

Медицинская геномика

Василий Евгеньевич Раменский
Анастасия Александровна Жарикова и Мария Ильинична Зайченко

ramensky@gmail.com, azharikova89@gmail.com

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

2024

План курса

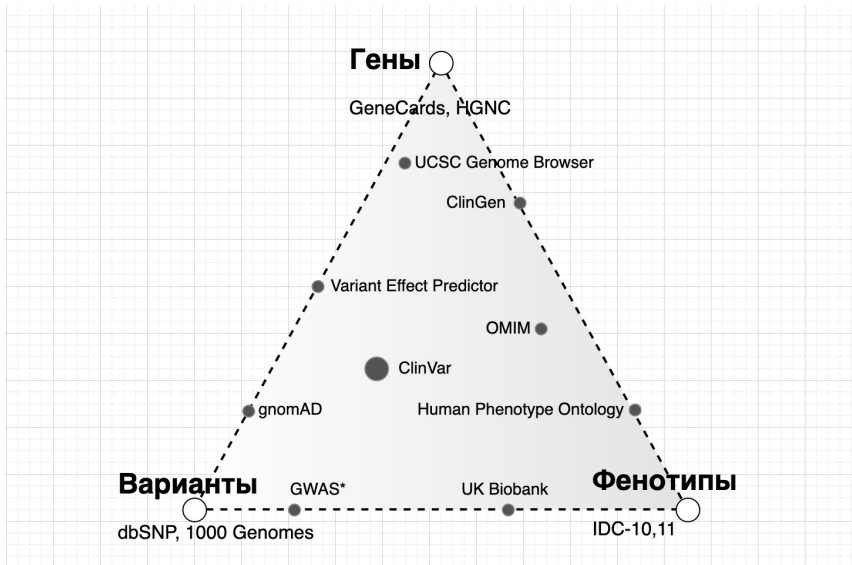
Часть 1: Кто виноват?

1. Механизмы и частота возникновения мутаций.
Структурные варианты генома и анеуплоидии
2. Передача мутаций
3. *Семинар*: фенотипы и болезни
4. Варианты во времени: некоторые основы популяционной генетики, ч.1
5. Варианты во времени: некоторые основы популяционной генетики, ч.2
6. Варианты в пространстве генов и их последствия
7. *Семинар*: гены и аннотация вариантов
8. Варианты в индивидуумах и популяциях

Часть 2: Что делать?

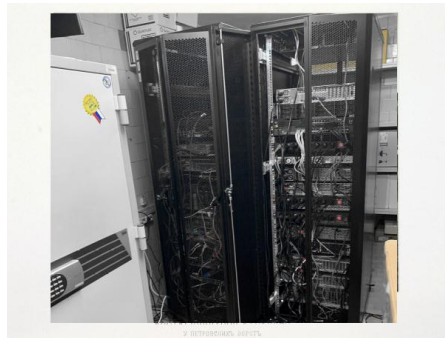
1. Моногенные (Менделевские) заболевания, ч.1: поиск генов
2. Моногенные (Менделевские) заболевания, ч.2: диагностика
3. *Семинар*: варианты в геноме человека
4. Некоторые основы количественной генетики
5. *Коллоквиум (контрольная работа)*
6. Комплексные (мультифакторные) заболевания

План курса



Медицинская геномика

- Обратная генетика: накопление большого количества геномных данных позволяет делать выводы о связи генотипа и фенотипа без экспериментального вмешательства.
- Данные [частично] доступны



Примечания

- Важная информация: *kodomo*
- Медицинская геномика = молекулярная генетика + популяционная генетика + медицинская генетика + статистическая генетика + генетическая эпидемиология + биоинформатика ⇒ Нет единого учебника
- Многие темы останутся незатронутыми, например, иммунология, патогены, микробиом, терапия, редактирование генома

- Определения
- Вопросы и задачи
- Слайды для самостоятельного изучения: 🏠
- Выводы, основные понятия, список литературы

1. T.Strachan, A.Read. Human Molecular Genetics. 2011. ISBN 0815345895.
2. J. Gillespie. Population genetics. A concise guide 1998 ISBN 0-8018-5764-6
3. A.Griffiths et al. An Introduction to Genetic Analysis. Freeman/Worth, 11 ed. 2015 ISBN 1464109486
4. Бочков Н.П., Пузырев В.П., Смирнихина С.А. Клиническая генетика. Учебник. Под ред. Н.П. Бочкова. ГЭОТАР-Медиа, 4-е издание, 2018. ISBN 978-5-9704-4628-7

