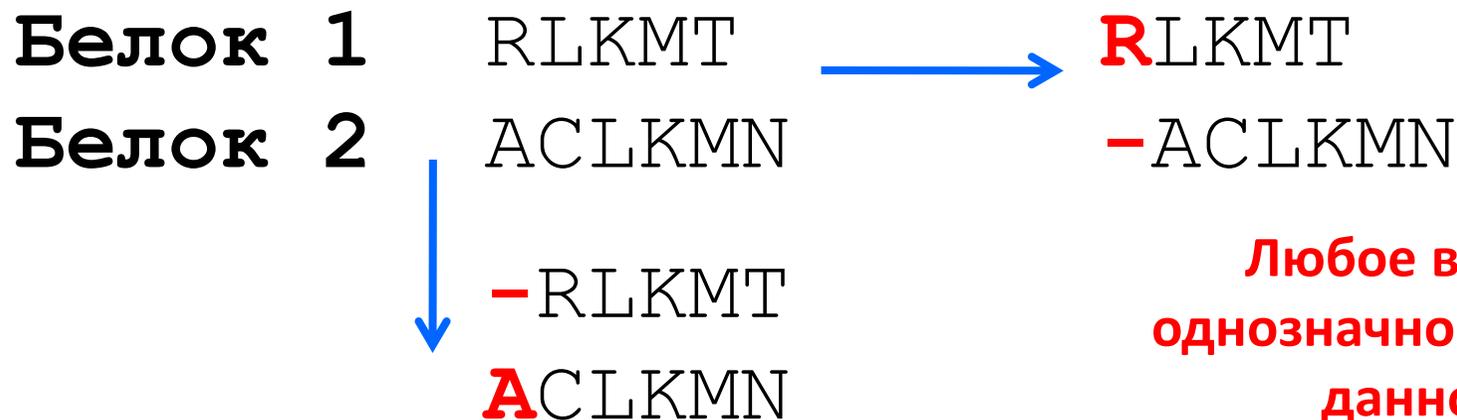
начать выравнивание с этой позиции

Динамическое программирование



Всегда есть два варианта
вставки гэпа:

- 1) вставить гэп в белок 1
- штраф за гэп
- 2) вставить гэп в белок 2
- штраф за гэп



Любое выравнивание
однозначно задается путем в
данной таблице

Динамическое программирование

		Белок 1					
		R	L	K	M	T	
Белок 2		0	-2	-4	-6	-8	-10
	A	-2					
	C	-4					
	L	-6					
	K	-8					
	M	-10					
	N	-12					

Пусть штраф за гэп будет
= 2

Верхняя строка и левый столбец
заполняются штрафами за гэпы

«**Концевые гэпы**»: можно
дополнительно оштрафовать за то,
что «финиш» не в конце матрицы
(т.е. нужно вставить концевые гэпы
в одну из последовательностей)

