

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

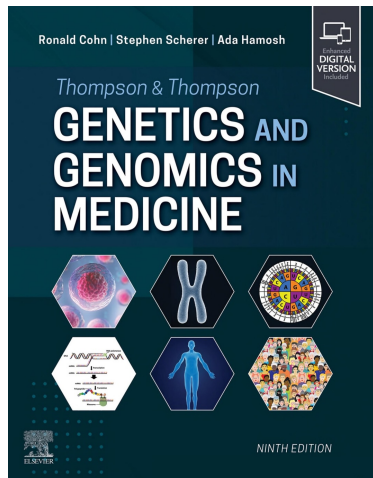
- Медицинская геномика $\neq$ Медицинская генетика
- = молекулярная генетика + популяционная генетика + медицинская генетика + статистическая генетика + генетическая эпидемиология + биоинформатика $\Rightarrow$ Нет единого учебника (?)
- За рамками курса: иммунология, патогены, микробиом, терапия, редактирование 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



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

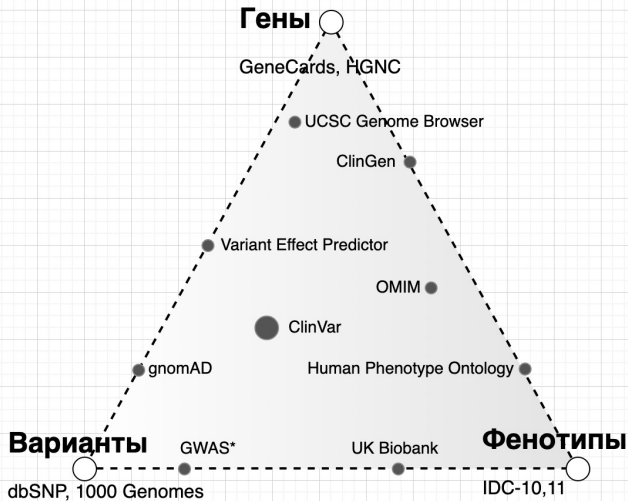
3. Комплексные (мультифакторные)

- 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

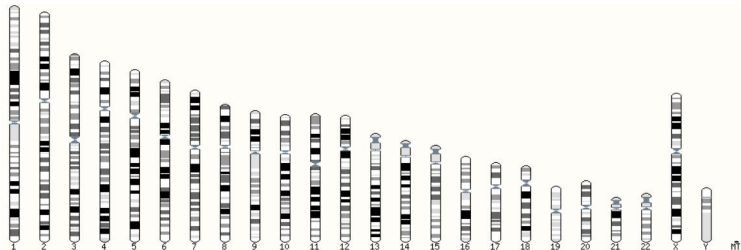
Кариотип человека 🏠



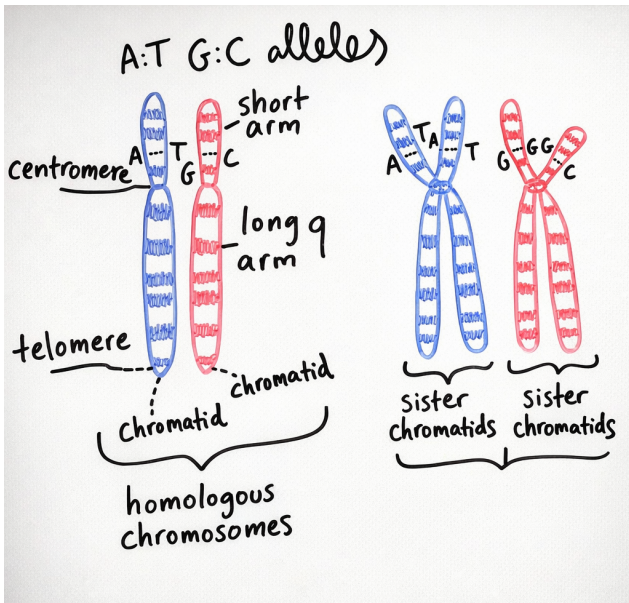
22 пары аутомсомных хромосом + 2 половые хромосомы

Кариотип человека 🏠

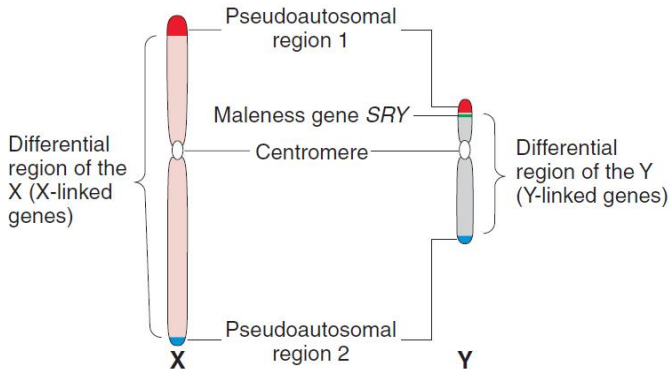
- **Эухроматин** (2.9 Гб) – участки ядерного генома, содержащие большое количество генов и являющиеся транскрипционно активными
- **Гетерохроматин** (0.2 Гб) – плотно уложенная (компактная) ДНК, транскрипционно неактивная, содержащая большое количество повторов. Находится в центромерах и теломерах
- **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