Лекция 1

Механизмы и частота возникновения мутаций.
Структурные варианты генома и анеуплоидии

План лекции

- 1 Медицинская геномика: описание курса
- 2 Кариотип и состав генома человека
- 3 Митоз и репликация ДНК
- 4 Точность репликации и частота возникновения мутаций
- 5 Мутации *de novo*
- 6 Структурные варианты и варианты числа копий
- 7 Анеуплоидии

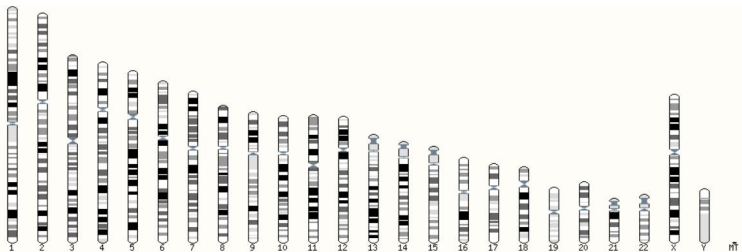
Кариотип человека 🏠



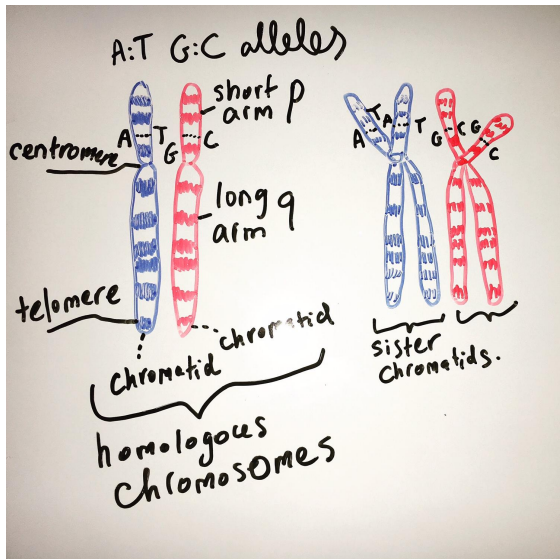
22 пары аутосомных хромосом + 2 половые хромосомы

Кариотип человека 🏠

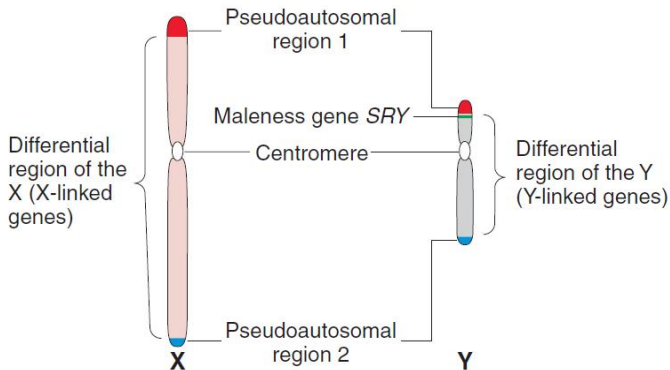
- **Эухроматин** (2.9 Гб) – участки ядерного генома, содержащие большое количество генов и являющиеся транскрипционно активными
- **Гетерохроматин** (0.2 Гб) – плотно уложенная (компактная) ДНК, транскрипционно неактивная, содержащая большое количество повторов. Находится в центромерах и теломерах
-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



