

Медицинская геномика

Василий Евгеньевич Раменский
Анастасия Александровна Жарикова

`ramensky@gmail.com, azharikova89@gmail.com`

НМИЦ Терапии и профилактической медицины
Факультет биоинженерии и биоинформатики МГУ
Институт искусственного интеллекта МГУ

2026

План курса

12 понедельников: 05.10 – 21.12, 8 тем + зачет:

1. Мутации *de novo* в геноме. Структурные варианты и анеуплоидии
2. Передача мутаций
3. Варианты в пространстве [генома] и их последствия
4. Варианты во времени: некоторые основы популяционной генетики
5. Варианты в индивидуумах и популяциях
6. Моногенные заболевания: поиск генов, диагностика
7. Некоторые основы количественной генетики
8. Комплексные (мультифакторные) заболевания
9. **Зачет**

Примечания:

- Короткие (3-10 мин) письменные тесты (квизы) в начале лекции,
- которые выполняются самостоятельно, без помощи Вселенского Разума.
- Для получения зачета нужно написать >60% (не менее 7) тестов,
- а для получения зачета-автомата — написать их хорошо

Примечания

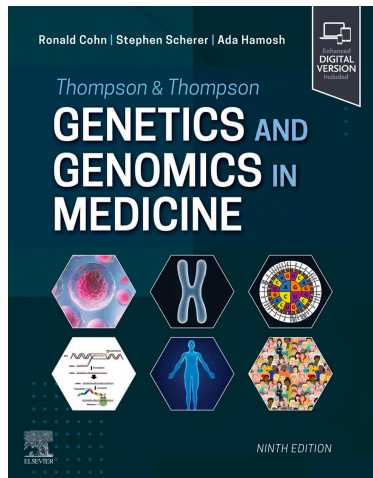
- Медицинская геномика $\neq$ Медицинская генетика
- = молекулярная генетика + популяционная генетика + медицинская генетика + статистическая генетика + генетическая эпидемиология + биоинформатика $\Rightarrow$ Нет единого учебника (?)
- За рамками курса: иммунология, патогены, микробиом, терапия, редактирование генома
- Слайды лекций: <https://kodomo.fbb.msu.ru/wiki/2023/7>

На слайдах:

- Prerequisites: 
- Слайды для самостоятельного изучения: 
- Вопросы и задачи
- Определения, выводы, основные понятия, список литературы

Учебники

1. Jackson *et al.* **The genetic basis of disease.** (2018) *Essays Biochem* 62:643–723.
<https://doi.org/10.1042/EBC20170053>
2. Thompson & Thompson. **Genetics and genomics in medicine.** (2024) ISBN 978-0-323-54762-8
3. Strachan, Read. **Human Molecular Genetics.** (2019) ISBN 978-0-815-34589-3.
4. Gillespie. **Population genetics. A concise guide.** (1998) ISBN 0-8018-5764-6
5. Griffiths *et al.* **An Introduction to Genetic Analysis.** (2015) ISBN 1464109486
6. Бочков и др. **Клиническая генетика.** (2018) ISBN 978-5-9704-4628-7



Тест 1 (05.10.26), 5 мин

Напишите на карточке **Ваше имя, фамилию и отчество, Ваш курс (или «Магистратура»)**, номер варианта и дайте ответ на вопросы:

Вариант 1

1. Дайте примерную оценку числа белок-кодирующих генов в геноме человека
2. Генотипы родителей Aa и Aa, у них 10 детей. Перечислите их наиболее вероятные генотипы
3. Приведите пример генетического заболевания

Вариант 2

1. Дайте примерную оценку длины генома человека (в парах нуклеотидов)
2. Генотипы родителей Aa и aa, у них 10 детей. Перечислите их наиболее вероятные генотипы
3. Приведите пример генетического заболевания

Генетические заболевания

OMIM Gene Map Statistics

OMIM Morbid Map Scorecard (Updated October 2nd, 2026) :

Total number of phenotypes* for which the molecular basis is known	7,762
Total number of genes with phenotype-causing mutation	5,093

* Phenotypes include (1) single-gene mendelian disorders and traits; (2) susceptibilities to cancer and complex disease (e.g., BRCA1 and familial breast-ovarian cancer susceptibility, [113705.0001](#), and CFH and macular degeneration, [134370.0008](#)); (3) variations that lead to abnormal but benign laboratory test values ("nondiseases") and blood groups (e.g., lactate dehydrogenase B deficiency, [150100.0001](#) and ABO blood group system, [110300.0001](#)); and (4) select somatic cell genetic disease (e.g., GNAS and McCune-Albright syndrome, [139320.0008](#) and IDH1 and glioblastoma multiforme, [147700.0001](#).)