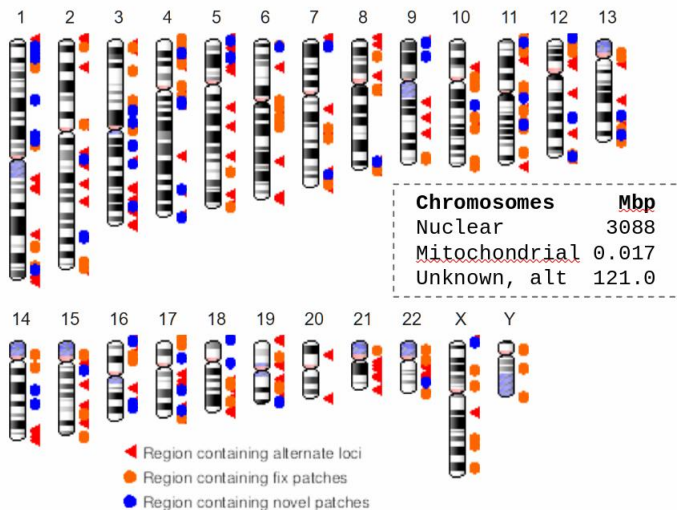


Женщины: XX, мужчины: XY
 Griffiths – *Introduction to Genetic Analysis*

Вопрос

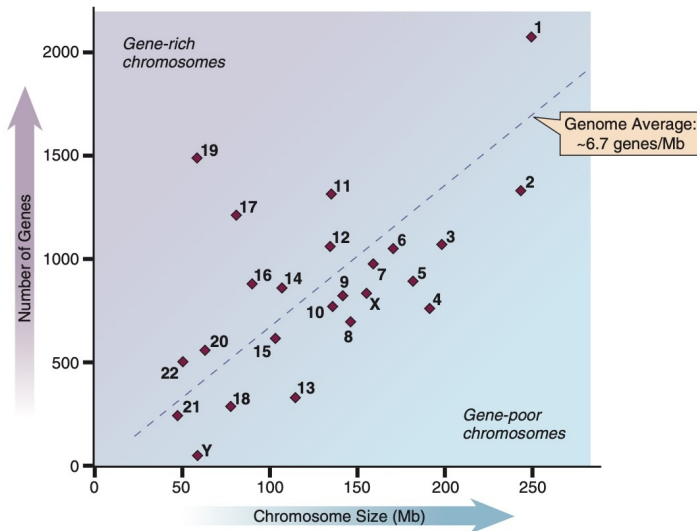
Как передаются хромосомы X и Y?

Референсный геном GRCh38.p13



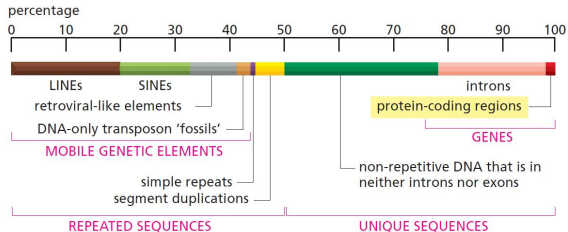
<https://www.ncbi.nlm.nih.gov/grc/human>

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

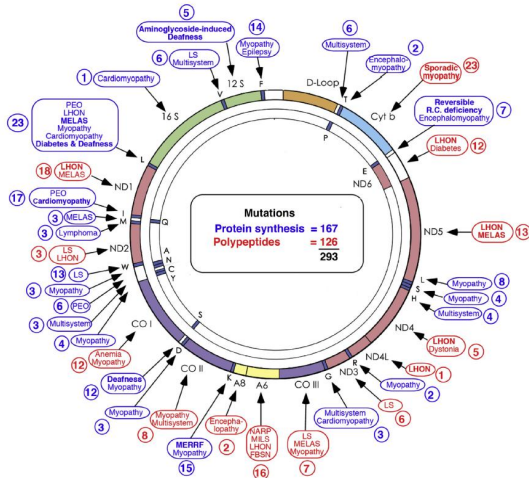
Table 1. Number of genes annotated in T2T-HG002v1.1 and T2T-CHM13v2.0

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.

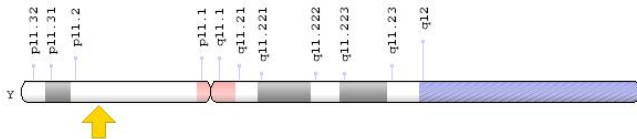
DOI:10.1016/j.cell.2026.06.016

-
- A circular map of the mitochondrial genome of *Drosophila melanogaster*. The map is divided into two main regions: the H strand (top half, light blue) and the L strand (bottom half, light orange). The H strand contains genes: *CYB*, *Thr*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, and *Thr*. The L strand contains genes: *7S DNA*, *0_H*, *P_H*, *P_L*, *0_L*, *ND1*, *Ile*, *fMet*, *Met*, *Val*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*, *Leu*, *Ser*, *His*, *ND4*, *ND4L*, *Arg*, *ND3*, *Gly*, *CO3*, *ATP6*, *ATP8*, *Lys*, *CO2*, *Asp*, *CO1*, *Ser*, *Trp*, *ND2*, *Ala*, *Asn*, *Cys*, *Tyr*, *Gln*, *Leu*, *Val*, *L2S*, *L6S*, *Pro*, *Glu*, *ND6*, *ND5*

DiMauro (2017) – *Mitochondrial Encephalomyopathies*

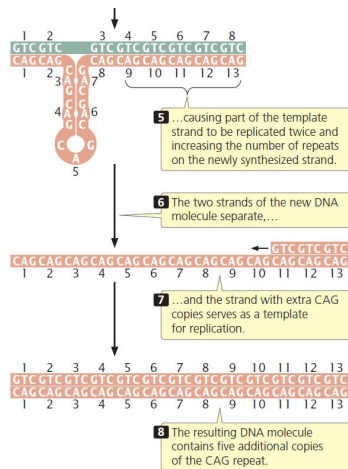
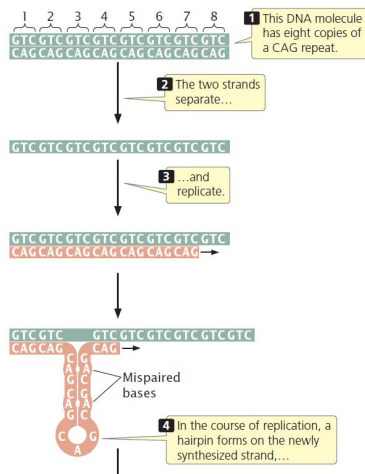
A propos: как узнать пол образца с помощью NGS?

- 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*)



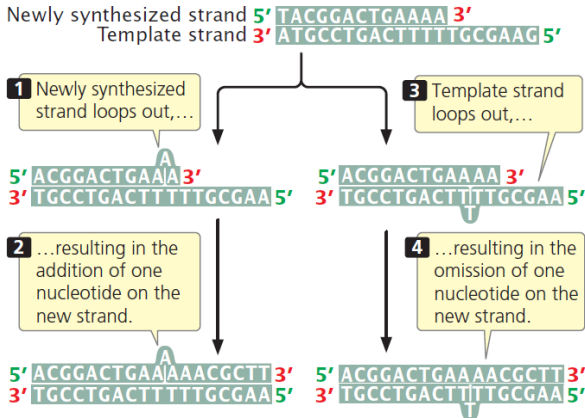
~
Точность репликации

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%):

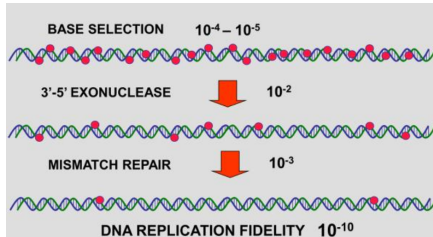
- Сбалансированные: инверсии и транслокации
- Несбалансированные (**варианты числа копий, CNV**): тандемные и диспергированные повторы, вставки, удаления (делеции)

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*

Мутации *de novo*

Мутации *de novo* (ДНМ)

Мутации *de novo* (ДНМ): наблюдаемые у ребенка, но не у родителей; возникают в клетках зародышевой линии (germline); могут быть выявлены путем секвенирования тройки мать-отец-ребенок.

