Вопросы

1. Перечислите все известные вам типы мутаций в геноме

2. Какова по порядку величины частота точечных мутаций в геноме из-за ошибок системы репликации (на нуклеотид на репликацию)?

3. Дайте пример заболевания, вызванного нарушениями в системе репарации ДНК. С дефектами какого гена ассоциировано это заболевание?

De novo mutations

De novo mutations (DNM) detected in a genome (exome), for example, by sequencing a mother-father-child trio

Overall dnSNV rate: 40-80 in a newborn

DNM rate variation: across the genome; in families; mutational clusters (within an individiual) and mutational hotspots (across individuals)

Factors contributing to DNM rate variation:

- sequence composition and functional context
- replication timing: early / late
- transcriptional activity and chromatin state
- the number of mitoses a cell has undergone (parental age)
- exposure to damaging agents
- the efficiency of the DNA repair
- the amount of time between mitoses

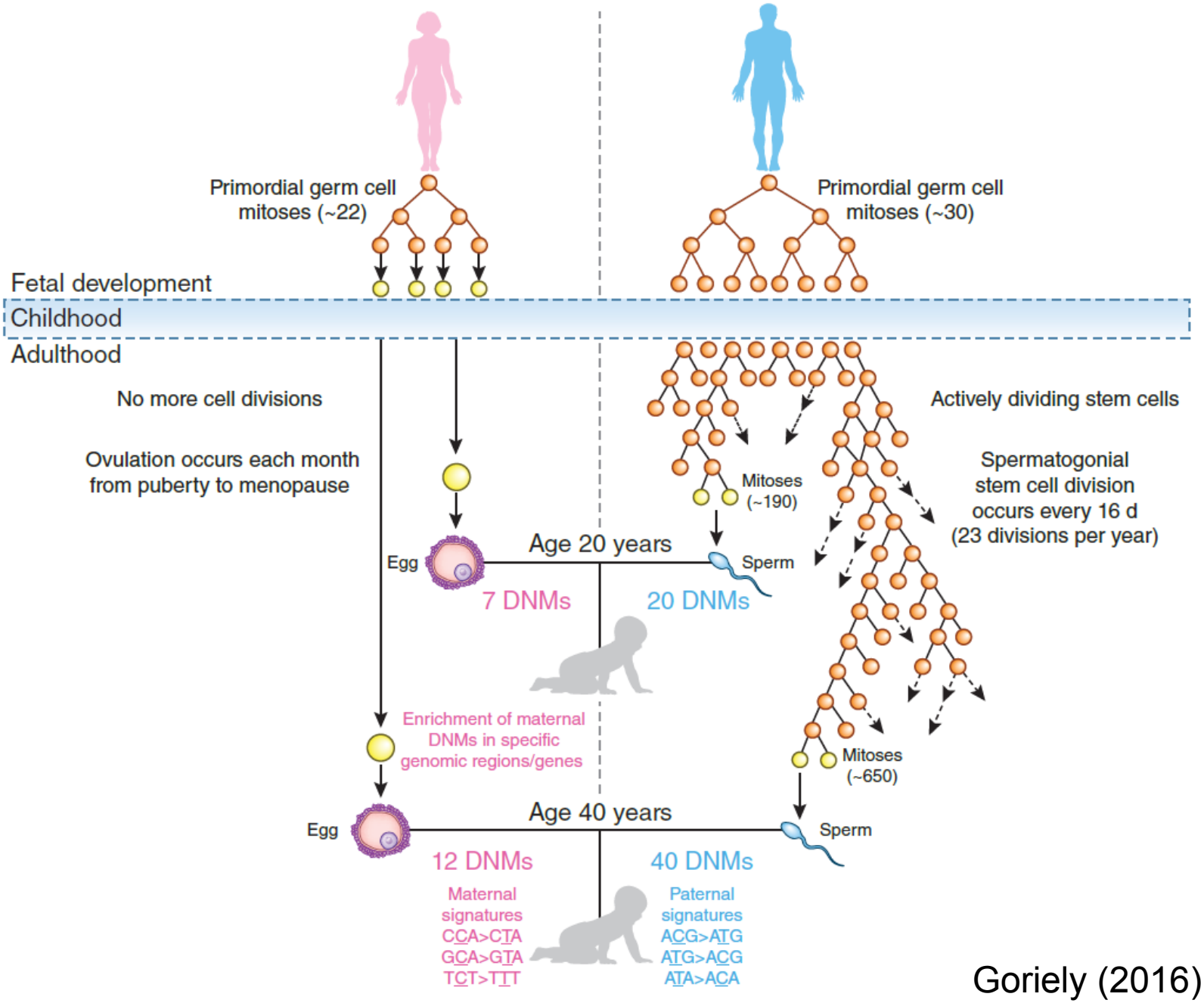
Mutation rates in disease-causing genes

Disorder	MIM	Mutations per Million Gametes	Signs and Symptoms (Phenotype)
X-linked			
Duchenne muscular dystrophy	310200	40–105	Muscle atrophy
Hemophilia A	306700	30–60	Severe impairment of blood clotting
Hemophilia B	306900	0.5–10	Mild impairment of blood clotting
Autosomal Dominant			
Achondroplasia	100800	10	Very short stature
Aniridia	106200	2.6	Absence of iris
Huntington disease	143100	<1	Uncontrollable movements, personality changes
Marfan syndrome	154700	4–6	Long limbs, weakened blood vessel walls
Neurofibromatosis type 1	162200	40–100	Brown skin spots, benign tumors under skin
Osteogenesis imperfecta	166200	10	Easily broken bones
Polycystic kidney disease	600666	60–120	Benign growths in kidneys
Retinoblastoma	180200	5–12	Malignant tumor of retina

Exercise: find genes in OMIM, explain the rate differences

Lewis – *Human genetics. Concepts and applications*

De novo mutations



Goriely (2016) *Nat Genet*

De novo mutations

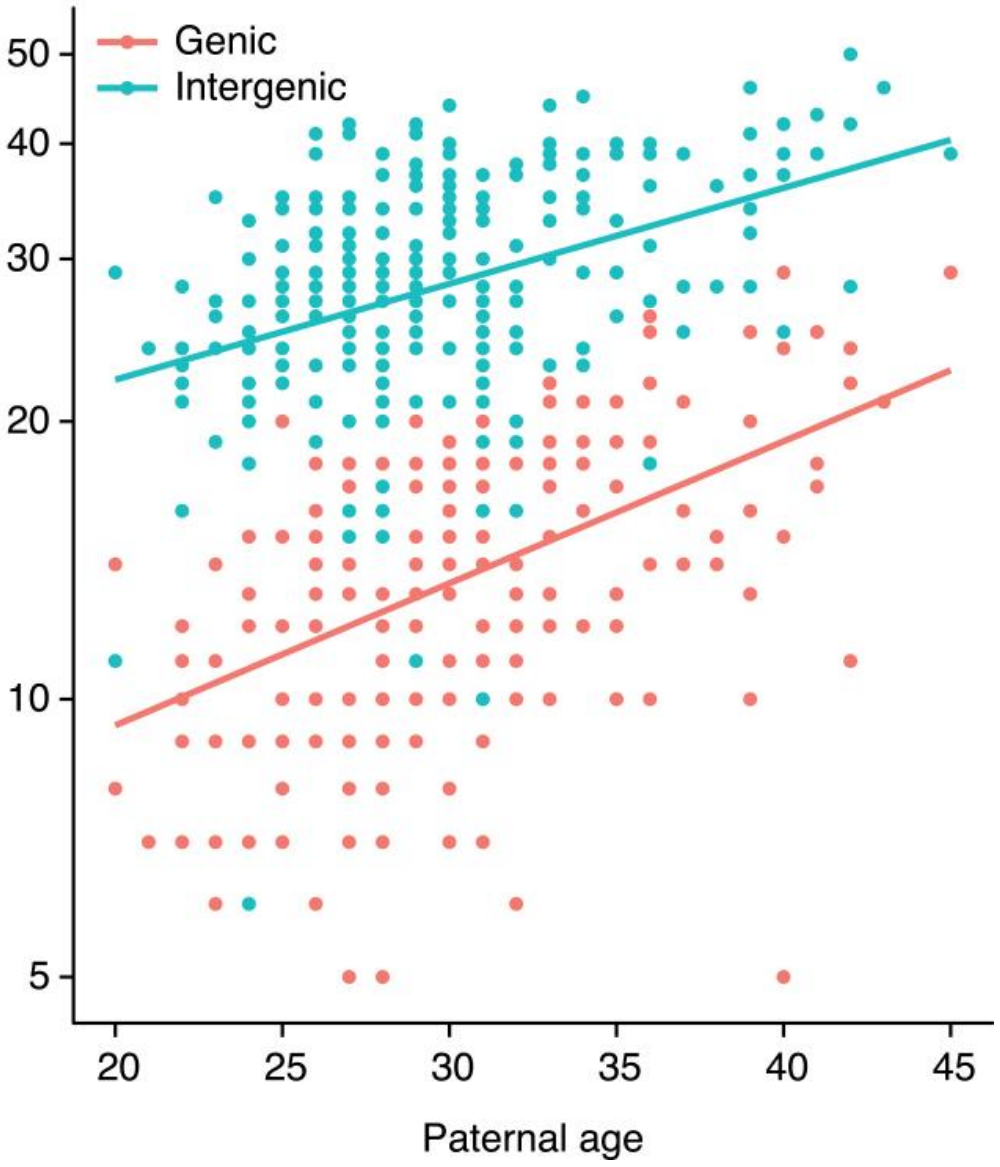
Paternal and maternal DNMs

+1-3 DNMs for each year of paternal age at conception.