Генетические заболевания

1. Хромосомные

- Причина: изменения на уровне целых хромосом или их фрагментов
- Частота: $\approx 50\%$ выкидышей в I триместре, 1:30 новорожденных
- Пример: синдром Ди Джорджи (OMIM:188400, делеция 22q11.2, 1:3,000-6,000 новорожденных)

2. Моногенные (менделевские)

- Причина: патогенные мутации в отдельных генах
- Частота: 1:300 новорожденных, 1:50 взрослых
- Пример: синдром Марфана (OMIM:154700, фибриллин-1 *FBN1*, 1:3,000-5,000 новорожденных)

Генетические заболевания

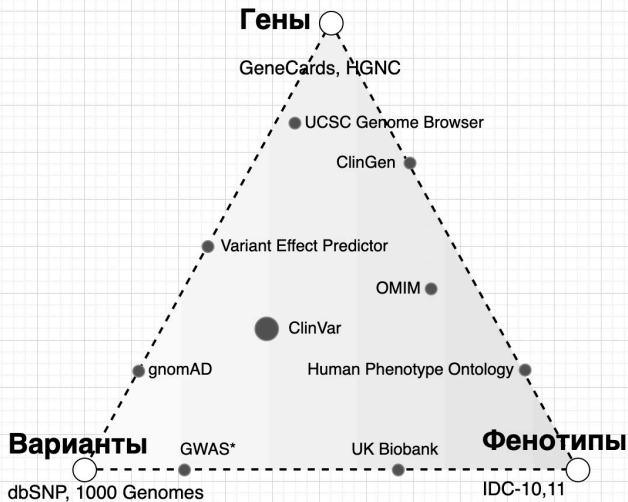
3. Комплексные (мультифакторные)

- Причина: совместное влияние [большого числа] различных вариантов
- Частота: $\approx 5\%$ новорожденных, $\approx 60\%$ взрослых
- Пример: болезнь Крона (OMIM:266600, 1:150-350 человек)

4. Митохондриальные (аномалии окислительного фосфорилирования)

- Причина: мутации в митохондриальной (70%) или ядерной (30%) ДНК
- Частота: 1:5,000 новорожденных
- Примеры: синдром Лея (Ли, подострая некротизирующая энцефаломиелопатия, OMIM:500017, 1:11,000-34,000 новорожденных)

Схема курса



Лекция 1

Мутации *de novo* в геноме. Структурные варианты и анеуплоидии

План лекции

- 1 Медицинская геномика: описание курса
- 2 Кариотип и состав генома человека
- 3 Точность репликации и частота мутаций
- 4 Мутации *de novo*
- 5 Дефекты в системе репарации
- 6 Структурные варианты и варианты числа копий
- 7 Анеуплоидии
- 8 Инактивация X-хромосомы
- 9 Выводы и литература

Prerequisites для этой лекции

- Митоз и репликация ДНК
- Экзогенные и эндогенные механизмы повреждения ДНК
- Механизмы рекомбинации и репарации ДНК