Параметры **needle**

-endweight

-endopen

-endextend

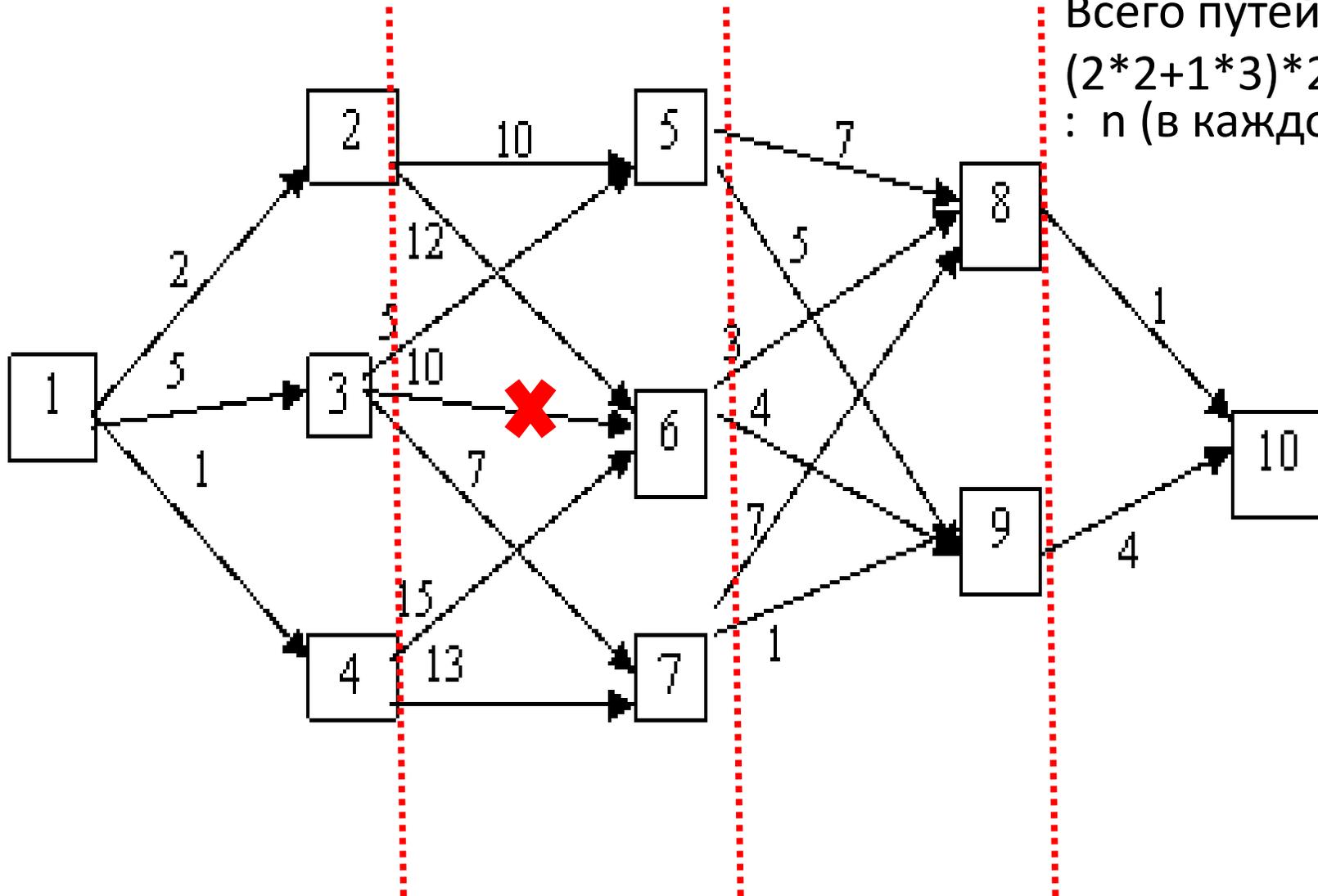
Белок 1 RLKMT

Белок 2 -----ACLKMN

RLKMT-----

-----ACLKMN

Найти кратчайший путь



Всего путей: $2^{(n-2)}$

$(2*2+1*3)*2 = 14$

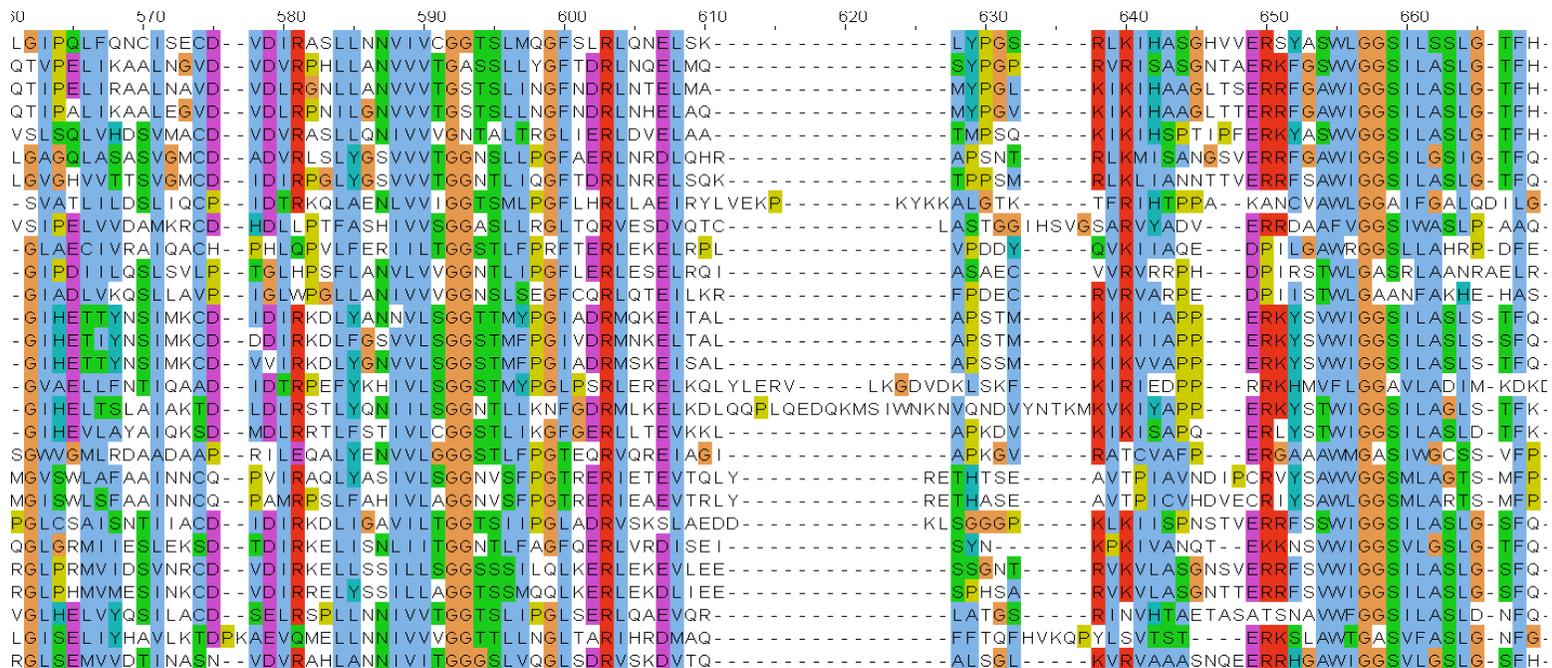
: n (в каждой вершине)

Пример хорошего выравнивания

```
11 DRREIRHIWDDWSSSF TDRRVAIVRAVFDDL FKHYP TSKALFERVKIDEPESGEF 66
8 DRHEVLDNWKGIWSAEFTGRRVAIGQAIFQELFALDPNAKGVFGRVNVDPKSEADW 63
8 DRREVQALWRSIWSAEDTGRRTLIGRLLFEELFEIDGATKGLFKRVNVDDTHSPEE 63
7 QRIVKVKQWVAQVYVSGES--RTDFAIDVFNNFFRTNPD-RSLFNRVNGDNVYSPEF 59
9 QRLKVKQWAKAYGVGHE--RVELGIALWKSMAQDNDARDLFRVHGEDVHSPAF 62
8 EGLKVKSEWGRAYGSGHD--REAFSQA IWRATFAQVPESRSLFKRVHGDDTSHPAF 61