+0.24 DNMs for each year of maternal age at conception (non-replicative DNA damage)

~80% of all DNMs are paternal

This effect varies considerably between families



Francioli (2015) *Nat Genet*

De novo mutation spectra

Transitions

- C:G>T:A, deamination of 5-methyl-C and C: 40%
- T:A>C:G, cause unknown: 25%

Note: CpG are only ~1% of the genome, so also at non-CpG; but transitions at CpG are ~18x more frequent than non-CpG

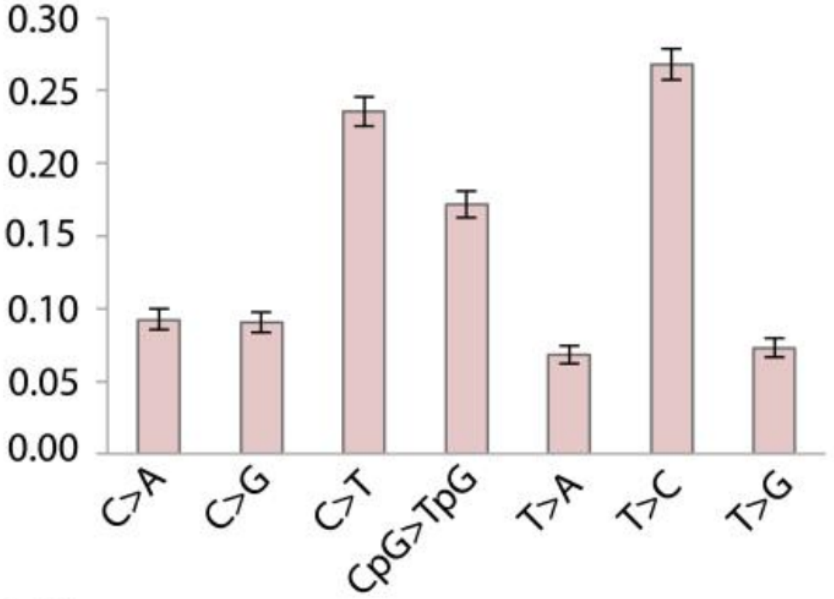
Transversions

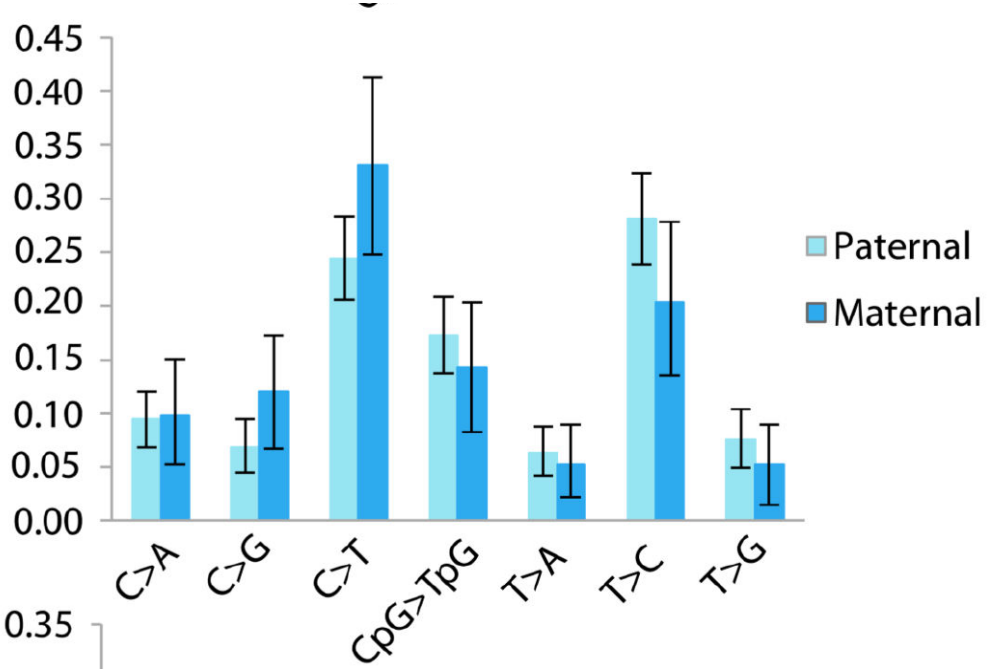
- G:C>T:A: 10%
- G:C>C:G: 10%
- A:T>C:G: <8%
- A:T>T:A: <8%

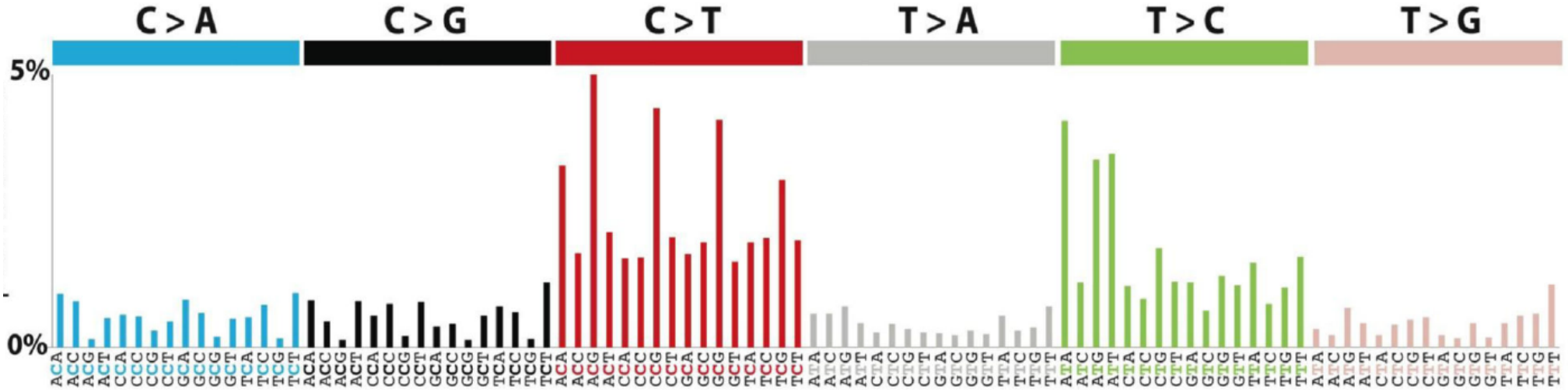
Note: transversions occur ~2.5x more frequently at CpG sites

Ohno (2019) *Genes & Genet Systems*

De novo mutation spectra







6,570 high confidence DNMs from 109 trios

Rahbari et al. (2016) *Nat Genet*

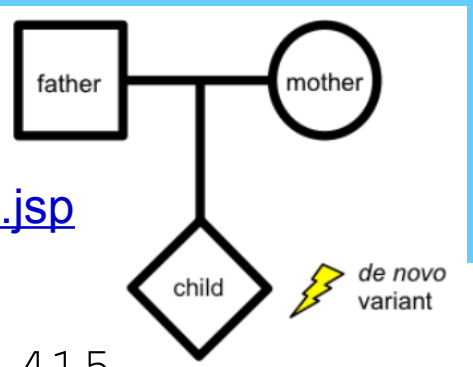
De novo mutation spectra

Validated mutational signatures found in human cancer
Each signature is displayed according to the 96 substitution classification defined by the substitution class and sequence context immediately 3' and 5' to the mutated base.