Количество *de novo* ОНВ: ~40-80 у каждого новорожденного.

Изменчивость частоты ДНМ: по всему геному; в семьях; в кластерах мутаций (в одном индивидууме) и в «горячих точках» мутагенеза.

Факторы, влияющие на изменение частоты ДНМ:

- состав последовательности и функциональный контекст
- время (timing) репликации: рано/поздно
- транскрипционная активность и состояние хроматина
- количество митотических делений клетки (возраст отца)
- влияние повреждающих факторов
- эффективность механизмов репарации ДНК
- количество времени между митотическими делениями

Частоты возникновения и количество ДНМ

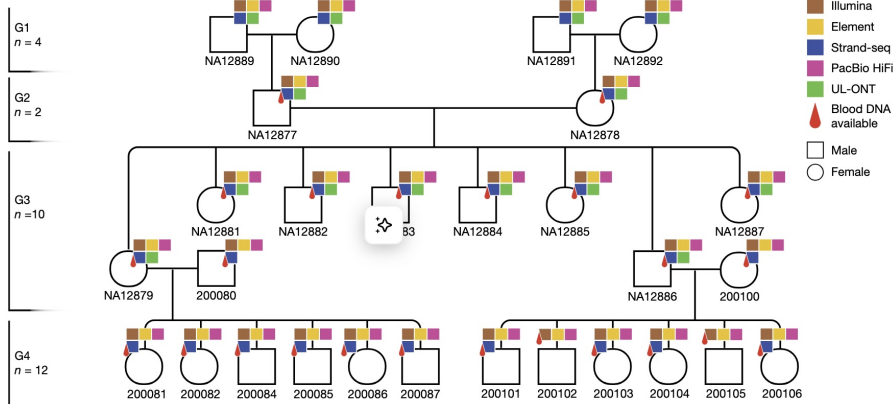
Частота возникновения точечных мутаций:
 10^{-10} на один нуклеотид за одну репликацию

Число ДНМ: $1.20 \cdot 10^{-8}$ на п.н. $\times 3.2 \cdot 10^9$ (длина генома) $\times 2$ (диплоид) ≈ 77

Тип	Частота на поколение	Число
ОНВ	1.20×10^{-8} на п.н. $1.66 \times 10^{-8} \ddagger$	44 - 82
Повторы динуклеотидов	1.20×10^{-4} на локус	-
Кодирующие ОНВ	-	1-2
Короткие инделы (<50 п.н.)	$0.53 - 1.59 \times 10^{-9}$ на п.н. $1.26 \times 10^{-8} \ddagger$	3-9
Длинные инделы	0.16	0.16
Варианты числа копий	$10^{-6} - 10^{-4}$ на локус	0.0154

$\ddagger$ Palamara (2015) *AJHG*

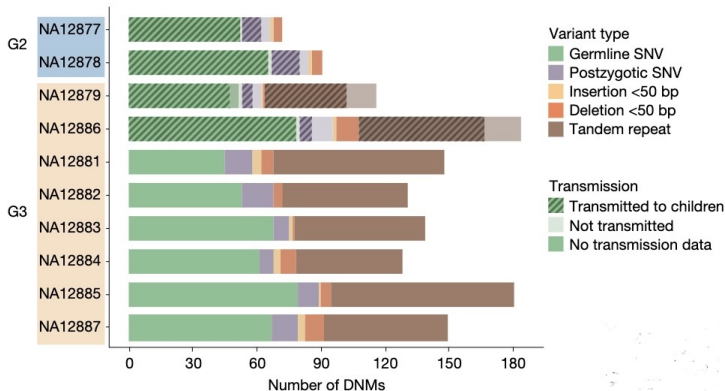
Выявление ДНМ в родословной CEPH1463



Porubsky (2025) *Nature* doi:10.1038/s41586-025-08922-2

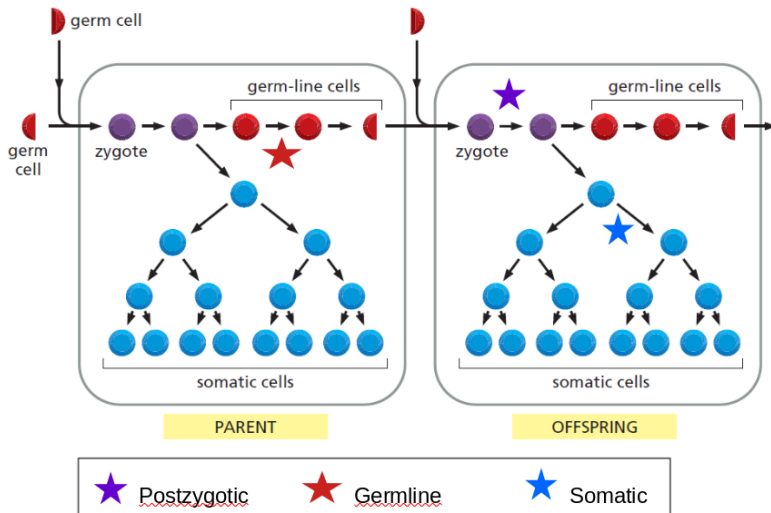
Количество ДНМ

- 98–206 ДНМ на поколение = 74.5 ОНВ + 7.4 инделов вне повторов + 65.3 инделов или структурных вариантов в повторах + 4.4 ДНМ в центромерах
- Мужчины: 12.4 ДНМ на Y-хромосоме