Кариотип и геном

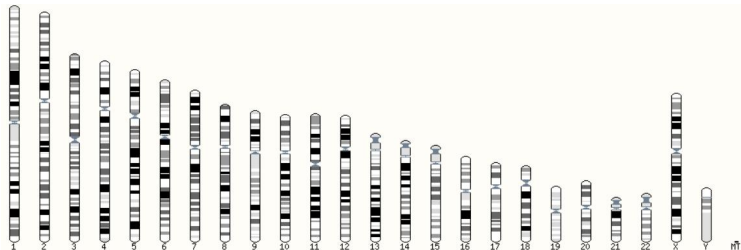
Кариотип человека 🏠



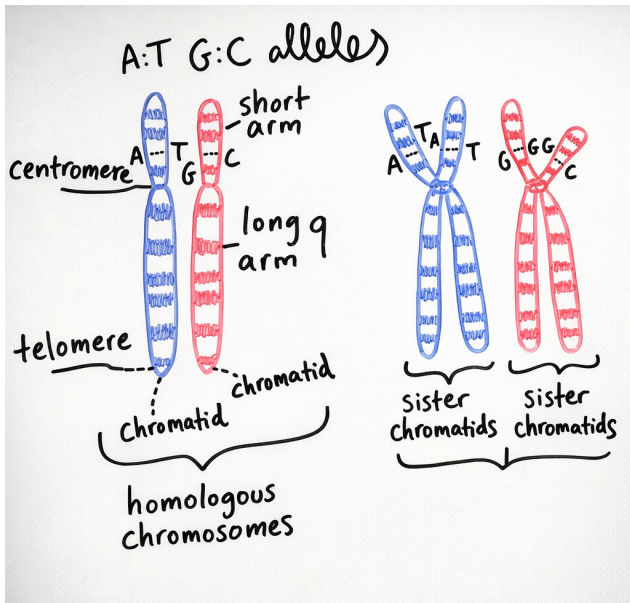
22 пары аутомсомных хромосом + 2 половые хромосомы

Кариотип человека 🏠

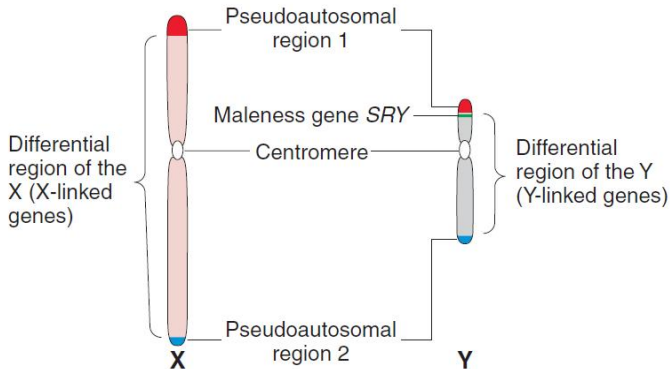
- **Эухроматин** (2.9 Гб) – участки ядерного генома, содержащие большое количество генов и являющиеся транскрипционно активными
- **Гетерохроматин** (0.2 Гб) – плотно уложенная (компактная) ДНК, транскрипционно неактивная, содержащая большое количество повторов. Находится в центромерах и теломерах
- **Метацентрические хромосомы** – хромосомы, у которых центромера находится примерно посередине, имеющие плечи одинаковой длины
- **Акроцентрические хромосомы** – ассиметричные хромосомы с плечами разной длины