Alexandrov (2013) *Nature*

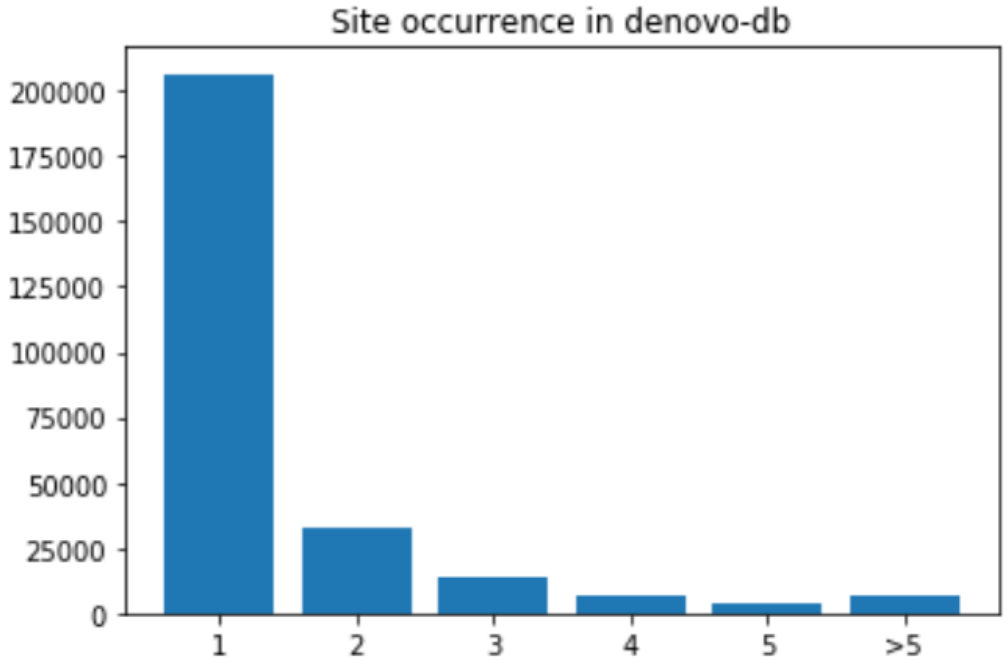
Denovo-db v.1.6.1

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

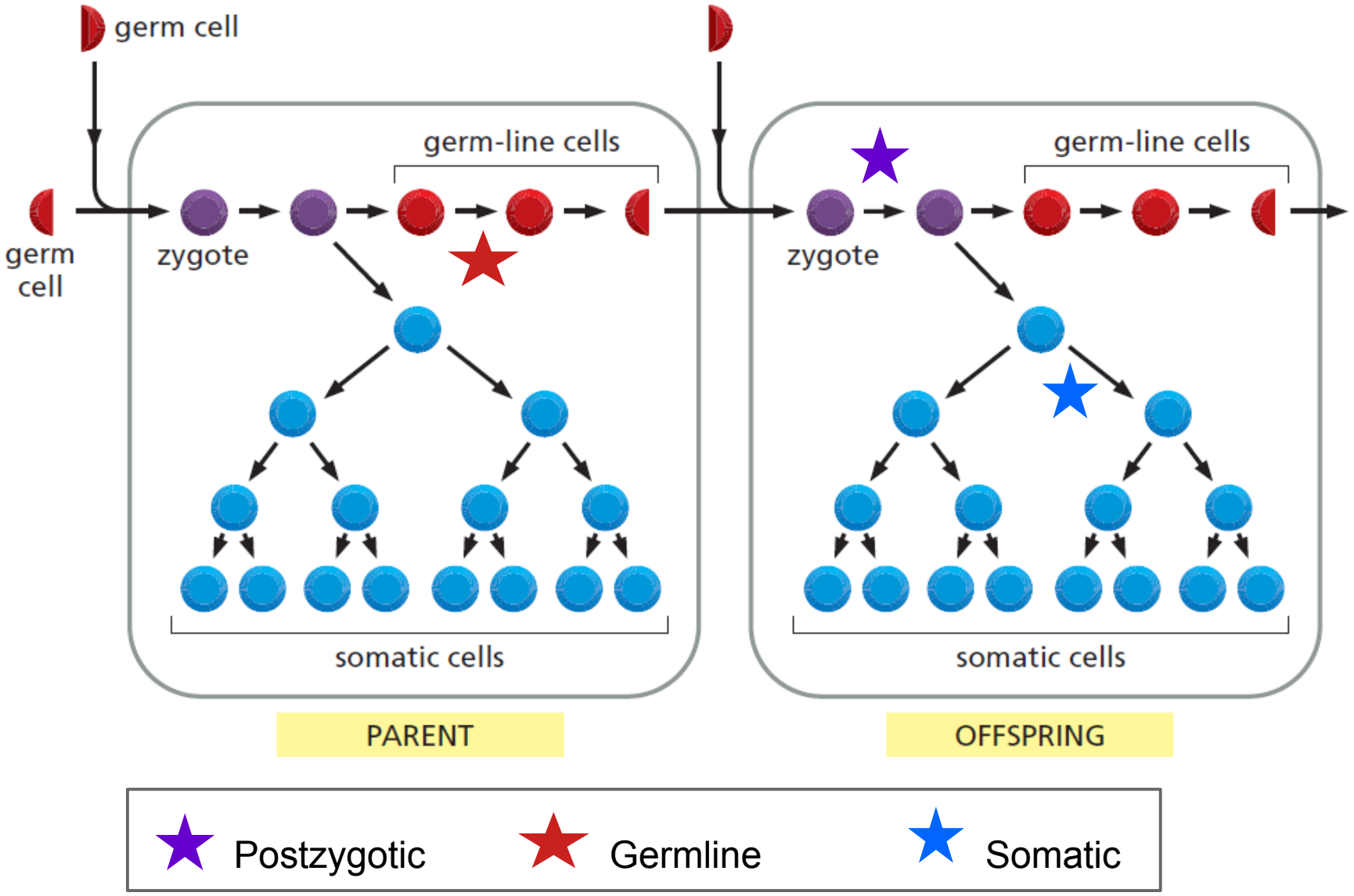


Variant	*Counts*	*%*
C>T	79,110	19.0
G>A	79,016	19.0
A>G	47,666	11.5
T>C	47,584	11.5
C>G	17,431	4.2
G>C	17,154	4.1
C>A	15,719	3.8
G>T	15,234	3.7
A>C	12,744	3.1
T>G	12,464	3.0
T>A	11,637	2.8
A>T	11,482	2.8
T>TA	739	0.2
CT>C	737	0.2
...		
Total	*415,515*	

Variants:	415,415
Samples:	11,518
Genes:	17,431
Sites:	270,506
Missense:	20,815
Pred. Damaging:	~75%



Mutation timing and mosaicism



Alberts - *Essential Cell Biology*, Fig 9-3

De novo mutations in human disease

- Ultra-rare individually, but significant collectively: 60-75% of all sporadic disease cases are DNMs
- More damaging than inherited; effect depends on timing
- **Severe pediatric disorders in outbred populations**: sporadic malformation syndromes (Schinzel–Giedion, Kabuki, Bohring–Opitz), neurodevelopmental (severe intellectual disability, ID), congenital heart disease (CHD)
- **Late-onset neurological and psychiatric disorders**: Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), schizophrenia (SCZ), epilepsy, Tourette syndrome (TS), autistic-spectrum disorder (ASD), and bipolar disorder (BP)
  Example: 10% SCZ cases have DNM CNV vs 1.26% controls
- **Inherited cancers**: Li-Fraumeni syndrome (TP53), familial adenomatous polyposis (APC), ~7% of non-somatic mutations are DNMs

Structural variants and CNVs

Structural variant (aka **chromosomal rearrangement)**: sequence change >1 kbp in size

- Balanced
  Inversion or translocation
- Unbalanced (aka **CNV, copy number variant**)
  Tandem or dispersed duplication, deletion, insertion

Mechanisms

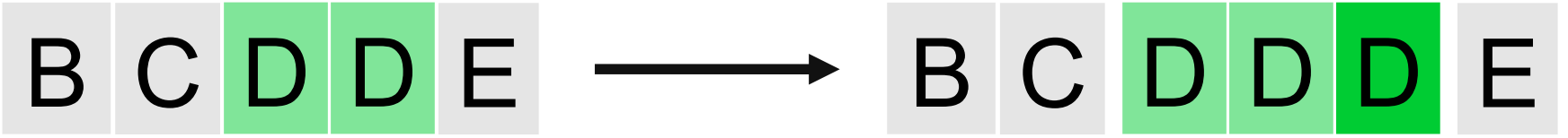
- Recombination: non-allelic homologous recombination (NAHR)
- Nonreplicative: Nonhomologous end joining (NHEJ) repair
- Replication-based:
  Fork stalling and template switching (FoSTeS)
  Microhomology-mediated break-induced replication (MMBIR)
- Retrotransposition (LINE1, Alu repeat)