6 QRFKVKHQWAEAFGTSHH--RLDFGLKLWNSIFRDAPEIRGLFKRVDGDNAYSAEF 59
7 QRLKVKRQWAEAYGSGND--REEFGHF IwthvfkDAPSARDLFRVVRGDN IHTPAF 60

67 KSHLVRVANGLDLLINLLDDTLV LQSHLGH LADQH IQRKGV TKEYFRG IGEAFA 120
64 KAHVIRVINGLDLAVNLLLEDPKALQEELKH LARQHRERSGVKAVYFDEMEKALL 117
64 FAHVLRV VNGLDTLIGVLGDSDTLNSLIDHLAEQHKARAGFKTVYFKEFGKALN 117
60 KAHMVRVFA GFDILISVLDDKPVLDQALAHYA AFHKQF-GTIP--FKAFGQTMF 110
63 EAHMARVFNGLDRVISSLTDEPVLNAQLEHLRQQHIKL-GITGHMFNLMRTGLA 115
62 IAHAEV LGGLDIAISTLDQPATLKEELDHLQVQHEGR-KIPDNYF DAFKTA I L 114
60 EAHAEV LGGLDMTISLLDDQA AFDAQLAHLKSQHAER-NIKADYYGVFVNELL 112
61 RAHATRV LGGLDMCIALLDDEGV LNTQLAHLASQ HSSR-GVSAAQYDVVEHSVM 113
```