Аутосомные хромосомы



Половые хромосомы

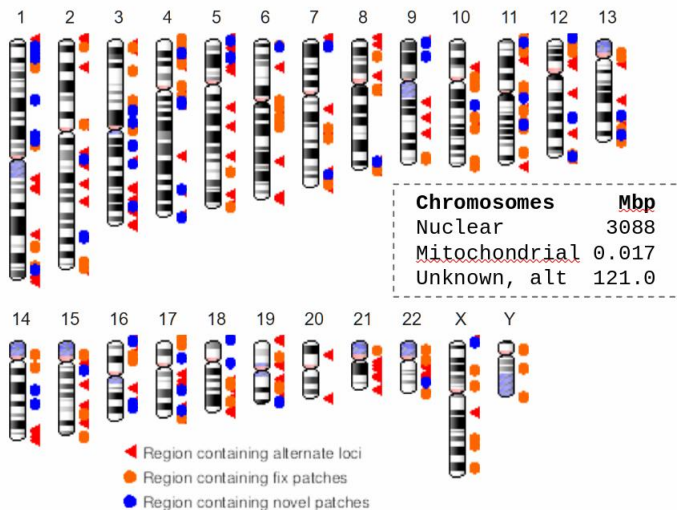


Женщины: XX, мужчины: XY
 Griffiths – *Introduction to Genetic Analysis*

Вопрос

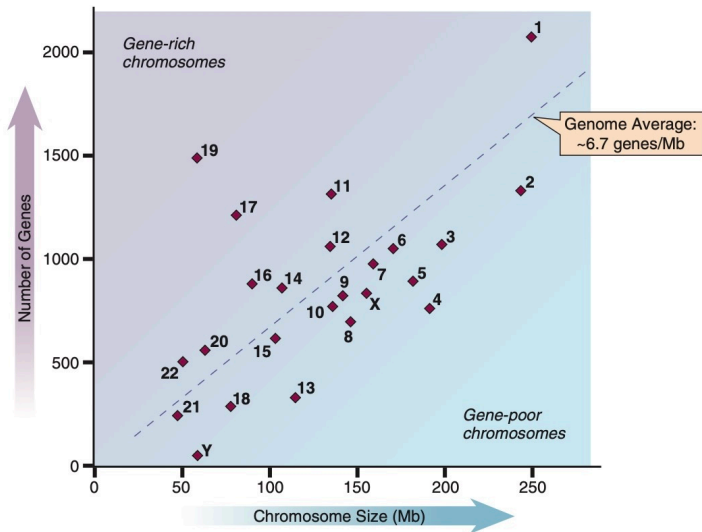
Как передаются хромосомы X и Y?

Референсный геном GRCh38.p13



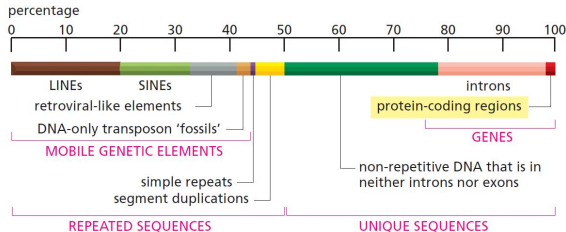
<https://www.ncbi.nlm.nih.gov/grc/human>

Размер и число генов на хромосомах человека



Thompson & Thompson

Состав генома человека



Alberts – *Essential Cell Biology*

Element	Transposition	Structure	Length	Copy number	Fraction of genome
LINEs	Autonomous	ORF1 ORF2 (pol) AAA	1–5 kb	20,000–40,000	21%
SINEs	Nonautonomous	AAA	100–300 bp	1,500,000	13%
DNA transposons	Autonomous	transposase	2–3 kb	300,000	3%
	Nonautonomous		80–3000 bp		

Griffiths – *Introduction to Genetic Analysis*

Состав генома человека

Участки	Длина, Мб	%	Описание
<u>Гены</u>	1,200	37.5	Участки генома, где происходит транскрипция
<u>Эзоны</u>	48	1.5	Транскрибируемые участки генома, которые остаются в составе РНК после сплайсинга
Другое (интроны, UTR)	1,152	36.0	Нетранслируемые участки кодирующей кДНК
<u>Уникальные и регуляторные последовательности</u>	510	15.9	
<u>Диспергированные повторы</u>	1,400	43.8	
LINE	640	20.0	~850,000 длинных диспергированных повторов (~7,000 п.н.). Ретротранспозоны, зачастую содержащие открытые рамки считывания, кодирующие машинерию обратной транскрипции. Часто находятся в неактивном состоянии
SINE, Alu повторы	420	13.1	~1.5 млн коротких диспергированных повторов (<500 п.н.). Ретротранспозоны, содержащие тРНК, мРНК, и рРНК, транспозиция которых зависит от других мобильных элементов
LTR ретротранспозоны	250	7.8	Мобильные элементы, содержащие длинные терминальные повторы (LTR), непосредственно фланкирующие внутреннюю кодирующую область
ДНК транспозоны	90	2.8	Мобильные элементы II класса, передвижение которых происходит за счет ДНК-интермедиатов
<u>Микросателлиты</u>	90	2.8	Участки в геномной последовательности, содержащие тандемные повторы длиной 2–10 п.н.
Всего	3,200	100	

Аннотация генов ENSEMBL GRCh38 v.99

Тип гена	Гены (тран- скрипты)	%	Описание
Белок-кодирующие гены	19,968 (153,197)	32.9	Гены, содержащие открытую рамку считывания
Псевдогены	15,263	25.2	Гены, являющиеся гомологами к известным белок-кодирующим генам, однако содержащие инделы со сдвигом рамки считывания и/или стоп-кодона, нарушающие открытую рамку считывания
Неподтвержденные	1,060	1.7	Необходима экспериментальная валидация
Рецепторы Т-клеток, иммуноглобулины	408	0.7	Подвергаются соматической рекомбинации перед транскрипцией
Гены РНК	23,977	39.5	
днкРНК	16,880		Некодирующие гены длиной > 200 п.н.
мяРНК	1,910		Процессинг предшественников мРНК
микроРНК	1,879		Маленькие РНК (~22 п.н.), ингибирующие экспрессию целевой мРНК
мякРНК	942		Пост-транскрипционная модификация других типов РНК
Другое	2,366		рРНК, тРНК, другие виды РНК
Всего	60,676 (227,818)		