Structural variants and CNVs

1. Unbalanced structural variants (CNVs)

Tandem duplication

Variable Number of Tandem Repeats



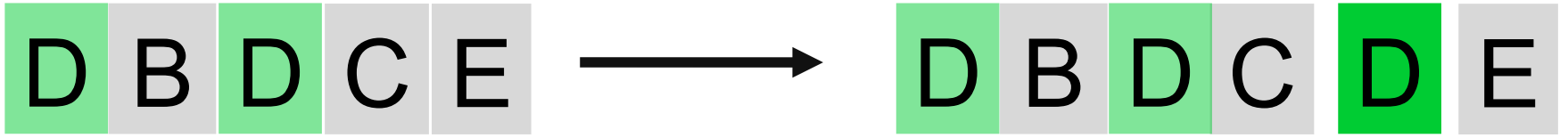
Dispersed duplication

Structural variants and CNVs

1. Unbalanced structural variants (CNVs)

Insertion

Repeat insertion



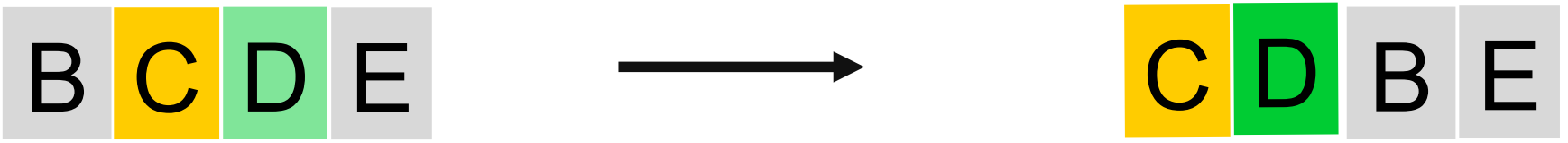
Deletion

Structural variants and CNVs

2. Balanced structural variants

Inversion

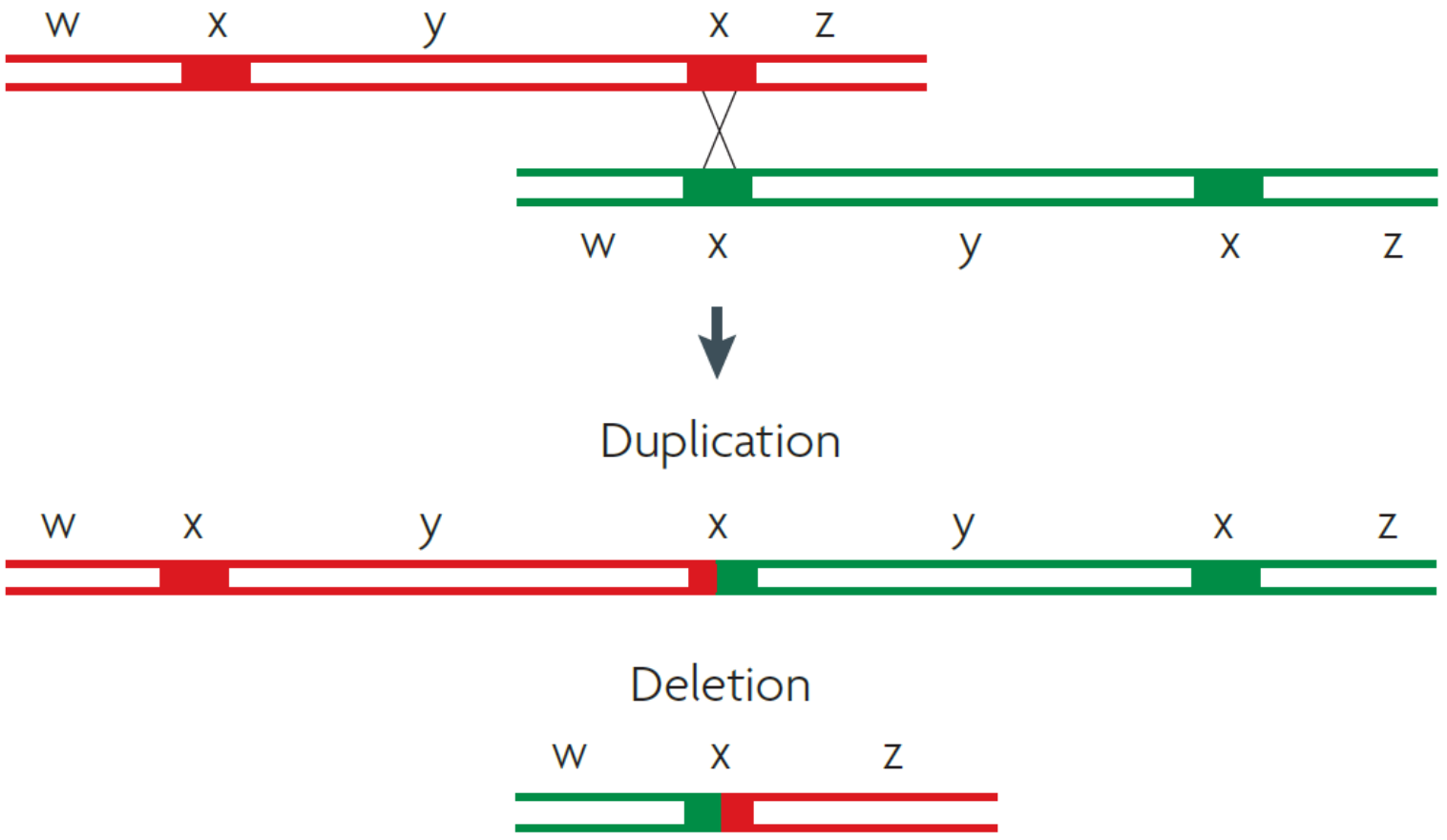
Intra-chromosomal translocation (ITX)



Inter-chromosomal translocation (CTX)

Mechanisms of chromosomal rearrangements

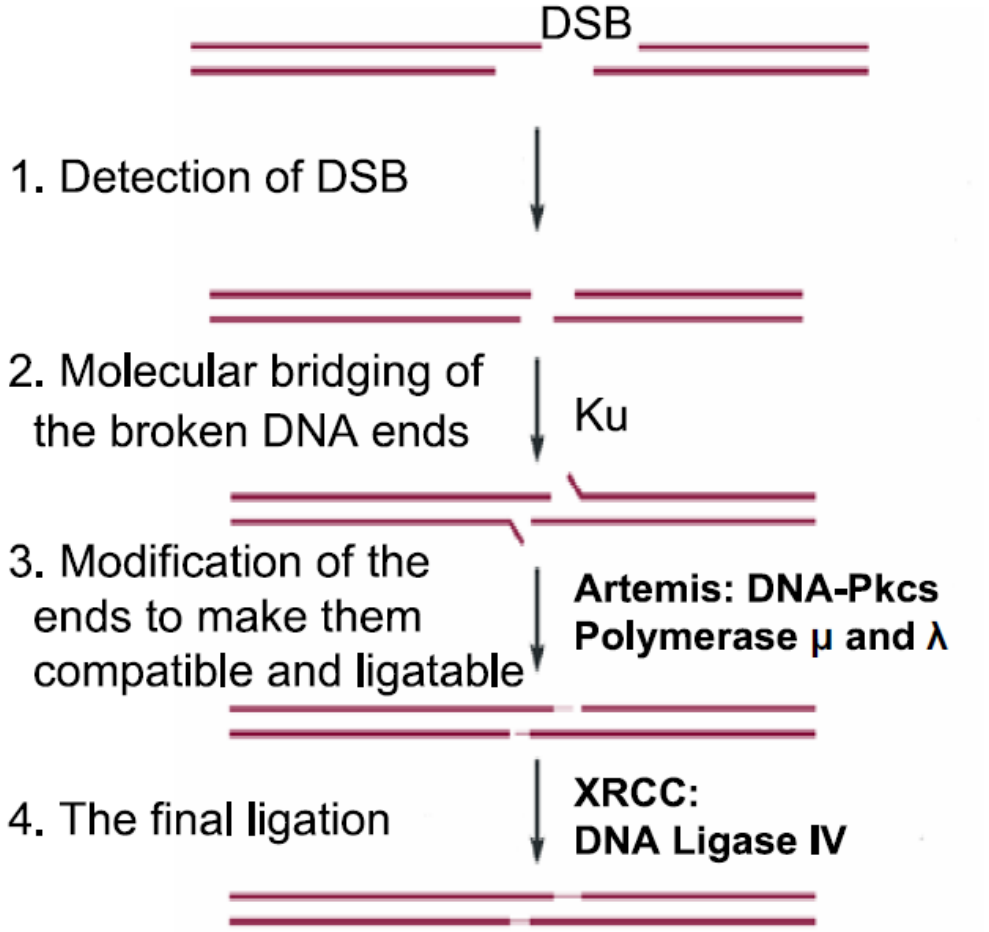
Non-allelic homologous recombination (NAHR)