Терминология 🏠



Половые хромосомы 🏠

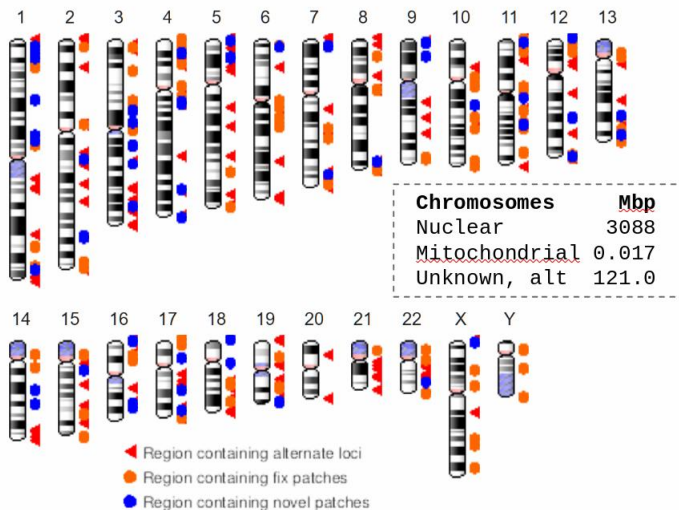


Женщины: XX, мужчины: XY
 (Griffiths – *Introduction to Genetic Analysis*)

Вопрос

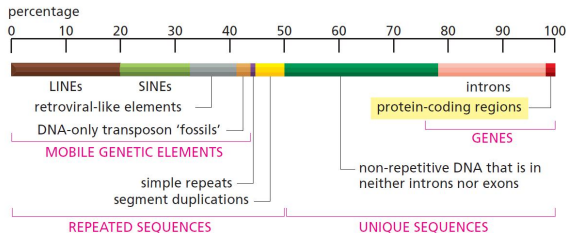
Как передается Y хромосома?

Референсный геном GRCh38.p13



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

Состав генома человека



Alberts – *Essential Cell Biology*

Element	Transposition	Structure	Length	Copy number	Fraction of genome
LINEs	Autonomous	ORF1 ORF2 (<i>pol</i>) 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	
<u>LINE</u>	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, мяРНК, 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-кодоны, 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)		

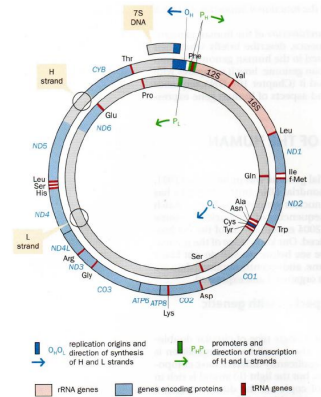
Аннотация генов ENSEMBL GRCh38 v.99

Хромосома	Примерная длина (п.н.)	Белок-кодирующие гены	РНК-кодирующие гены	Псевдогены
1	248956422	2047	1964	1233
2	242193529	1303	1605	1033
3	198295559	1075	1160	768
4	190214555	753	984	732
5	181538259	881	1200	710
6	170805979	1041	989	803
7	159345973	989	977	893
8	145138636	670	1041	629
9	138394717	778	786	678
10	133797422	728	880	568
11	135086622	1312	1053	815
12	133275309	1036	1197	627
13	114364328	321	586	378
14	107043718	820	857	519
15	101991189	613	986	513
16	90338345	867	1033	467
17	83257441	1185	1198	531
18	80373285	269	608	246
19	58617616	1474	895	514
20	64444167	543	594	250
21	46709983	231	403	183
22	50818468	492	513	332
X	156040895	843	640	872
Y	57227415	63	108	392
Митохондриальная	16569	13	24	

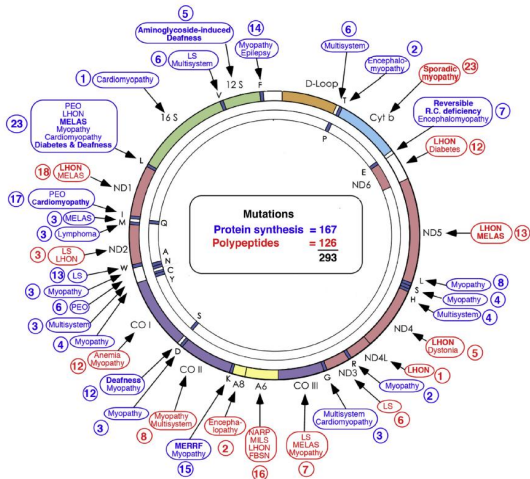
Jackson (2018) *Essays Biochem*

Митохондриальный геном

- **мтДНК:** кольцевая молекула ДНК длиной 16 659 п.н., состоящая из двух нитей; содержит H- и L-цепи; аналогична бактериальной ДНК
- Содержится только в яйцеклетках, наследуется от матери; в каждой клетке $\sim 10^3$ штук; сильно гетерогенная
- 37 генов: 22 тРНК + 2 рРНК + 13 кодирующих
- 13 полипептидов являются частью **митохондриального регуляторного комплекса** (сахара \Rightarrow АТФ) вместе с ядерными генами
- мтДНК, \sim автономна, свой генетический код
- Стоп-кодоны: TAA, TAG, AGA, AGG
- **Митохондриальные заболевания:** группа разнообразных наследственных аномалий окислительного фосфорилирования, появляющихся из-за мутаций в митохондриальной (70%) или ядерной (30%) ДНК
- Известно ~ 300 точечных мутаций в мтДНК, вызывающих различные заболевания