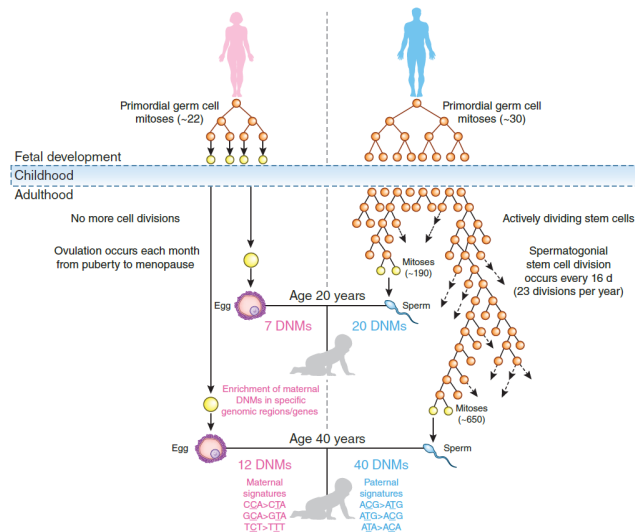


Porubsky (2025) *Nature* doi:10.1038/s41586-025-08922-2

Время появления ДНМ и мозаицизм



Происхождение ДНМ: материнские и отцовские

Goriely (2016) *Nat Genet*

Происхождение ДНМ: материнские и отцовские

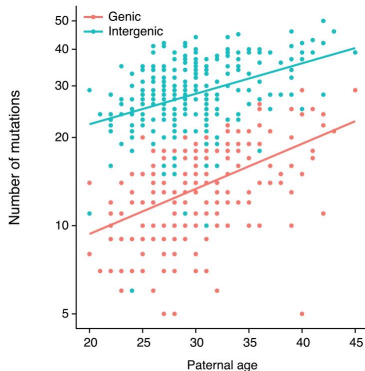
Материнские и отцовские

+1-3 ДНМ на один год возраста отца на момент зачатия.

+0.24 ДНМ на один год возраста матери на момент зачатия (повреждение ДНК, произошедшее не во время репликации).

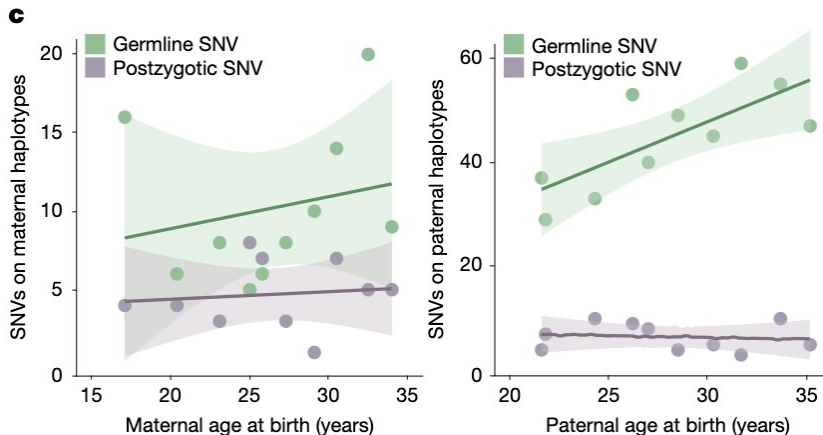
~ 80% всех ДНМ приходят от отца ребенка.

Эффект обладает большой изменчивостью между семьями.



Francioli (2015) *Nat Gene*

Происхождение ДНМ: материнские и отцовские



Porubsky (2025) *Nature* doi:10.1038/s41586-025-08922-2

Спектр ДНМ

Транзиции:

- C:G>T:A, дезаминирование 5-метил-C и C: 40%
- T:A>C:G, причина неизвестна: 25%

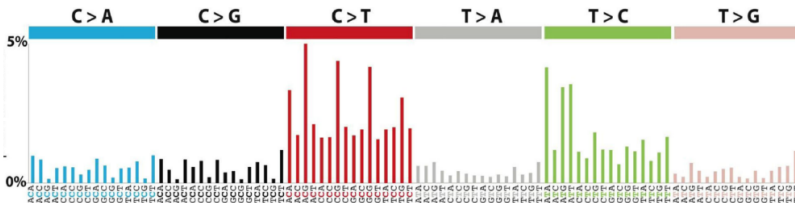
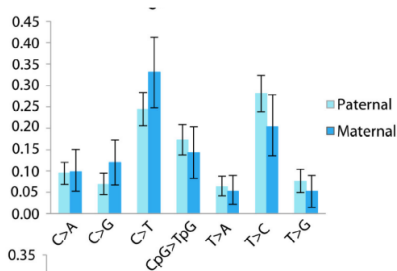
CpG составляют $\sim 1\%$ генома, так что и в не-CpG участках; однако, транзиции в CpG происходят в $\sim 18x$ раз чаще чем в не-CpG участках.

Трансверсии:

- G:C>T:A: 10%
- G:C>C:G: 10%
- A:T>C:G: $< 8\%$
- A:T>T:A: $< 8\%$

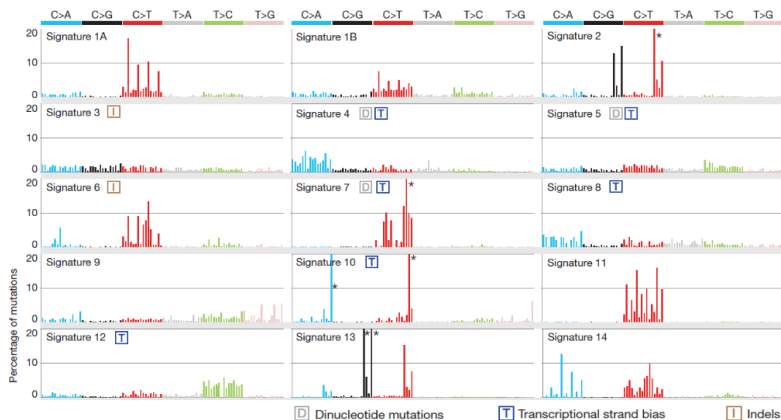
Трансверсии происходят в $\sim 1.5x$ раз чаще в CpG сайтах.

Ohno (2019) *Genes & Genet Systems*



6,570 ДНМ из 109 трио
Rahbari et al. (2016) *Nat Genet*

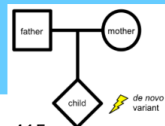
Спектр ДНМ



Подтвержденные мутационные сигнатуры образцов рака человека. Каждая сигнатура представлена в соответствии с классификацией на 96 замен, определенных классом самой замены и контекстом последовательности вокруг мутировавшего основания.