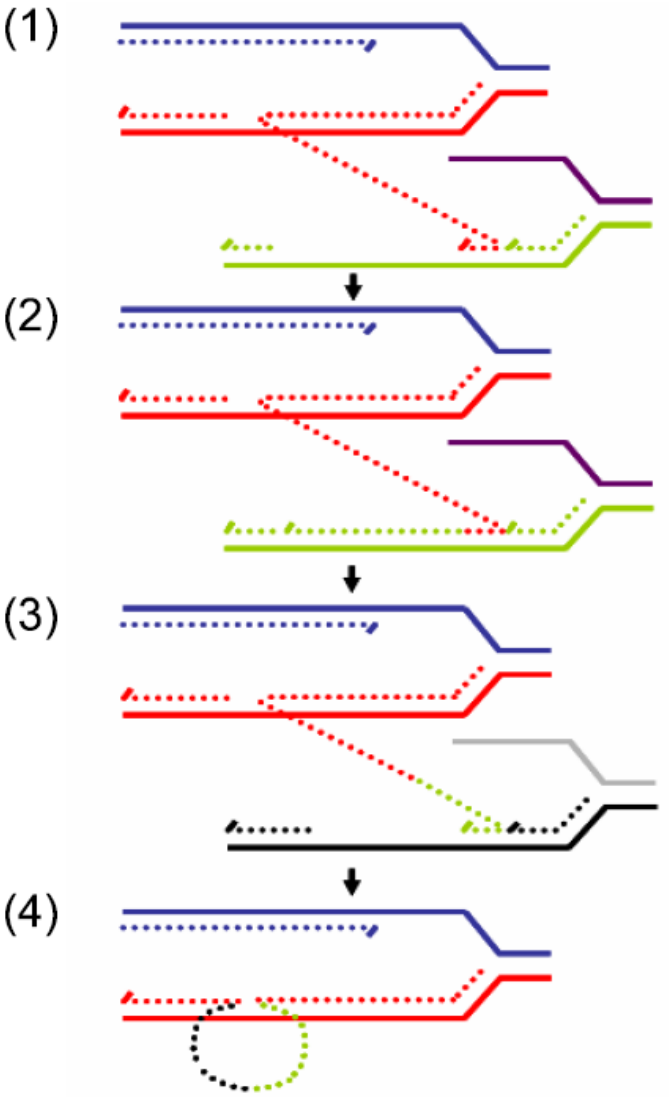
Hastings (2009) *Nat Rev Genet*

Mechanisms of chromosomal rearrangements

Nonhomologous end joining (NHEJ)

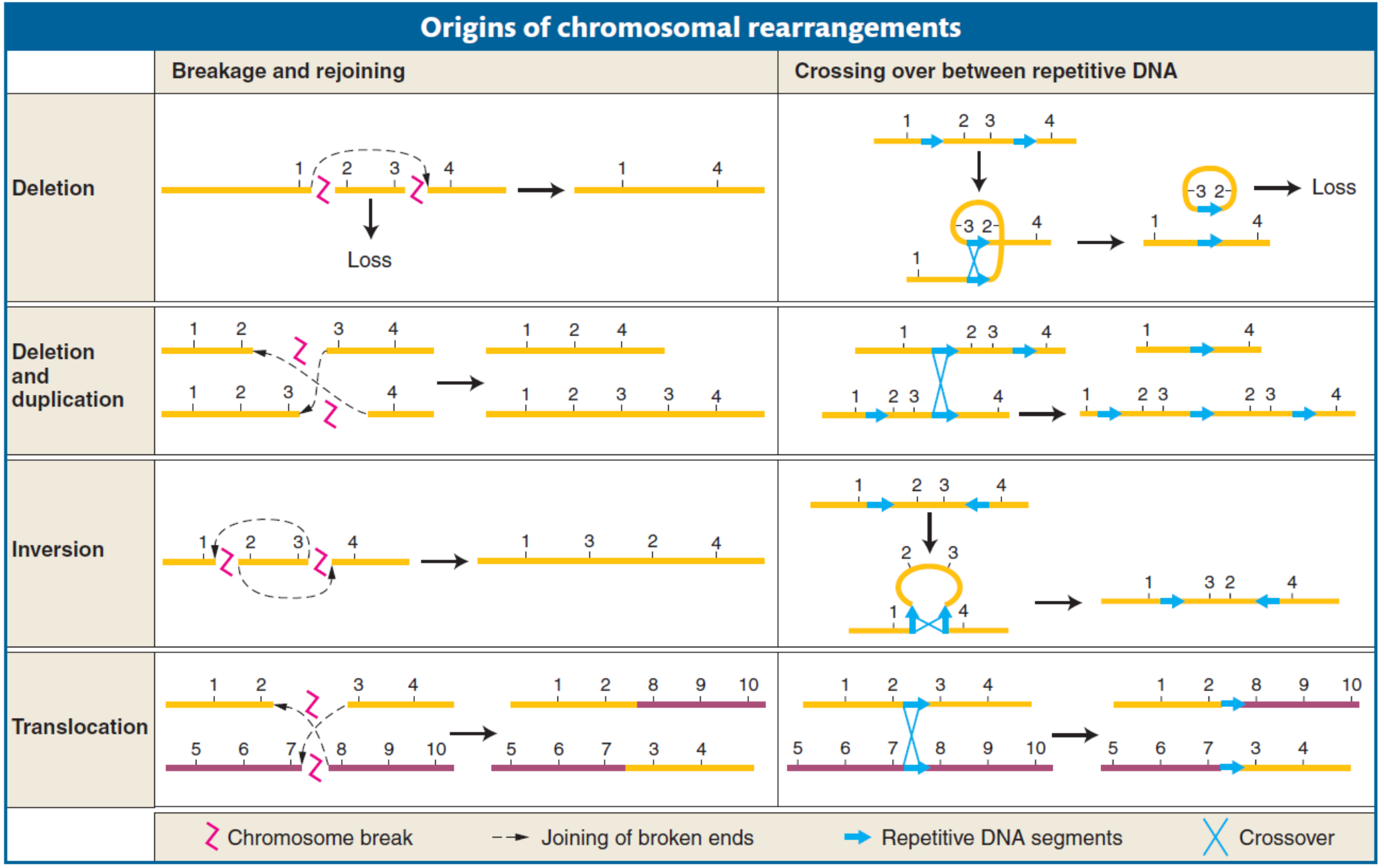


Fork stalling and template switching (FoSTeS)



Gu(2008) *Pathogenetics*

Mechanisms of chromosomal rearrangements



Griffiths -- *Introduction to Genetic Analysis*

Am J Hum Genet. 2021 Apr 1;108(4):597-607. doi: 10.1016/j.ajhg.2021.02.012. Epub 2021 Mar 5.

De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families

Jonathan R Belyeu [1], Harrison Brand [2], Harold Wang [2], Xuefang Zhao [2], Brent S Pedersen [1], Julie Feusier [3], Meenal Gupta [1], Thomas J Nicholas [1], Joseph Brown [1], Lisa Baird [1], Bernie Devlin [4], Stephan J Sanders [5], Lynn B Jorde [6], Michael E Talkowski [7], Aaron R Quinlan [8]

Abstract

Each human genome includes *de novo* mutations that arose during gametogenesis. While these germline mutations represent a fundamental source of new genetic diversity, they can also create deleterious alleles that impact fitness. Whereas the rate and patterns of point mutations in the human germline are now well understood, far less is known about the frequency and features that impact de novo structural variants (dnSVs).

Am J Hum Genet. 2021 Apr 1;108(4):597-607. doi: 10.1016/j.ajhg.2021.02.012. Epub 2021 Mar 5.

De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families

Jonathan R Belyeu [1], Harrison Brand [2], Harold Wang [2], Xuefang Zhao [2], Brent S Pedersen [1], Julie Feusier [3], Meenal Gupta [1], Thomas J Nicholas [1], Joseph Brown [1], Lisa Baird [1], Bernie Devlin [4], Stephan J Sanders [5], Lynn B Jorde [6], Michael E Talkowski [7], Aaron R Quinlan [8]

Introduction

Several mechanisms, including replication infidelity, genomic damage, non-allelic recombination, and double-strand break repair, are known to create de novo mutations (DNMs) in the human germline. These mutations contribute to genomic diversity and often are primary targets in the analysis of rare, dominant genetic disorders. There is therefore a long-standing interest in understanding the frequency at which DNMs occur and the patterns that affect these rates. Numerous studies have measured the rate of germline de novo single-nucleotide variants (dnSNVs) and small insertion-deletion mutations (indels) at approximately 70 events per individual, and it has been established that the majority of these small point mutations arise on the paternal gamete. The frequency of single-nucleotide and insertion-deletion DNMs increases with parental age, especially paternal age.

Am J Hum Genet. 2021 Apr 1;108(4):597-607. doi: 10.1016/j.ajhg.2021.02.012. Epub 2021 Mar 5.