Или так ...



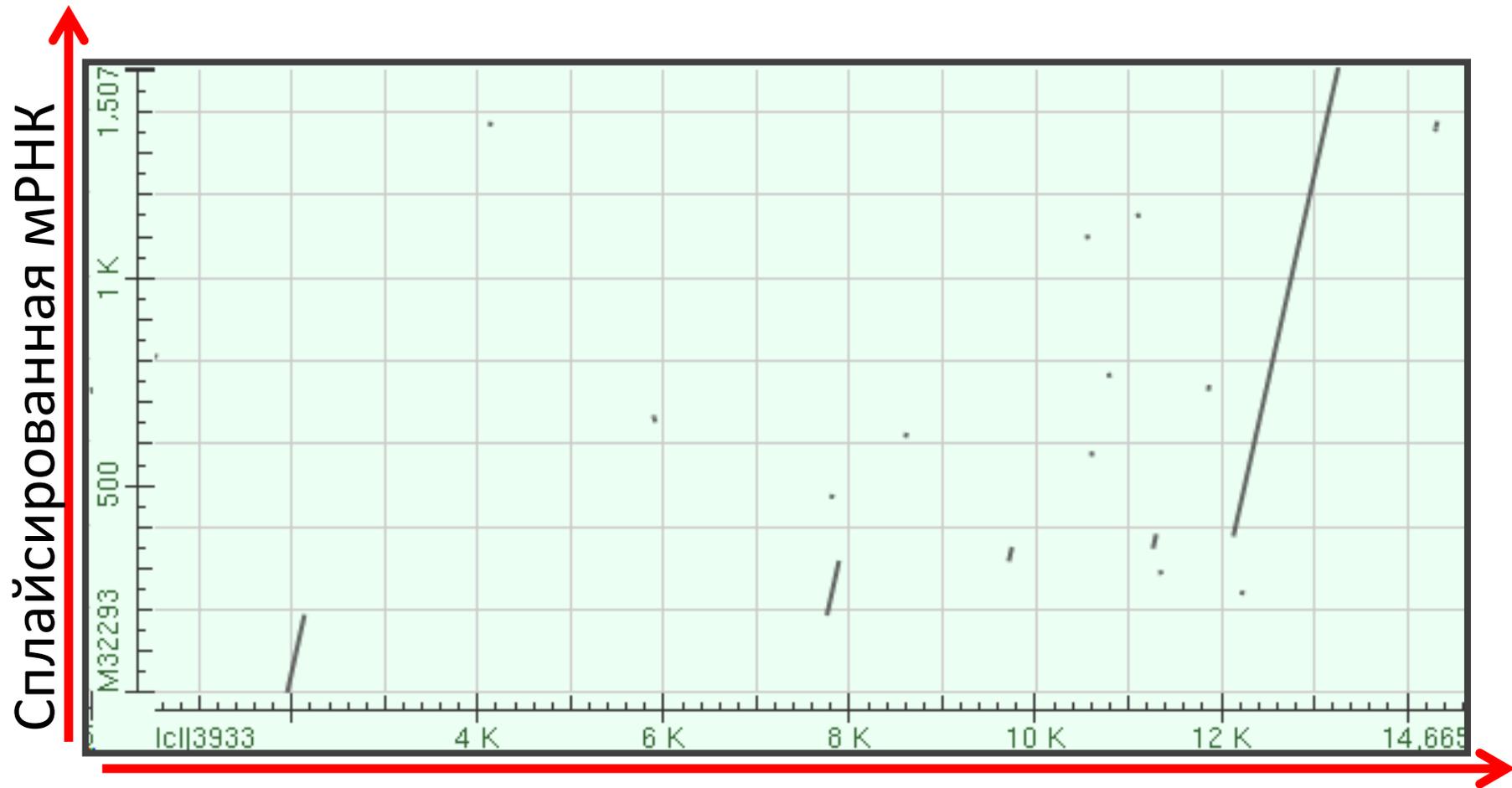
Раскраска – ClustalX

	10	20	30	40	50	60	70
3	VSLAGRMMF	KRVMGKAS	FCNIQDL	KGRIQIY	VARDSIG	EESYADF	FKK-YDV
	GDILGVK	GIVFKTK	TGGEIS	IHAAE	VTL		
'	VRLTGR	LMAVREH	GKASFAD	IEDMSG	RIQLYAK	EDNLGED	LYELFMS-LD
	IGDQIG	IEGLLF	KTGRGE	VTVQI	ENFELL		
!	VSIAGR	LMSRRDM	GKANFID	VHDATGR	MQVYVR	VNDVGEES	FAFVKK-WD
	IGDIVG	VKGVK	KVFRTR	RGEIS	IHATA	VTL	
'	ATVAGR	IMTKRG	GKGVGF	AHLQDR	KGQIQI	YVRKDT	VGEEDYET
	ADLGD	FIGVT	GTIMKT	DTGEV	TIKPTS	ITQL	
	VTLAGR	LMAKRG	GQKAGF	SNIQDL	TGRIQL	YTRLDDL	GEEEFKRF
	LDIGD	IIGVT	GKVFKT	QKGEIS	IWVKS	YKLL	
3	VAVAGR	IMTKRG	GKAGFA	HIQDVT	GQIQIY	VRQDDV	GEEQYEL
	SDLGD	IVGVR	GTMFKT	KVGELS	IKVSS	YEFLL	
'	VSLAGR	MMSKR	VMGKAS	FANLQDR	DGRMQL	YVSRNDL	GDEAYADF
	MDIGD	IIGVK	GFAFRT	KTGEIS	SVHVKE	LTL	
	VKIAGR	IMAIRG	HGKA	AFDQD	NTGKI	QVYIRK	DVSDDT
	LDIGD	IVGIE	GNVFKS	KRGEIS	SVS	VNNMTIL	
'	VSIAGR	MMSKR	VMGKAS	FCNIQDL	KGNIQV	YVARDN	VGEDSYAD
	SDIGD	IYGVK	GFAFRT	KTGEIS	IHAAE	ITLL	
'	VTLAGR	IMTKR	SHGKAS	FADIQDR	DGRIQI	YVRYDN	VGGKNEY
	LDIGD	IIGVT	GEVFKT	KTGEI	TINVK	DFKLL	