Аннотация генов ENSEMBL GRCh38 v.99 🏠

Хромосома	Примерная длина, т.п.н.	Белок- кодирующие гены	РНК- кодирующие гены	Псевдогены
1	248,956	2,047	1,964	1,233
2	242,193	1,303	1,605	1,033
3	198,295	1,075	1,160	768
4	190,214	753	984	732
5	181,538	881	1,200	710
6	170,805	1,041	989	803
7	159,345	989	977	893
8	145,138	670	1,041	629
9	138,394	778	786	678
10	133,797	728	880	568
11	135,086	1,312	1,053	815
12	133,275	1,036	1,197	627
13	114,364	321	586	378
14	107,043	820	857	519
15	101,991	613	986	513
16	90,338	867	1,033	467
17	83,257	1,185	1,198	531
18	80,373	269	608	246
19	58,617	1,474	895	514
20	64,444	543	594	250
21	46,709	231	403	183
22	50,818	492	513	332
X	156,040	843	640	872
Y	57,227	63	108	392
MT	16.6	13	24	

Jackson (2018) *Essays Biochem*

Диплоидный геном человека HG002 // Hansen (2026) *Cell*

- T2T, кроме 38.8 млн п.н. (0.65%) участков рибосомной ДНК
- Точность Q68.9 (одна ошибка на 8 млн п.н.) для 99.4% генома
- Гаплотип-специфичные гены
- Нефункциональные гены на материнском и отцовском гаплотипах: 129 и 121

Table 1. Number of genes annotated in T2T-HG002v1.1 and T2T-CHM13v2.0

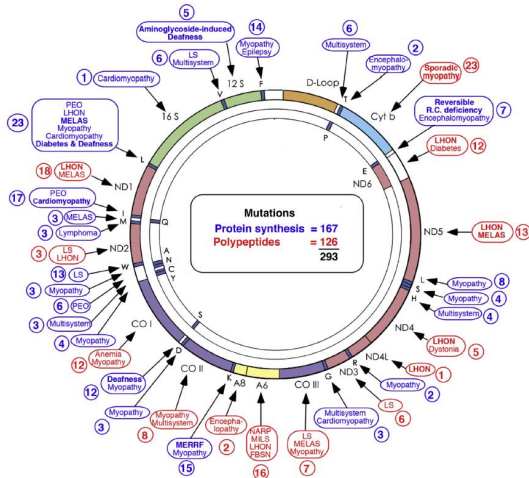
Gene type	HG002 MAT 1–22	HG002 PAT 1–22	HG002 MAT ChrX	HG002 PAT ChrY	CHM13 All Chrs
Protein coding	19,117	19,086	839	104	20,008
Non-coding (incl. lncRNA)	21,809	21,730	560	175	22,474
Pseudogene	16,072	16,031	926	415	17,281
Total	56,998	56,847	2,325	694	59,763

Gene biotypes were identified using the attribute "gene_biotype" from the T2T-CHM13 reference annotation. The T2T-CHM13 counts include all autosomes as well as both ChrX and ChrY.

DOI:10.1016/j.cell.2026.06.016

-
- The diagram illustrates the circular mitochondrial genome of *Drosophila melanogaster*. The H strand (heavy strand) is represented by the outer circle, and the L strand (light strand) is the inner circle. The 7S DNA is shown as a small segment. The origin of replication for the H strand (O_H) is indicated by a blue arrow, and for the L strand (O_L) by a blue arrow. The promoters for the H strand (P_H) and L strand (P_L) are indicated by green arrows. The genes are color-coded: rRNA genes are pink, genes encoding proteins are blue, and tRNA genes are red. The genes shown include: Thr, Pro, Val, Leu, Ile, Met, ND1, Gln, Trp, ND2, Ala, Asn, Cys, Tyr, Ser, CO1, Asp, CO2, Lys, ATP6, ATP8, CO3, Gly, Arg, ND3, ND4L, ND4, Leu, Ser, His, ND5, Glu, ND6, and Cyb.