De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families

Jonathan R Belyeu [1], Harrison Brand [2], Harold Wang [2], Xuefang Zhao [2], Brent S Pedersen [1], Julie Feusier [3], Meenal Gupta [1], Thomas J Nicholas [1], Joseph Brown [1], Lisa Baird [1], Bernie Devlin [4], Stephan J Sanders [5], Lynn B Jorde [6], Michael E Talkowski [7], Aaron R Quinlan [8]

In contrast, precise estimates of germline mutations affecting the structure of the human genome (structural variants [SVs]) have been far more difficult to discern.

De novo SVs (dnSVs) largely arise from mutational mechanisms that are distinct from those responsible for point mutations. The larger size of SVs, defined here and in many other studies as variants affecting at least 50 base pairs, increases the likelihood that any given SV will impact protein-coding genes or other critical genomic regions. Understanding the selective constraints on dnSV-specific mechanisms is essential because a broad spectrum of balanced, unbalanced, and complex structural mutations are known to underlie many developmental disorders. However, dnSVs are predicted to occur several hundred-fold less frequently than point mutations, requiring a much larger sample size to achieve accurate estimates of dnSV rates.

Am J Hum Genet. 2021 Apr 1;108(4):597-607. doi: 10.1016/j.ajhg.2021.02.012. Epub 2021 Mar 5.

De novo structural mutation rates and gamete-of-origin biases revealed through genome sequencing of 2,396 families

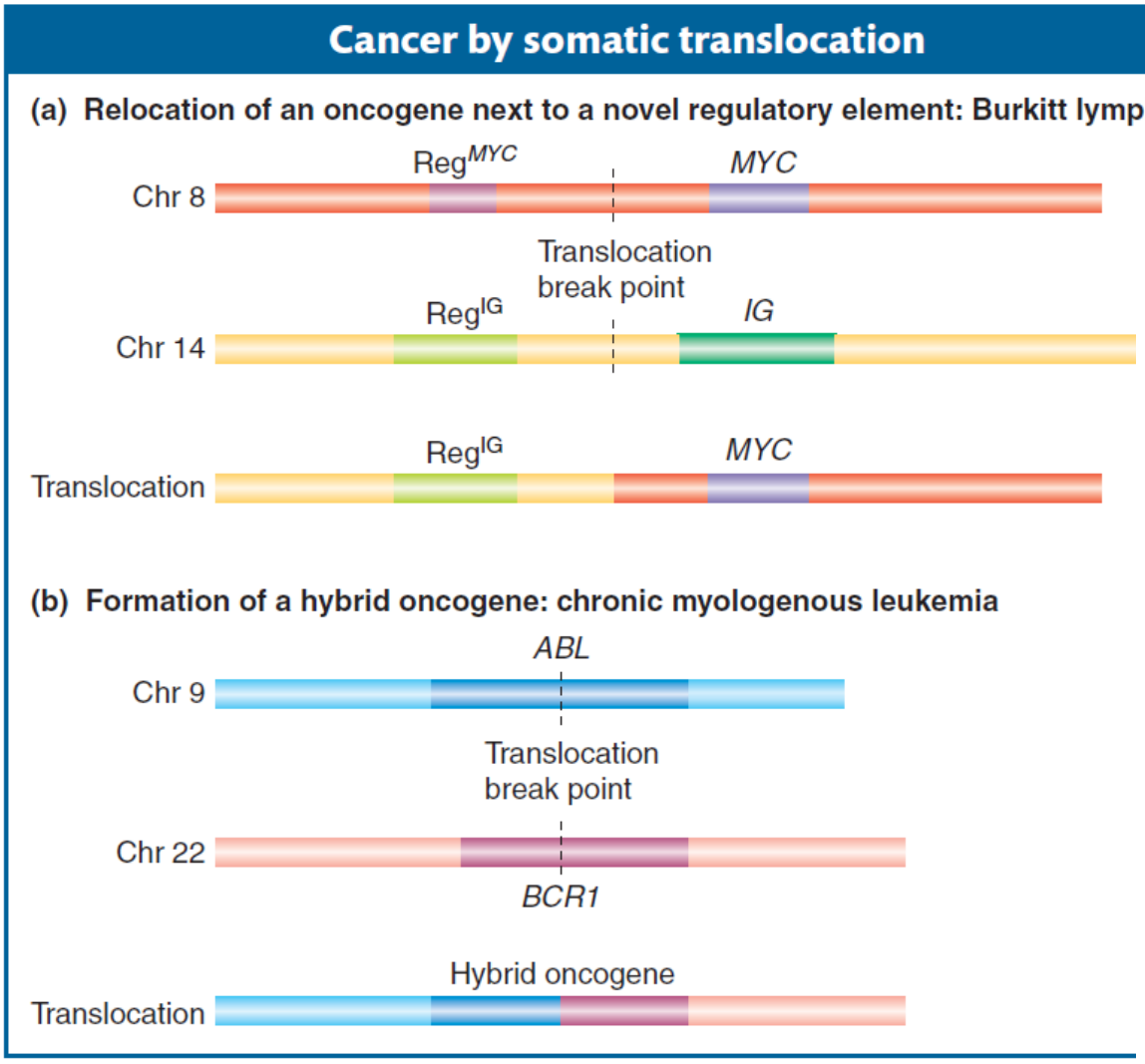
Jonathan R Belyeu [1], Harrison Brand [2], Harold Wang [2], Xuefang Zhao [2], Brent S Pedersen [1], Julie Feusier [3], Meenal Gupta [1], Thomas J Nicholas [1], Joseph Brown [1], Lisa Baird [1], Bernie Devlin [4], Stephan J Sanders [5], Lynn B Jorde [6], Michael E Talkowski [7], Aaron R Quinlan [8]

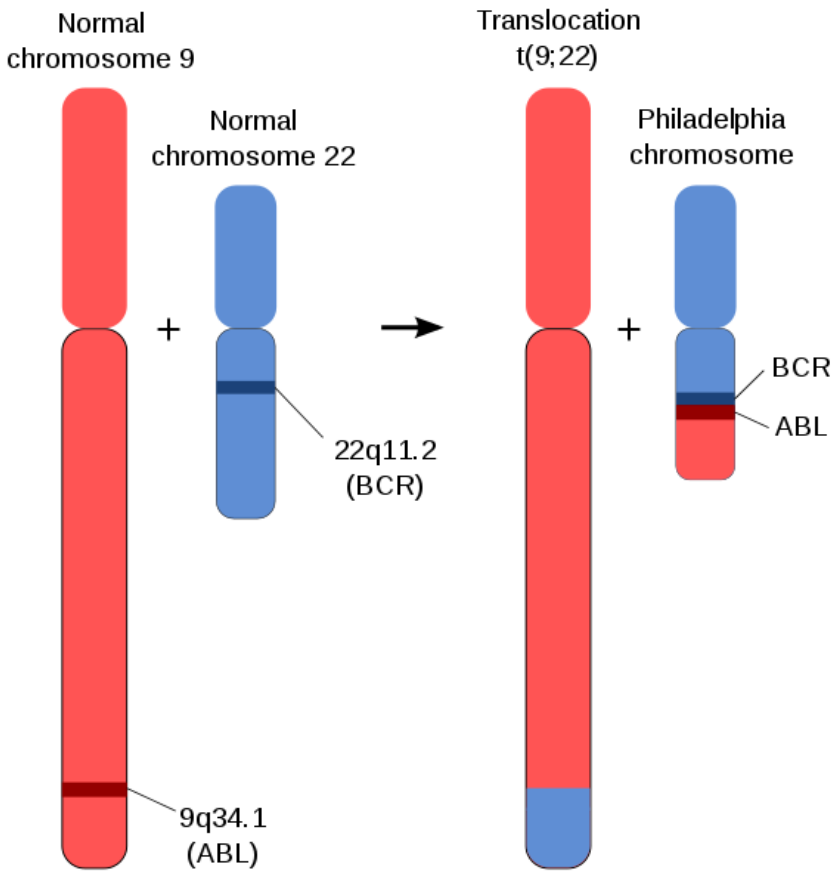
- Family-based study of germline mutations among 9,599 human genomes from 33 multigenerational CEPH-Utah families and 2,384 families from the Simons Foundation Autism Research Initiative; short-read WGS

- dnSV rate: 0.160 events per genome in unaffected individuals, 0.206 per genome) in ASD-affected individuals.
- In both probands and unaffected samples, ~73% of dnSVs arose in paternal gametes
- Most de novo structural mutations to be caused by mutational mechanisms that do not require sequence homology.
- No statistically significant correlation between parental age and dnSV in offspring.