'	VSVAGR	IMTKR	VKGKVG	FTHIQDR	RFHQL	QIYIRK	DAIGEDAY
	ADLGD	IIGIK	GTFRTN	TGELS	VKATE	FTL	
'	VAIAGR	LMTKR	GKAGFA	AHVQDL	AGQIQI	YVRKDQ	VGDDEFD
	ADLGD	IVGVE	GVMFKT	NTGELS	VKAKK	FTL	
?	VKIAGR	MMSKR	DQGVV	FTHIQD	FSGQIQ	AYIRKDD	VGEDWFEL
	FDIGD	IVGIK	GTVFRTR	RGEIS	SVHANE	VQIL	
'	VVIAGR	MIAKR	GAGKVI	FADFRD	VSGKIQ	VYAKR	DDLGEN-YP
	ADLGD	FLGIK	GIMKTE	AGELTV	LATEL	THL	
	VSLSGR	IMAKR	GHGKAG	FAHIQD	SSGQVQ	IYTRVDN	VGEEDYEL
	LDIGD	IALKG	PVFMTR	KGEIT	VEVKE	FHLQ	
!	VKVAGR	LMSKR	GQGVV	FADLADL	PGKIQ	LFIKIDN	VGEEALKE
	FDLGD	WVAAT	GEVFKT	KMGEV	SVKVT	SFELI	
	-IVAGR	LMTKR	GKGVGF	AHLQD	REGQIQ	LYVRKDE	VGEENEY
	ADLGD	FLGVE	GEIMKT	DMGELS	IKAKK	LTFLL	
	VTLAGR	IMSKR	THGKAS	FADIQD	REGRIQ	IYVKYDT	VGEKNEY
	LDIGD	IIGVT	GEVFKS	KTGEI	TRVKD	FKLL	
	VTIAGR	MMSKR	GKGVGF	ADIRDR	SGKMQI	YVRKDV	VGDDNYM
	ADLGD	ILGIS	GEVMKT	DSGEL	TVKAN	HVTHL	
9	VTIAGR	ITAKR	VMGKAS	FANIQD	SRGNIQ	VYARLNDL	GAEYARLEK
	LDIGD	IIGCY	GKVFRT	RKGEI	TVHID	SYKLL	
3	VTIAGR	MMAKR	GHGKAS	FINIQD	KDGRIQ	VYVREDK	IGEKEYEL
	YDRGD	ILEIQ	GEIFKT	NKGEV	SVKALE	VKLL	
'	VTIAGR	MMTKR	GKGVGF	AHLQD	LEGQIQ	IYVRKDS	VGDDQYE