Alexandrov (2013) *Nature*

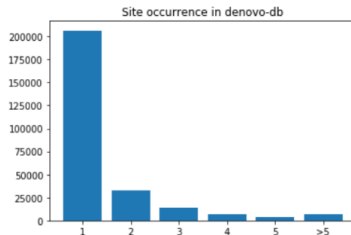
Denovo-db v.1.6.1



<i>Variant</i>	<i>Counts</i>	<i>%</i>
C>T	79,110	19.0
G>A	79,016	19.0
A>G	47,666	11.5
T>C	47,584	11.5
C>G	17,431	4.2
G>C	17,154	4.1
C>A	15,719	3.8
G>T	15,234	3.7
A>C	12,744	3.1
T>G	12,464	3.0
T>A	11,637	2.8
A>T	11,482	2.8
T>TA	739	0.2
CT>C	737	0.2
...		
<i>Total</i>	<i>415,515</i>	

Variants: 415,415
 Samples: 11,518
 Genes: 17,431
 Sites: 270,506

Missense: 20,815
 Pred. Damaging: ~75%



<http://denovo-db.gs.washington.edu/denovo-db/index.jsp>

ДНМ и заболевания человека

- ДНМ крайне редко поодиночке, но значимы в целом; являются причиной 60–75% всех спорадических случаев заболеваний
- Более болезнетворны, чем наследственные мутации, эффект зависит от времени появления
- **Тяжелые детские заболевания в популяциях без инбридинга:** спорадические случаи синдромов пороков развития (синдромы Шинца, Кабуки, Боринга-Опица), нарушения развития нервной системы (серьезные случаи умственной отсталости), пороки сердца.
- **Неврологические и психиатрические заболевания, развивающиеся во взрослом возрасте:** болезнь Паркинсона, боковой амиотрофический склероз, шизофрения, эпилепсия, синдром Туретта, расстройства аутистического спектра и биполярное расстройство.
Пример: в 10% случаев шизофрении наблюдались варианты числа копий *de novo*, против 1.3% у контролей
- **Наследственные формы рака:** синдром Ли-Фраумени (*TP53*), семейный аденоматозный полипоз, ~ 7% несоматических мутаций являются ДНМ.

Частота возникновения мутаций 🏠

S : mutation rate per nucleotide per cell division

K : the average number of germline cell divisions per generation, from zygote to zygote (~ 30 in females, ~ 60 – 500 in males)

N : genome size

Mutation rate per genome: $S \times K \times N$

$\sim 10^{-10}$ per nucleotide per cell division (or $\sim 10^{-8}$ per generation, because there are ~ 100 cell divisions and rounds of DNA replication per human generation $\Rightarrow \sim 100$ *de novo* mutations in a newborn

1) $\sim 1\%$ of all newborns being affected by a serious disease due to a *de novo* mutation. If the mutation rate were 100 times higher, 10^{-8} per cell division, we would immediately **go extinct**.

2) 10^{14} cells in human body $\Rightarrow$ total number of somatic mutations in each person ?

Crumbling Genome: The Impact of Deleterious Mutations on Humans, First Edition. Alexey S. Kondrashov.
© 2017 John Wiley & Sons, Inc. Published 2017 by John Wiley & Sons, Inc.

Частота возникновения мутаций 🏠

Genes Genet. Syst. (2019) **94**, p. 13–22

Spontaneous *de novo* germline mutations in humans and mice: rates, spectra, causes and consequences

Mizuki Ohno*

The human body consists of approximately 10^{14} cells and undergoes approximately 10^{16} cell divisions in a lifetime, resulting in **over 10^{15} cumulative mutations per individual** (Frank, 2014).

If 10^6 stem cells in intestinal tissue generate transient daughter cells once a week with a mutation rate of approximately 10^{-9} per nucleotide per cell division, the intestinal epithelium of a 60-year-old human would have accumulated more than 10^9 independent mutations. Thus, **nearly every genomic site is likely to be mutated in at least one cell in this organ** (Lynch, 2010a, 2010b).

Частота мутаций в болезнетворных генах 🏠

Заболевание	МIM	Кол-во мутаций на миллион гамет	Признаки и симптомы (фенотип)
Сцепленные с X-хромосомой			
Мышечная дистрофия Дюшена	310200	40-105	Мышечная атрофия
Гемофилия А	306700	30-60	Сильные нарушения свертываемости крови
Гемофилия В	306900	0.5-10	Легкие нарушения свертываемости крови
Аутосомно-доминантные			
Ахондроплазия	100800	10	Очень низкий рост
Аниридия	106200	2.6	Отсутствие радужной оболочки глаза
Болезнь Гентингтона	143100	<1	Неконтролируемые движения, изменения характера
Синдром Марфана	154700	4-6	Длинные конечности, тонкие стенки сосудов
Нейрофиброматоз 1 типа	162200	40-100	Коричневые пятна на коже, доброкачественные опухоли под кожей
Несовершенный остеогенез	166200	10	Ломкие кости
Поликистоз почек	600666	60-120	Доброкачественные образования в почках
Ретинобластома	180200	5-12	Злокачественная опухоль сетчатки

Lewis – *Human genetics. Concepts and applications*

Задача

Найдите гены заболеваний в OMIM; объясните разницу в количестве ДНМ

Репарация ДНК

Заболевания, вызванные дефектами в системе репарации ДНК 🏠

Заболевание	Симптомы	Генетические дефекты
Ксеродерма пигментная	Кожа в пятнах, похожих на веснушки, чувствительность к солнечному свету, предрасположенность к раку кожи	Дефекты эксцизионная репарация нуклеотидов
Синдром Коккейна	Карликовость, чувствительность к солнечному свету, преждевременное старение, глухота, умственная отсталость	Дефекты эксцизионная репарация нуклеотидов
Трихотиодистрофия	Ломкие волосы, кожные аномалии, низкий рост, характерные черты лица	Дефекты эксцизионная репарация нуклеотидов
Наследственный непололипозный рак толстой кишки	Предрасположенность к раку толстой кишки	Дефекты в репарации ошибочно спаренных нуклеотидов
Анемия Фанкони	Повышенная пигментация кожи, аномалии костей, сердца и почек, предрасположенность к лейкемии	Предположительно развивается из-за дефектов в репарации кросслинкинга
Атаксия-телеангиэктазия	Дефекты в координации, суженные кровеносные сосуды в коже и глазах, иммунодефициты, чувствительность к ионизирующей радиации	Дефекты в обнаружении повреждений ДНК и ответе на них
Синдром Фраумени	Ли-Предрасположенность к раку множества различных тканей	Дефекты в ответе на повреждение ДНК

Pierce – *Genetics Essentials. Concepts and Connections*

Задача

Найдите гены, соответствующие заболеваниям, в базе OMIM.