Conclusion: dnSVs have different mechanisms than dnSNVs

Chromosomal rearrangements and disease





The *MYC* proto-oncogene is a transcription factor that plays a role in cell cycle progression, apoptosis and cellular transformation. The *ABL* proto-oncogene encodes a protein kinase in a cell proliferation signaling pathway. The Bcr1-Abl fusion protein has a permanent kinase activity, regardless of the initiating signal.

Griffiths -- *Introduction to Genetic Analysis*

Chromosomal rearrangements and disease

Syndrome	Chromosomal location and key genes (if identified)	Typical size of deletion/duplication	Estimated incidence among live-births	Typical phenotypic features (not exhaustive, and not all these features are seen in all cases)
Di George syndrome/22q11 deletion syndrome	**22q11.2** *TBX1, COMT*	3 Mb deletion (90% of cases)	1/4000	Congenital heart defects, cleft palate, developmental delay, learning difficulty, increased risk of mental illness, recurrent infections
Williams syndrome/Williams–Beuren syndrome	**7q11.3** *CLIP2, ELN, GTF2I, GTF2IRD1, LIMK1*	1.5–1.8 Mb deletion	1/7500 to 1/10000	Supravalvular aortic stenosis, joint problems and loose skin, mild to moderate intellectual disability, characteristic 'elfin' facial appearance
Smith–Magenis syndrome	**17p11.2** *RAI1*	Approximately 3.6 Mb deletion	1/15000 to 1/25000	Mild to moderate intellectual disability, disturbed sleep patterns, behaviour problems including aggression and self-harm
Cri-du-chat syndrome	**5p15.2** *CTNND2*	Approximately 5–40 Mb deletion	1/15000 to 1/50000	Cat-like cry, microcephaly, severe psychomotor problems and severe intellectual disability
Wolf–Hirschhorn syndrome	**4p16.3** *NSD2, LETM1, MSX1*	Approximately 5–18 Mb deletion	1/50000	Characteristic 'Greek warrior helmet' facial appearance, delayed growth and development, mild to severe intellectual disability
Potocki–Lupski syndrome	**17p11.2** *RAI1*	Approximately 3.6 Mb duplication	1/25000	Developmental delay, mild to moderate learning disability, behavioural problems
Cat eye syndrome/Schmid–Fraccaro syndrome	**22q11** *ADA2, CECR2*	2–5 Mb duplication or triplication	1/50000 to 1/150000	Preauricular skin tags or pits, ocular coloboma, anal atresia with fistula, heart and renal malformations

Jackson (2018) *Essays in Biochemistry*

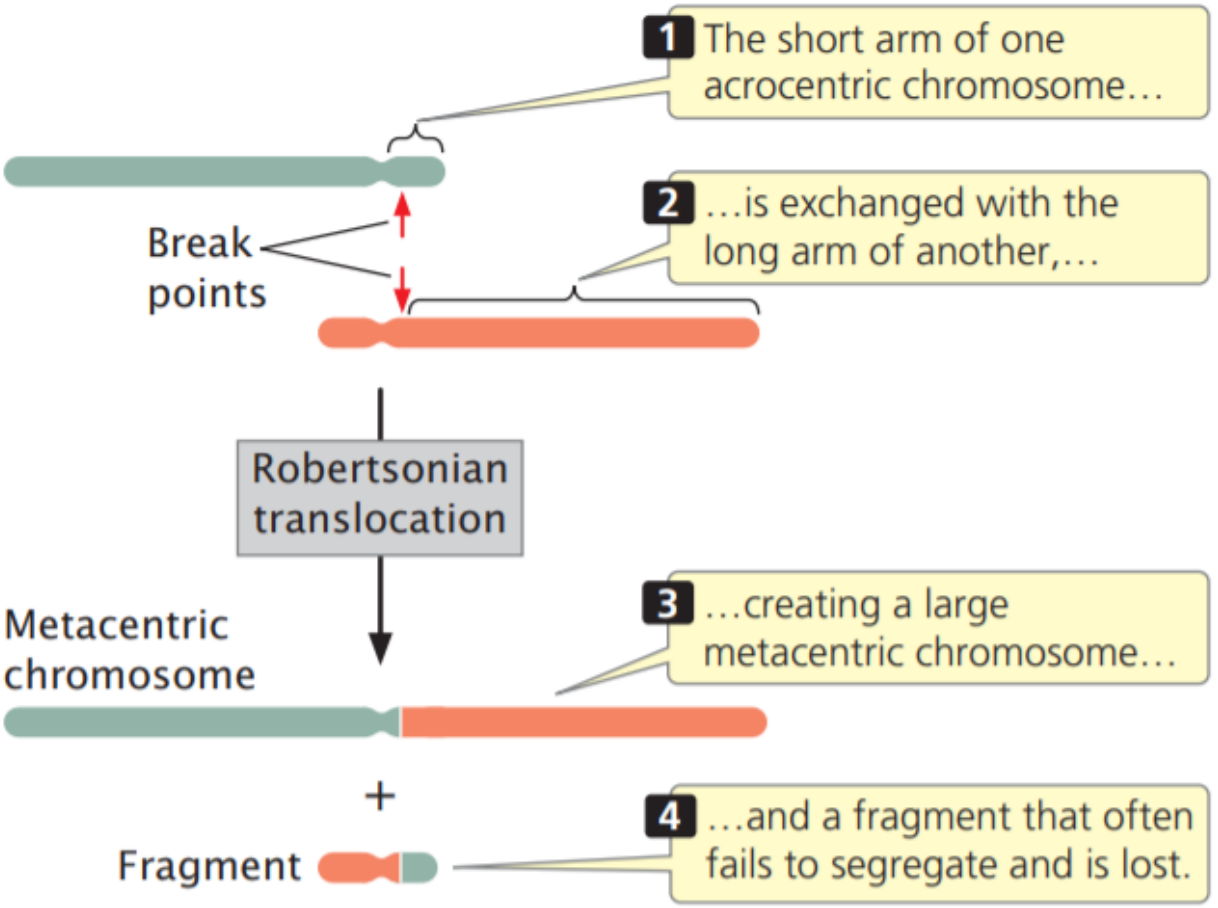
Aneuploidy

Aneuploidy: wrong number of complete chromosomes: nullisomy, monosomy, trisomy. Results from aberration in mitosis or meiosis

- Major cause of spontaneous abortions (~30% of all conceptions)
- Detected in ~0.3-0.6% live human births

Name	Karyotype	Frequency
Turner syndrome	XO (Females, X monosomy)	1:2000- 1:2500
Klinefelter syndrome	XXY (XXXY, XXXXY, XXYY)	1:1000
Poly-X females	XXX	1:1000
Down syndrome	Trisomy 21	1:1100
Edwards syndrome	Trisomy 18	1:6000
Patau syndrome	Trisomy 13	1:7000-1:14000
Trisomy 8	Trisomy 8	1:25000 – 1:50000