	SDLGD	LIGVT	GKVFKT	NVGELS	VKATS	SFELL	
3	VSLAGR	IMSKR	DQGVIF	FTHIQD	IDGRLQ	IYIRMD	DLGQAMFEL
	FDVGD	IIGVQ	GKVFRT	KRGEIS	IHAKE	VTL	
	VSIAGR	VMTTRR	KGVGF	TNVQD	VNGQIQ	LYVRKDD	VGEEAYE
	CDLGD	IVGIS	GYVFKT	NVGELS	SVHVT	SFTLL	
	ATIAGR	LVTKR	GKGVGF	AHLQD	REGQIQ	IYVRKDA	VGEENEY
	ADLGD	FLGVE	GEVMRT	DMGELS	IKATHI	THL	
	VSVAGR	MLSKR	GHGKAS	FVDILDP	KGIQIY	IKIDEV	GEEKYQEF
	DI	GDIGV	KGEVFK	THKGEIS	IKAKE	IEML	
?	VSIAGR	IMAKR	GHGKAS	FCDIM	DELGEI	QLYVKM	DEVGSEAYEL
	LDIGD	ILGVI	GIVFKT	HSGEIS	SVKVK	TFELL	
?	VRVAGR	VMGLRR	HGRAAF	ADLQDV	SGRLQL	LARVGP	LPEDEYRAF
	LDRGD	WIGA	ETVIRTR	RGEIS	SVEVT	GFVLL	
	VRIAGR	LVARR	GHGKAT	FAVLND	VTGNIQ	IYAKYD	VLGPEAYD
	VDLGD	ILGVE	GTVFRT	NRGEI	TVLAE	QVTFL	
'	ASVAGR	MVSKR	GKGVGF	AHLQD	REGQIQ	IYVRKDE	VGEDNYD
	ADLGD	FLGVT	GEIMKT	NTGEV	SVKAE	KVTLL	
	VTIAGR	IMSKR	GHGKAS	FADIQDR	DGRIQI	LYFKMD	VLGKNEY
	LDIGD	IIGVT	GEVFKS	KTGEV	TRVQD	FKLL	
?	VTIAGR	LMAKR	GMGKAS	FAQLLDR	SGQIQI	YVRQDT	VGEELSKV
	ADIGD	MIGVT	GVVFKT	KTGELS	VKAKEL	SYL	

КОНЕЦ

презентации

Карта сходства последовательностей



Кодирующая последовательность ДНК из курицы