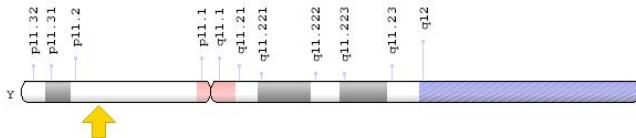
Болезнетворные мутации в мтДНК 🏠



DiMauro (2017) – *Mitochondrial Encephalomyopathies*

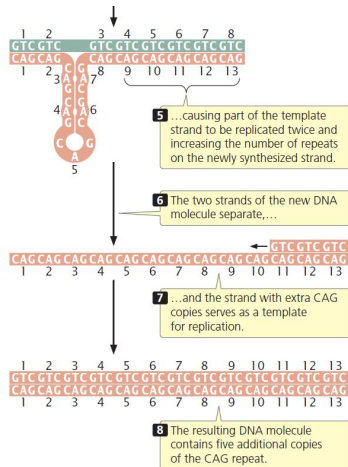
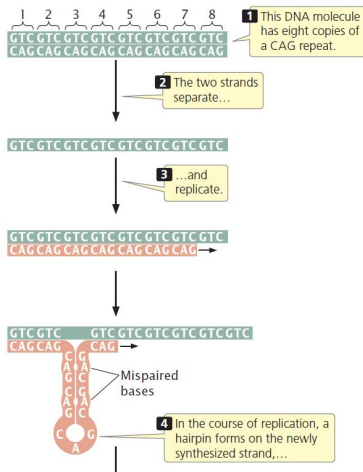
A propos: как узнать пол образца с помощью NGS?

- Соотношение гетерозигот/гомозигот на X хромосоме (Het/Hom):
 $\text{Het/Hom} < 0.8 \Rightarrow \text{М}$, $\text{Het/Hom} > 0.8 \Rightarrow \text{Ж}$; неопределенность: 0.5–1.0
- Ген **SRY** (Sex-determining Region Y) – не содержит интронов, кодирует фактор развития семенников, который принимает участие в развитии организма по мужскому типу у млекопитающих
- Ген амелогенина человека: **AMELX** и **AMELY**. Последовательности коротких плеч X и Y хромосом схожи на 84%. В судебно-медицинской практике при определении половой принадлежности используется вставка/делеция 6 п. н. в первом интроне генов **AMELX** и **AMELY** (Tzvetkov (2010) *Pharmacogenomics*)



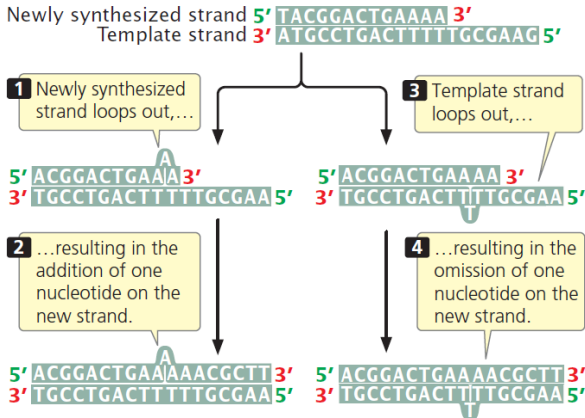
~
Точность репликации

Увеличение числа повторов при репликации



Pierce – *Genetics Essentials. Concepts and Connections*

Увеличение числа повторов при репликации



13.13 Insertions and deletions may result from strand slippage.

Pierce – *Genetics Essentials. Concepts and Connections*

Увеличение числа повторов и связанные с ним заболевания

Заболевание	Повтор	Количество копий в повторе	
		Норма	Заболевание
Спинальная и бульбарная мышечная атрофия (болезнь Кеннеди)	CAG	11–33	40–62
Синдром ломкой X-хромосомы	CGG	6–54	50–1500
Синдром Якобсона	CGG	11	100–1000
Спиноцеребеллярная атаксия (несколько видов)	CAG	4–44	21–130
Аутосомно доминантная церебральная атаксия	CAG	7–19	37–220
Миотоническая дистрофия	CTG	5–37	44–3000
Болезнь Хантингтона	CAG	9–37	37–121
Атаксия Фридрейха	GAA	6–29	200–900
Дентаторубро-паллидолюисова атрофия	CAG	7–25	49–75
Болезнь Унферрихта-Лундборга	CCCCGCCCGCG	2–3	12–13

Pierce – *Genetics Essentials. Concepts and Connections*

Задача

Найдите гены, соответствующие заболеваниям, в базе OMIM.