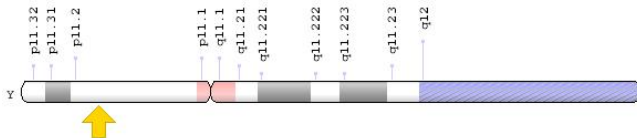
Болезнетворные мутации в мтДНК 🏠




DiMauro (2017) – *Mitochondrial Encephalomyopathies*

A propos: как узнать пол образца NGS?

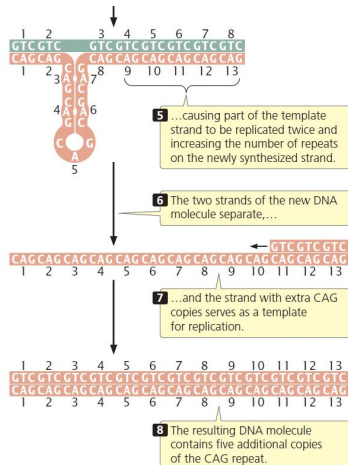
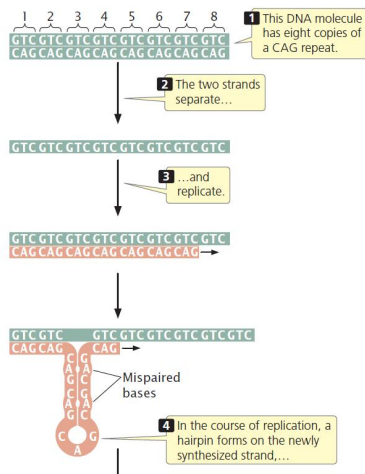
- Соотношение гетерозигот/гомозигот на X хромосоме (Het/Hom):
 $\text{Het/Hom} < 0.8 \Rightarrow \text{М}$, $\text{Het/Hom} > 0.8 \Rightarrow \text{Ж}$; неопределенность: 0.5–1.0
- Ген **SRY** (Sex-determining Region Y) – не содержит интронов, кодирует фактор развития семенников, который принимает участие 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*)



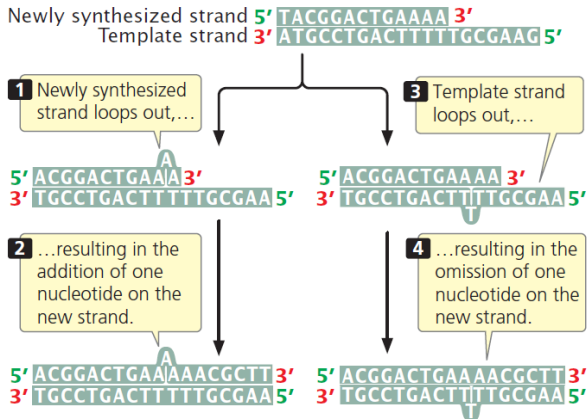
Prerequisites 

- Митоз!

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-паллидолюсисовая 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



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

Определения

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

- Транзиции (G>A, C>T)
- Трансверсии (G>G, A>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 \sim 1кб и более (2%):

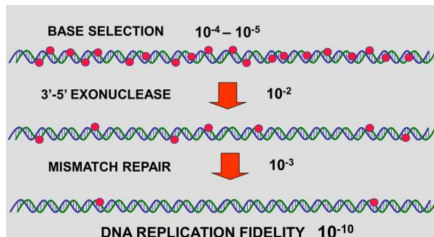
- Сбалансированные: инверсии и транслокации
- Несбалансированные (**варианты числа копий, CNV**): тандемные и диспергированные повторы, вставки, удаления (делеции)

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 *пруффринга* 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*

Генетические заболевания, связанные с 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 неполовой 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 кросслинkinга
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.