Заболевания, вызванные дефектами в системе репарации ДНК 🏠

Location	Phenotype	Phenotype MIM number	Inheritance	Phenotype mapping key	Gene/Locus	Gene/Locus MIM number
17p13.1	Li-Fraumeni syndrome	151623	<u>AD</u>	<u>3</u>	TP53	191170

- **Синдром Ли-Фраумени** – редкое аутосомно-доминантное наследственное заболевание, носители которого имеют predisposition к развитию рака.
- Риск развития любого инвазивного рака (кроме рака кожи): ~ 50% к 30 годам (против 1% у населения в целом) и ~ 90% к 70 годам.
- Синдром связан с герминальной мутацией в **гене-онкосупрессоре *TP53***, кодирующем транскрипционный фактор P53, который в норме участвует в контроле деления и роста клеток. *TP53* обычно экспрессируется при наличии клеточного стресса, например, при повреждении ДНК, таким образом останавливая клеточный цикл для участия в репарации ДНК (если это возможно) или вызове апоптоза в случае необратимого повреждения ДНК.

Заболевания, вызванные дефектами в системе репарации ДНК 🏠

Пигментная ксеродермия – редкое аутосомно-рецессивное заболевание, включающее в себя **аномальную пигментацию кожи и высокую чувствительность к солнечному свету**. Люди с этим заболеванием также имеют **сильную предрасположенность к раку кожи**, с частотой встречаемости в 1,000-2,000 раз больше, чем у здоровых людей.

Клетки большинства больных пигментной ксеродермией имеют дефекты эксцизионной репарации нуклеотидов $\Rightarrow$ множество пиримидиновых димеров (появившихся из-за УФ-излучения солнца) остаются неисправленными, вследствие чего могут привести к раку.



Pierce – *Genetics Essentials. Concepts and Connections*

Структурные варианты

Структурные варианты

Структурные варианты (хромосомные перестройки): изменения последовательности $> 1\text{кб}$ длиной.

- Сбалансированные: инверсии и транслокации
- Несбалансированные (**CNV, варианты числа копий**): Тандемные и диспергированные дупликации, делеции, инсерции

Механизмы:

- Рекомбинация: неаллельная гомологичная рекомбинация (NAHR)
- Нерепликативные: репарация негомологично соединенных концов (NHEJ)
- Репликативные:
 - Остановка вилки репликации и изменение матрицы (FoSTeS)
 - Опосредованная микрогомологией и индуцированная разрывами репликация (MMBIR)
- Ретротранспозиции (LINE1, Alu-повторы)

Несбалансированные структурные варианты (CNV)

Tandem duplication



Variable Number of Tandem Repeats



Dispersed duplication



Несбалансированные структурные варианты (CNV)

Insertion

B C D E

B C G D E

Repeat insertion

D B D C E



D B D C D E

Deletion

B C D E

B E

Сбалансированные структурные варианты

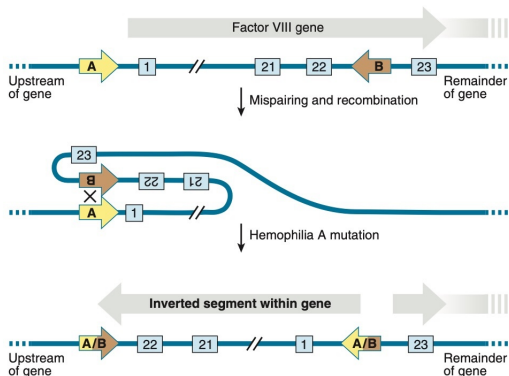
Inversion

Intra-chromosomal translocation (ITX)Inter-chromosomal translocation (CTX)



(а) Протоонкоген *MYC*: ТФ, участвует клеточном цикле, апоптозе и трансформации клеток. (б) Протоонкоген *ABL*: протеинкиназа, участвующая в сигнальном пути клеточной пролиферации. Гибридный белок BCR1-ABL обладает постоянной киназной активностью вне зависимости от наличия инициирующего сигнала.

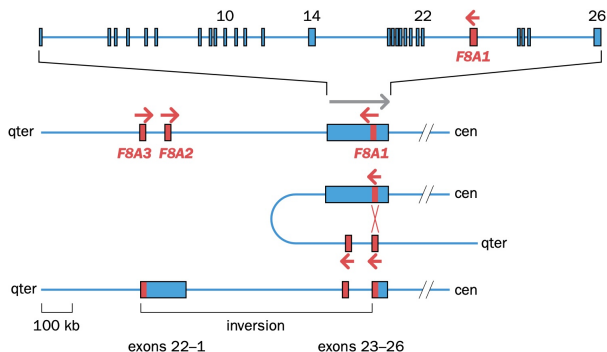
Пример инверсии части гена *F8* как причины гемофилии А 🏠



Thompson & Thompson

В ходе мейоза у этого участка X-хромосомы отсутствует гомологичный партнер. Повторы A и B могут спариваться друг с другом, образуя петлю. Кроссинговер между спаренными повторами *F8* приводит к инверсии сегмента размером 500 тыс. п.н. Каждый отдельный экзон и прилегающие к нему последовательности интронов остаются неповрежденными.

Пример инверсии части гена *F8* как причины гемофилии А 🏠



Strachan & Read

В ходе мейоза у этого участка X-хромосомы отсутствует гомологичный партнер. Повторы A и B могут спариваться друг с другом, образуя петлю. Кроссинговер между спаренными повторами *F8* приводит к инверсии сегмента размером 500 тыс. п.н. Каждый отдельный экзон и прилегающие к нему последовательности интронов остаются неповрежденными.

Анеуплоидии

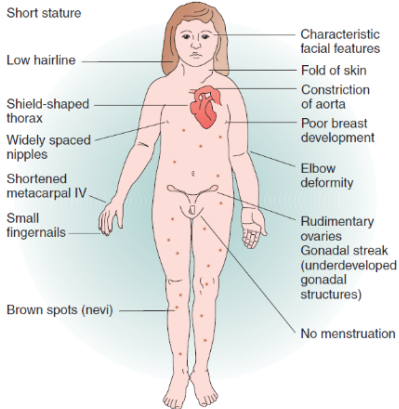
Анеуплоидия: неверное количество целых хромосом: нуллисомия, моносомия, трисомия и т.д.