OMIM®

Online Mendelian Inheritance in Man®

An Online Catalog of Human Genes and Genetic Disorders

Updated February 12, 2021

Search OMIM for clinical features, phenotypes, genes, and r



Dissected OMIM Morbid Map Scorecard (Updated February 12th, 2021) :

Class of phenotype	Phenotype	Gene *
Single gene disorders and traits	5,740	4,006
Susceptibility to complex disease or infection	694	499
"Nondiseases"	151	119
Somatic cell genetic disease	231	130
*Some genes may be counted more than once because mutations in a gene may cause more than one phenotype and the phenotypes may be of different classes (e.g., activating somatic BRAF mutation underlying cancer, 164757.0001 . and germline BRAF mutation in Noonan syndrome, 164757.0022 .)		

Определения

Мутации – случайные изменения в последовательностях ДНК.

Мутации являются причиной всей генетической изменчивости человека и генетических заболеваний.

Механизмы мутаций:

- Спонтанные ошибки репликации
- Эндогенные (спонтанные) повреждения ДНК: дезаминирование, депуринизация
- Экзогенные (индуцированные) повреждения ДНК: радиация, воздействие химическими соединениями

Варианты :=

- мутации (недавние изменения)
- полиморфизмы (закрепились в популяции)
- целенаправленно вызванные изменения

Варианты в геноме: определения

Однонуклеотидные варианты (ОНВ): изменение одного нуклеотида ДНК (90%):

- Транзиции ($G \rightarrow A$, $C \rightarrow T$)
- Трансверсии ($G \rightarrow G$, $A \rightarrow T$)

Короткие делеции: удаление нескольких (< 50 п. н.) нуклеотидов (6%):

- Делеции уникальных последовательностей
- Сокращение количества коротких повторов

Короткие инсерции: вставка нескольких (< 50 п. н.) нуклеотидов (2%):

- Вставка уникальных последовательностей
- Увеличение количества коротких повторов

Структурные варианты: изменение последовательности ~ 1 кб и более (2%):

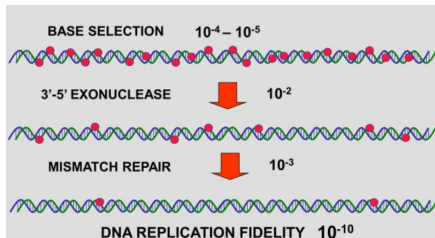
- Сбалансированные: инверсии и транслокации
- Несбалансированные (**варианты числа копий, CNV**): тандемные и диспергированные повторы, вставки, удаления (делеции)

Анеуплоидия: неверное количество целых хромосом – нуллисомия, моносомия, трисомия

Механизмы обеспечения точности репликации

Частота возникновения точечных мутаций:
 10^{-10} на один нуклеотид за одну репликацию

1. **ДНК-полимераза:** частота ошибки $\sim 10^{-5}$
2. **Система прюффридинга 3'-5'** убирает 99% ошибок, $\Rightarrow$ ее частота ошибки $\sim 10^{-2}$
3. **Комплекс репарации** убирает и восстанавливает ДНК-фрагменты около ошибочно спаренных нуклеотидов; частота ошибки $\sim 10^{-3}$



Fijalkowska (2012) *FEMS Microbiology Rev*

Механизмы обеспечения точности репликации

Частота возникновения точечных мутаций:
 10^{-10} на один нуклеотид за одну репликацию

TABLE 6–1 ERROR RATES

US Postal Service on-time delivery of local first-class mail	13 late deliveries per 100 parcels
Airline luggage system	1 lost bag per 200
A professional typist typing at 120 words per minute	1 mistake per 250 characters
Driving a car in the United States	1 death per 10^4 people per year
DNA replication (without mismatch repair)	1 mistake per 10^7 nucleotides copied
DNA replication (including mismatch repair)	1 mistake per 10^9 nucleotides copied

Alberts – *Essential Cell Biology*