Генетические заболевания, связанные с 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>	3	TP53	191170

- **Синдром Ли-Фраумени** – редкое аутосомно-доминантное наследственное заболевание, носители которого имеют предрасположенность 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.


Генетические заболевания, связанные с 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*

Prerequisites 

- Экзогенные и эндогенные механизмы повреждения ДНК
- Механизмы репарации ДНК

Мутации *de novo*

Мутации *de novo* (DNM): наблюдаемые у ребенка, но не у родителей; могут быть выявлены путем секвенирования тройки мать-отец-ребенок.

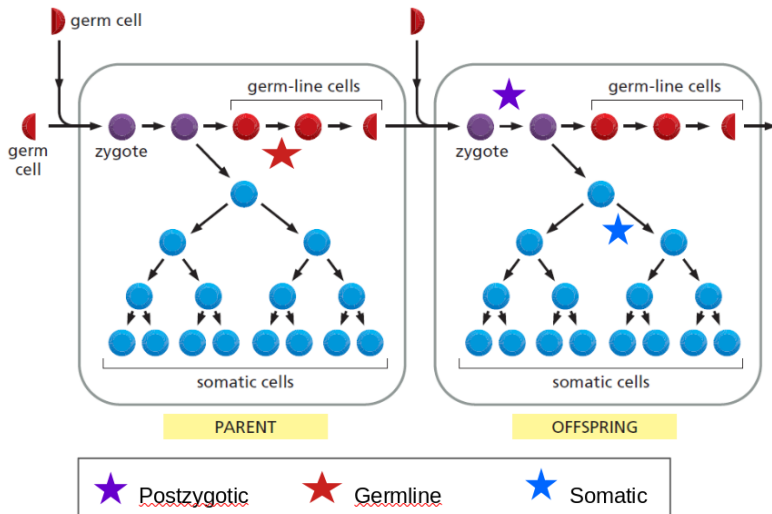
Количество *de novo* ОНВ: 40-80 у каждого новорожденного.

Вариация частоты DNM: по всему геному; в семьях; в кластерах мутаций (в одном индивидууме) и в "горячих точках" мутагенеза.

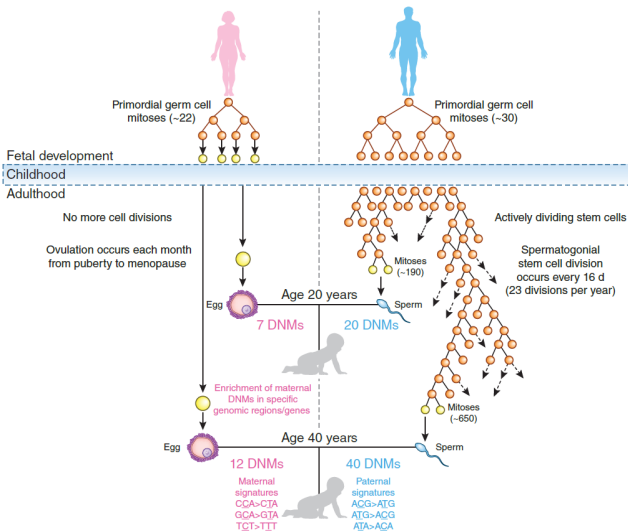
Факторы, влияющие на изменение частоты DNM:

- 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

Время появления и мозаицизм



Происхождение: материнские и отцовские

Goriely (2016) *Nat Genet*

Мутации *de novo*

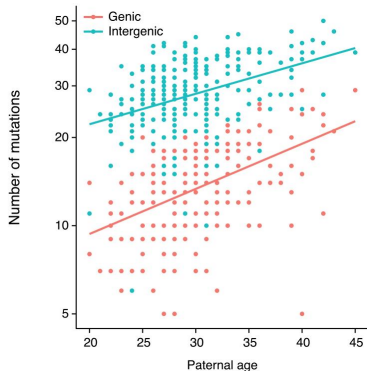
Материнские и отцовские DNM

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

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

~ 80% всех DNM приходят от отца ребенка.

Эффект обладает большой вариабельностью между семьями.