- Результат ошибок при мейозе (гаметогенез) или митозе (соматическое деление при раннем эмбриональном развитии)
- Ошибки при митозе могут приводить к мозаицизму: появлению у эмбриона двух кариотипически различных линий клеток
- Основная причина спонтанных аборт ($\sim 30\%$ от всех зачатий)
- Выявляются у $\sim 0.3 - 0.6\%$ новорожденных

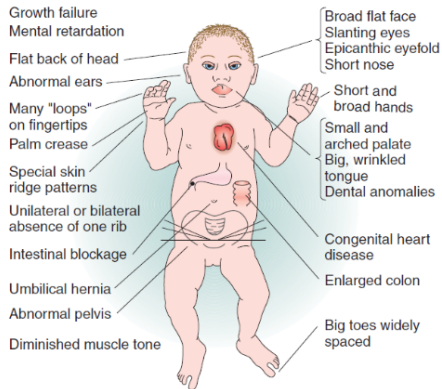
Название	Кариотип	Частота встречаемости
Синдром Тернера	X0 (Моносомия по X хромосоме у женщин)	1:2000-1:2500
Синдром Клайнфельтера	XXY (XXXY, XXXXY, XXXYY)	1:1000
Трисомия по X-хромосоме	XXX	1:1000
Синдром Дауна	Трисомия по 21 хромосоме	1:1100
Синдром Эдвардса	Трисомия по 18 хромосоме	1:6000
Синдром Патау	Трисомия по 13 хромосоме	1:7000-1:14000
Трисомия по 8 хромосоме	Трисомия по 8 хромосоме	1:25000-1:50000

Примеры

Characteristics of Turner syndrome (XO)



Characteristics of Down syndrome (trisomy 21)



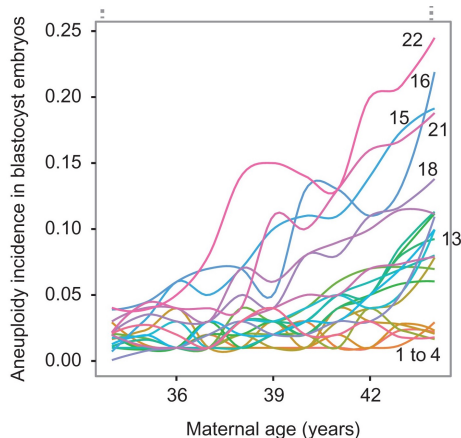
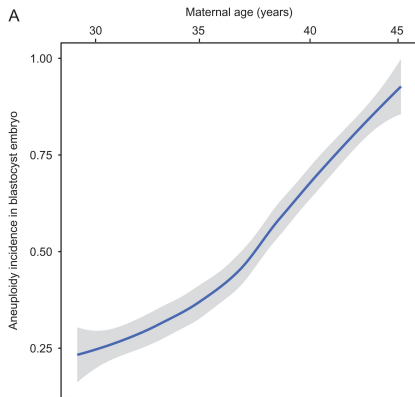
Griffiths– *Introduction to Genetic Analysis*

Примеры

Aneuploidy	Common name	Estimated incidence among life-births	Symptoms can include
Trisomy 13	Patau syndrome	Approximately 1:16000	Severe intellectual disability, heart defects, brain or spinal cord abnormalities, small or poorly developed eyes, extra fingers or toes, cleft lip and palate, weak muscle tone
Trisomy 18	Edwards syndrome	Approximately 1:5000	Intrauterine growth retardation, low birth weight, heart defects and abnormalities of other organs, small, abnormally shaped head, small jaw and mouth, clenched fists, severe intellectual disability
Trisomy 21	Down syndrome	Approximately 1:800	Mild to moderate intellectual disability, characteristic facial appearance, weak muscle tone, heart defects, digestive abnormalities, hypothyroidism, increased risk of hearing and vision problems, leukaemia, Alzheimer's disease
Trisomy X	Triple X syndrome	Approximately 1:1000	Increased height, increased risk of learning disabilities, delayed development of speech, language and motor skills, weak muscle tone, behavioural and emotional difficulties, seizures, kidney abnormalities
47,XXY		Approximately 1:1000	Increased height, increased risk of learning disabilities, delayed development of speech, language, and motor skills, weak muscle tone, hand tremors, seizures, asthma, scoliosis, behavioural and emotional difficulties
47,XXY	Klinefelter syndrome	1:500 to 1:1000	Small testes, low testosterone levels, delayed and incomplete puberty, breast enlargement, reduced facial and body hair, infertility, increased height, increased risk of breast cancer, learning disabilities, delayed speech and language development
48,XXXY		Approximately 1:18000 to 1:40000	Small testes, low testosterone levels, delayed and incomplete puberty, breast enlargement, reduced facial and body hair, infertility, increased height, tremors, dental problems, peripheral vascular disease, deep vein thrombosis, asthma, type 2 diabetes, seizures, heart defects, delayed speech and language development, learning disabilities
45,X	Turner syndrome	Approximately 1:2500	Short stature, early loss of ovarian function, infertility, absence of puberty, webbing of the neck, skeletal abnormalities, kidney problems, heart defects

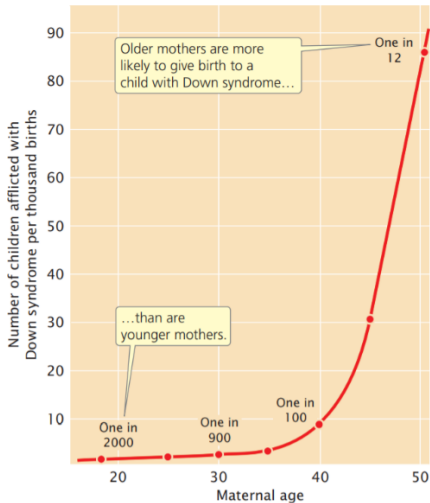
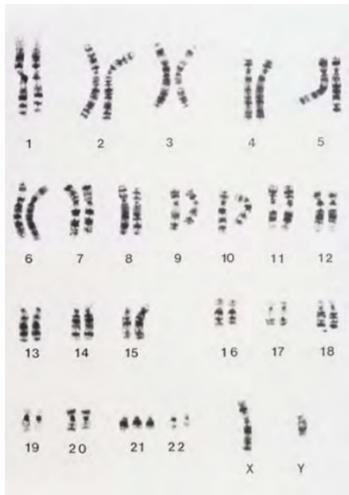
Jackson (2018) (*Essays in Biochemistry*)