Francioli (2015) *Nat Gene*

Спектр мутаций *de novo*

Транзиции:

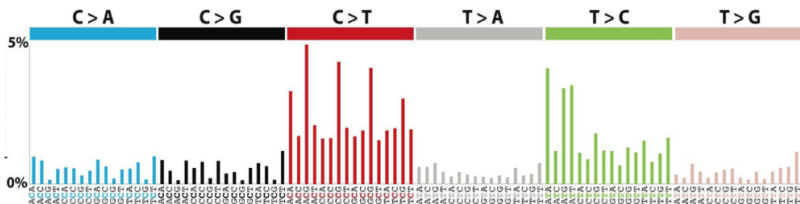
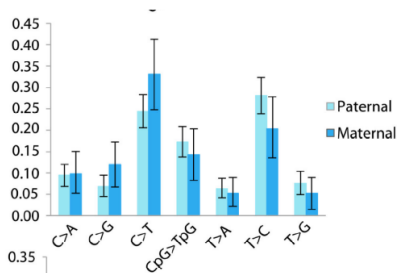
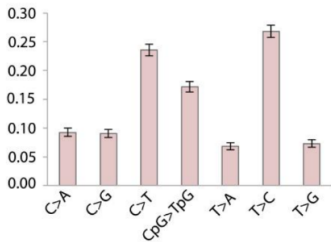
- C:G > T:A, дезаминирование 5-метил-C и C: 40%
- T:A > C:G, причина неизвестна: 25%

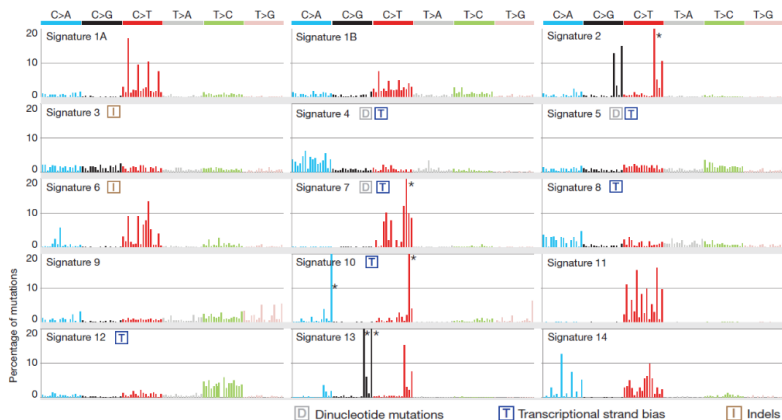
CpG составляют $\sim 1\%$ генома, так что и в не-CpG участках; однако, транзиции в CpG происходят в $\sim 18x$ раз чаще чем в не-CpG участках.

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*

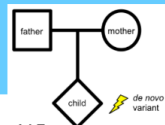
Спектр мутаций *de novo*

Спектр мутаций *de novo*

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*

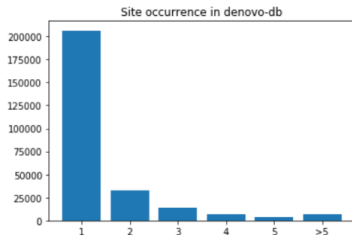
Denovo-db v.1.6.1



	<i>Variant Counts</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>

Мутации *de novo* и заболевания человека

- DNM крайне редки 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 DNM.

Скорость возникновения мутаций и следствия 🏠

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.

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 генах, вызывающих различные заболевания 🏠

Заболевание	MIM	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, соответствующие заболеваниям, в базе OMIM; объясните разницу в частоте возникновения мутаций.

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)

Структурные варианты

Tandem duplication

B C D E

B C C D E

Variable Number of Tandem Repeats

B C D D E



B C D D D E

Dispersed duplication

B C D E

B C D E C

Несбалансированные структурные варианты (CNV)

Структурные варианты

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

Несбалансированные структурные варианты (CNV)

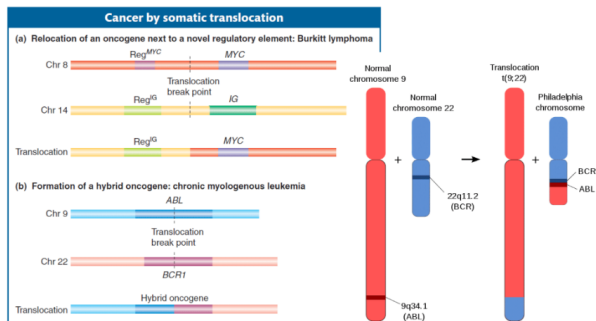
Структурные варианты

Inversion

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

Сбалансированные структурные варианты

Хромосомные перестройки и заболевания



Griffiths – *Introduction to Genetic Analysis*

Протоонкоген *MYC* является транскрипционным фактором, участвующим в клеточном цикле, апоптозе и трансформации клеток. Протоонкоген *ABL* кодирует протеинкиназу, участвующую в сигнальном пути клеточной пролиферации. Гибридный белок Bcr1-ABL обладает постоянной киназной активностью вне зависимости от наличия иницирующего сигнала.

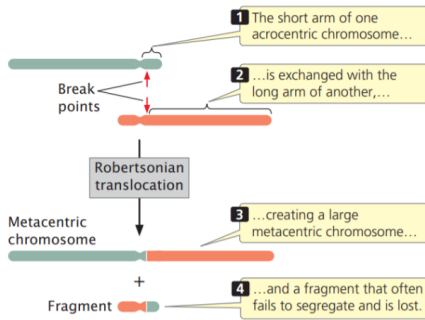
Анеуплоидии

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

- Результат ошибок при митозе или мейозе
- Основная причина спонтанных аборт ($\sim 30\%$ от всех зачатий)
- Выявляются у $\sim 0.3 - 0.6\%$ новорожденных

Название	Кариотип	Частота встречаемости
Синдром Тернера	X0 (у женщин, моносомия по X хромосоме)	1:2000-1:2500
Синдром Клайнфельтера	XXY (XXXУ, XXXXY, XXYY)	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	ХТрисомия по 13 хромосоме	1:7000-1:14000
Трисомия по 8 хромосоме	Трисомия по 8 хромосоме	1:25000-1:50000

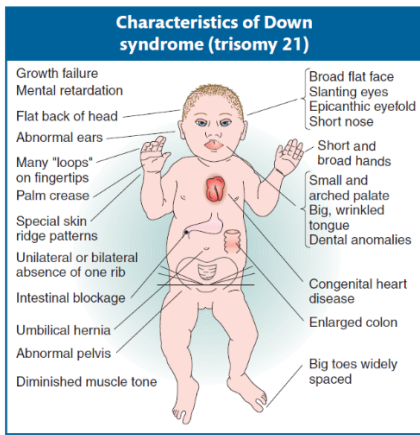
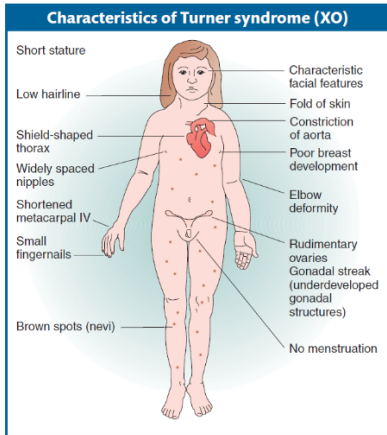
Робертсоновская транслокация



Pierce – *Genetics Essentials. Concepts and Connections*

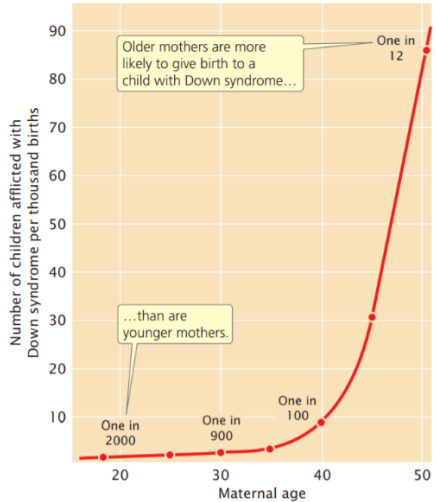
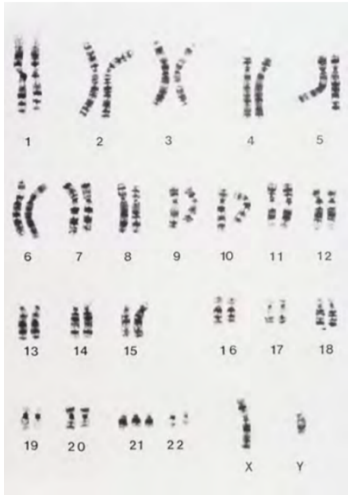
Робертсоновская транслокация: длинные плечи акроцентрических хромосом (13, 14, 15, 21) объединяются общей центромерой, из-за чего появляется хромосома с двумя длинными плечами и, обычно, вторая хромосома с двумя короткими плечами. Поражает $\sim 1/1000$ новорожденных.