Анеуплодия и возраст матери // Capalbo et al (2017) Human Repr Upd



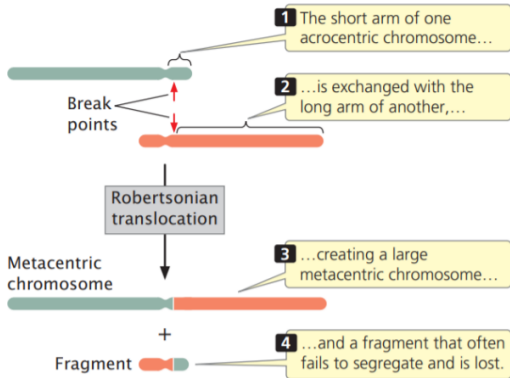
// Источник: Константин Юрьевич Боярский, *Эволюционная гинекология*
youtu.be/AK9g8DZLmc8

Синдром Дауна: трисомия по 21 хромосоме



Pierce – *Genetics Essentials. Concepts and Connections*

Робертсоновская транслокация



Pierce – *Genetics Essentials. Concepts and Connections*

Робертсоновская транслокация ($\sim 1/1000$ новорожденных): акроцентрические хромосомы (13, 14, 15, 21). Одна из причин ($\sim 1-3\%$) мужского бесплодия.

Инактивация X-хромосомы

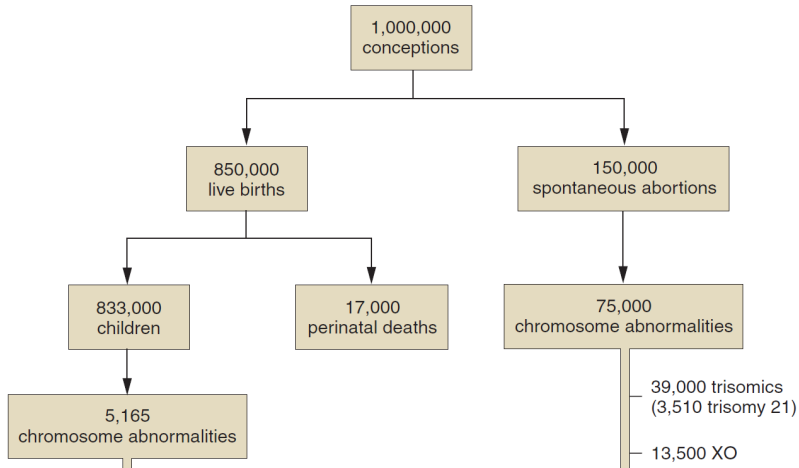
Инактивация X-хромосомы: одна из двух X-хромосом конденсирована в каждой клетке эмбриона женского пола.

- Происходит на ранних стадиях развития
- Случайна в разных клетках
- Сохраняется при последующих делениях клетки, но не между поколениями
- Ткани женского организма похожи на лоскутное одеяло, в котором в половине случаев инактивирована материнская, а в половине случаев — отцовская хромосома
- Инактивация X-хромосомы обеспечивает сравнимые уровни экспрессии $\sim 1,500$ генов, находящихся на X-хромосоме, у мужчин и женщин



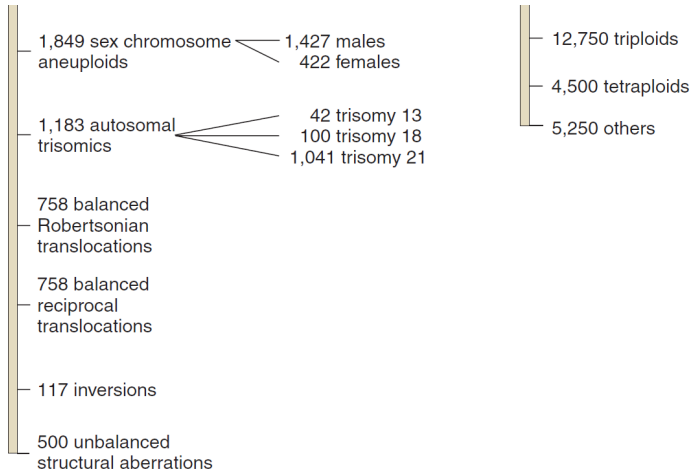
Выводы

Судьба 1 млн прижившихся человеческих зигот



Griffiths – *Introduction to Genetic Analysis*

Судьба 1 млн прижившихся человеческих зигот



Griffiths – *Introduction to Genetic Analysis*

Выводы

- Несколько механизмов обеспечивают высокую точность репликации ДНК, включая выбор нуклеотидов, проверку (proofreading) и репарацию неверно спаренных нуклеотидов.
- Несмотря на это, мутации появляются ввиду спонтанных ошибок репликации, а также действия эндогенных и экзогенных факторов повреждения ДНК
- Частота возникновения мутаций у человека является компромиссом между вымиранием и необходимостью эволюционных изменений
- Существуют различные виды мутаций *de novo*: однуклеотидные варианты, структурные варианты и анеуплоидии. Их частоты возникновения и последствия сильно различаются.

Список литературы

- Seplyarskiy VB, Sunyaev S (2021) The origin of human mutation in light of genomic data. *Nat Rev Genet* 1–15.
- Ohno, M. (2019). Spontaneous de novo germline mutations in humans and mice: rates, spectra, causes and consequences. *Genes Genet Syst* 94, 13–22.
- Acuna-Hidalgo, R., Veltman, J.A., and Hoischen, A. (2016). New insights into the generation and role of de novo mutations in health and disease. *Genome Biol* 17.
- Weckselblatt, B., and Rudd, M.K. (2015). Human structural variation: mechanisms of chromosome rearrangements. *Trends Genet* 31, 587–599.
- Zhang, F., Gu, W., Hurles, M.E., and Lupski, J.R. (2009). Copy Number Variation in Human Health, Disease, and Evolution. *Annu Rev Genomics Hum Genet* 10, 451–481.
- Jackson, M., Marks, L., May, G.H.W., and Wilson, J.B. (2018). The genetic basis of disease. *Essays Biochem* 62, 643–723
- Belyeu, J.R., Brand, H., et al. (2021). De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families. *Am J Hum Genet* 108, 597–607.