Анеуплоидии



Griffiths– *Introduction to Genetic Analysis*

Синдром Дауна



Pierce – *Genetics Essentials. Concepts and Connections*

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*)

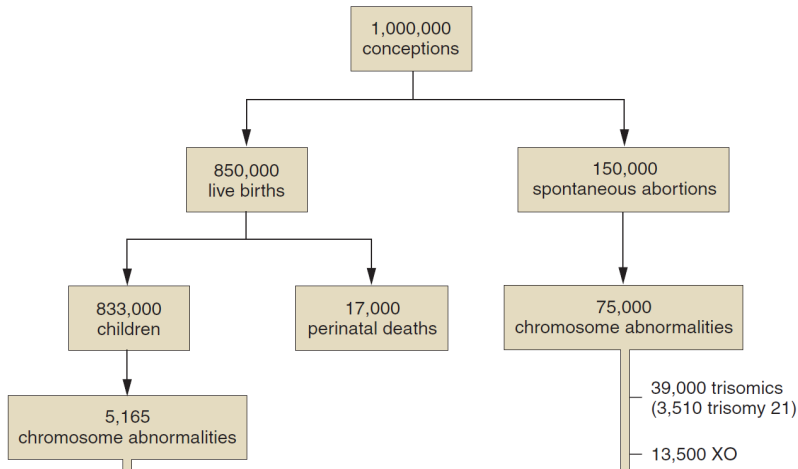
Инактивация X-хромосомы

Одна из двух X-хромосом становится инактивированной и конденсированной в каждой клетке эмбриона женского пола.

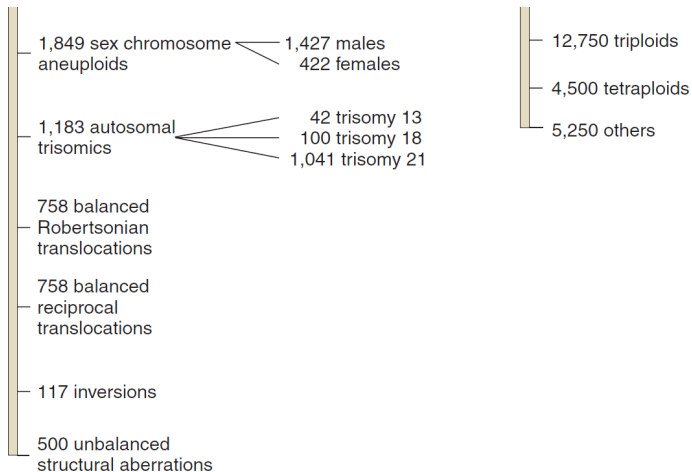
- Происходит на ранних стадиях развития
- Случайна в разных клетках
- Сохраняется при последующих делениях клетки, но не между поколениями
- Ткани женского организма похожи на лоскутное одеяло, в котором в половине случаев инактивирована материнская, а в половине случаев — отцовская хромосома
- Инактивация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



Судьба 1 млн прижившихся человеческих зигот

Griffiths – *Introduction to Genetic Analysis*

Судьба 1 млн прижившихся человеческих зигот

Griffiths – *Introduction to Genetic Analysis*

Частоты возникновения и количества мутаций *de novo*

Тип	Частота возникновения на поколение	Всего у индивидуума
ОНВ	$1.20 \cdot 10^{-8}$ на п. н.; $1.66 \cdot 10^{-8*}$	44 - 82
Повторы динуклеотидов	$1.20 \cdot 10^{-4}$ на локус	-
Кодирующие ОНВ	-	1-2
Короткие инделы (<50 п. н.)	$0.53 - 1.59 \cdot 10^{-9}$ на п. н.; $1.26 \cdot 10^{-8*}$	3-9
Длинные инделы	0.16	0.16
CNV	$10^{-6} - 10^{-4}$ на локус на поколение	0.0154

* Palamara (2015) *AJHG*

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 эндогенных и экзогенных повреждений 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. *The American Journal of Human Genetics* 108, 597–607.