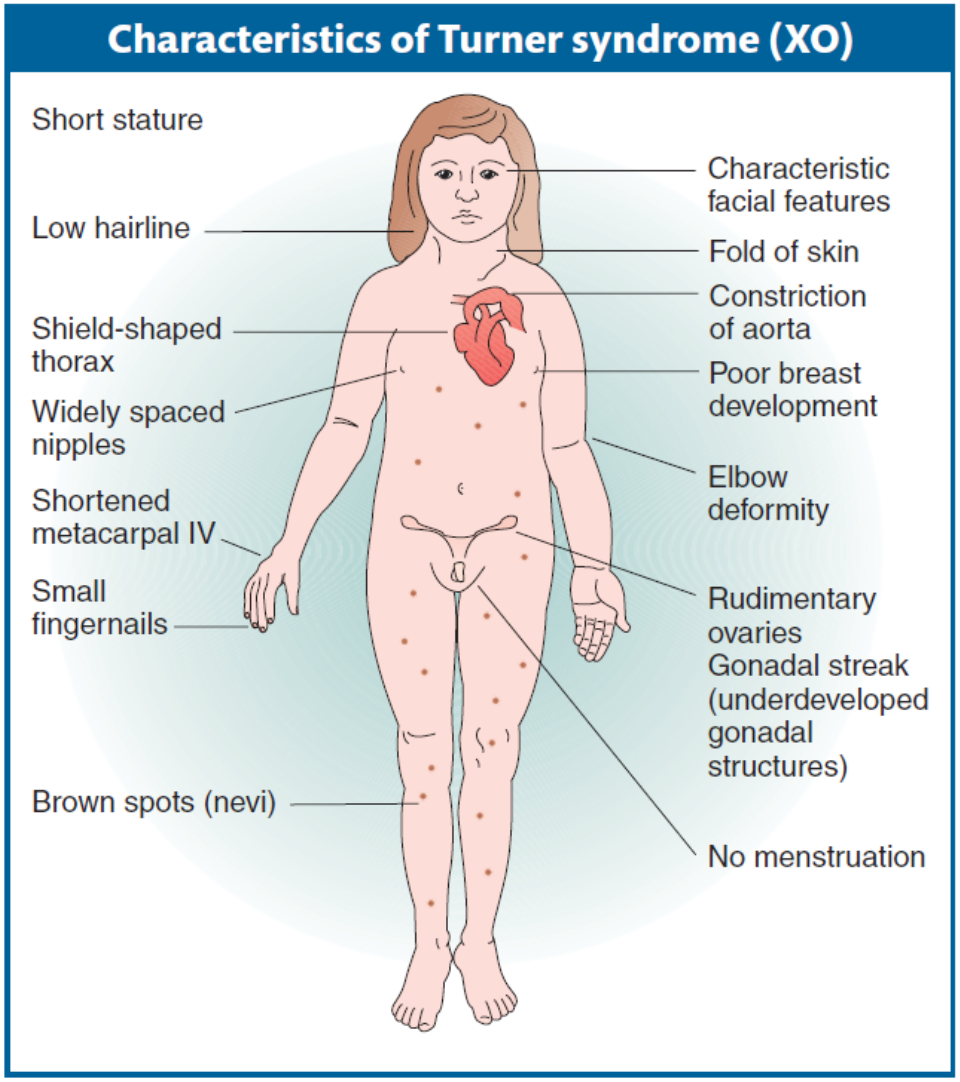
Robertsonian translocation

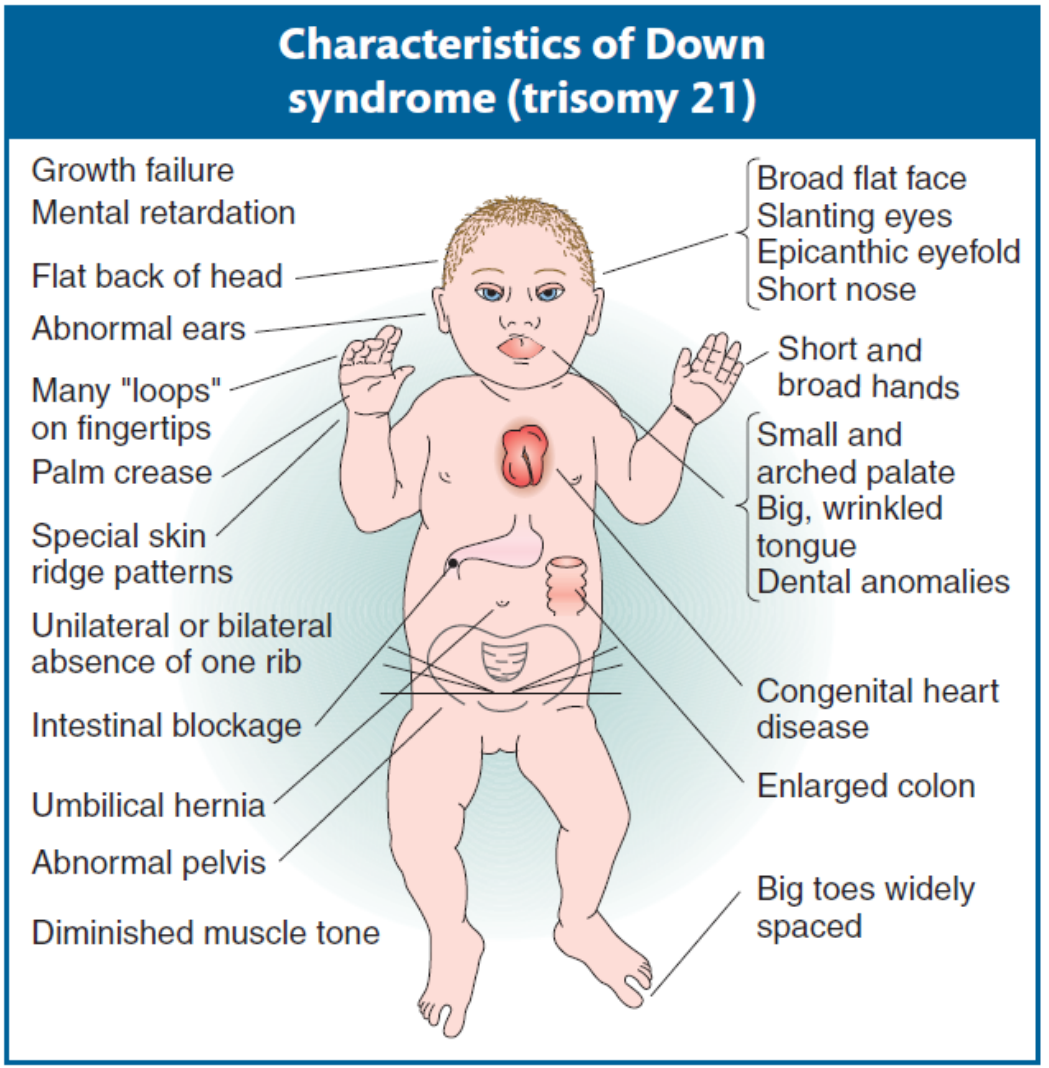


Robertsonian translocation: the long arms of two acrocentric chromosomes (13,14,15,21) become joined to a common centromere, resulting in a chromosome with two long arms and usually another chromosome with two short arms. **Affects ~1/1000 newborns.**

Pierce -- *Genetics Essentials. Concepts and Connections*

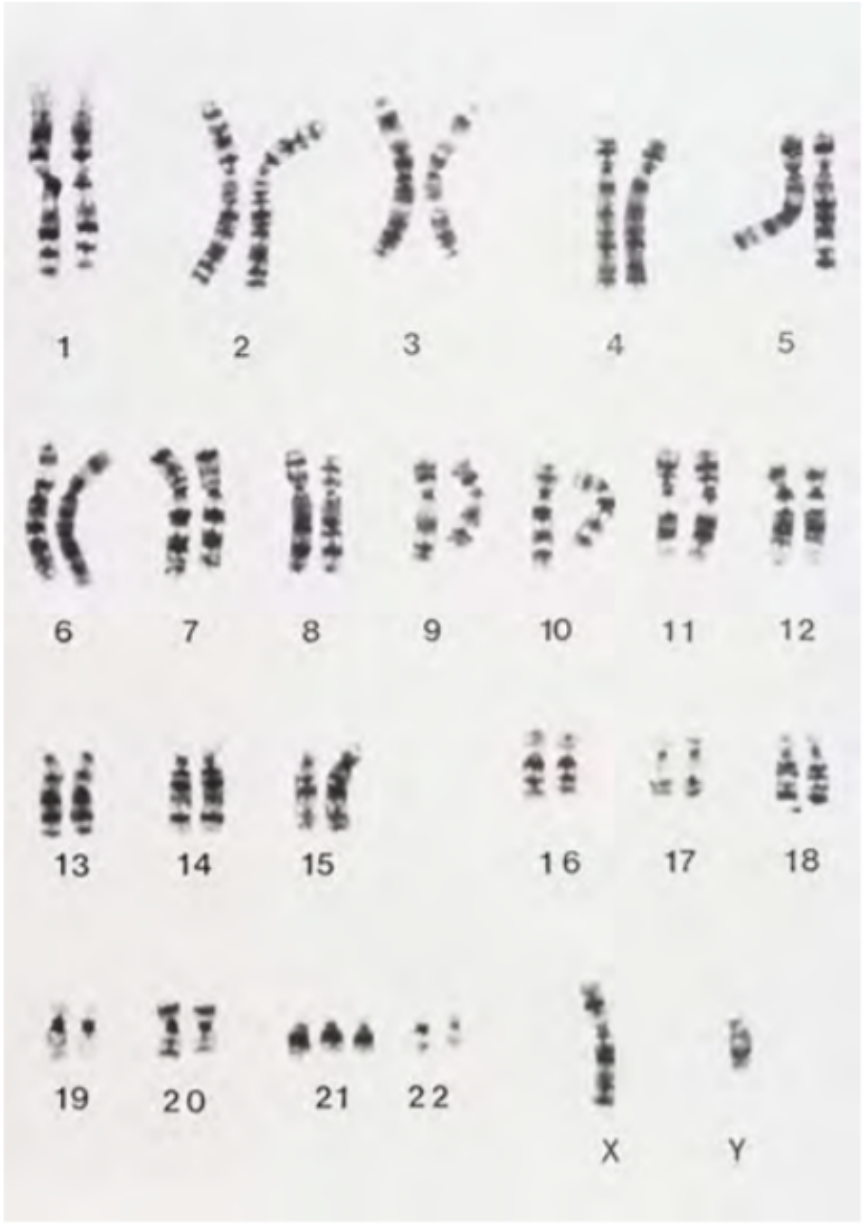
Aneuploidy

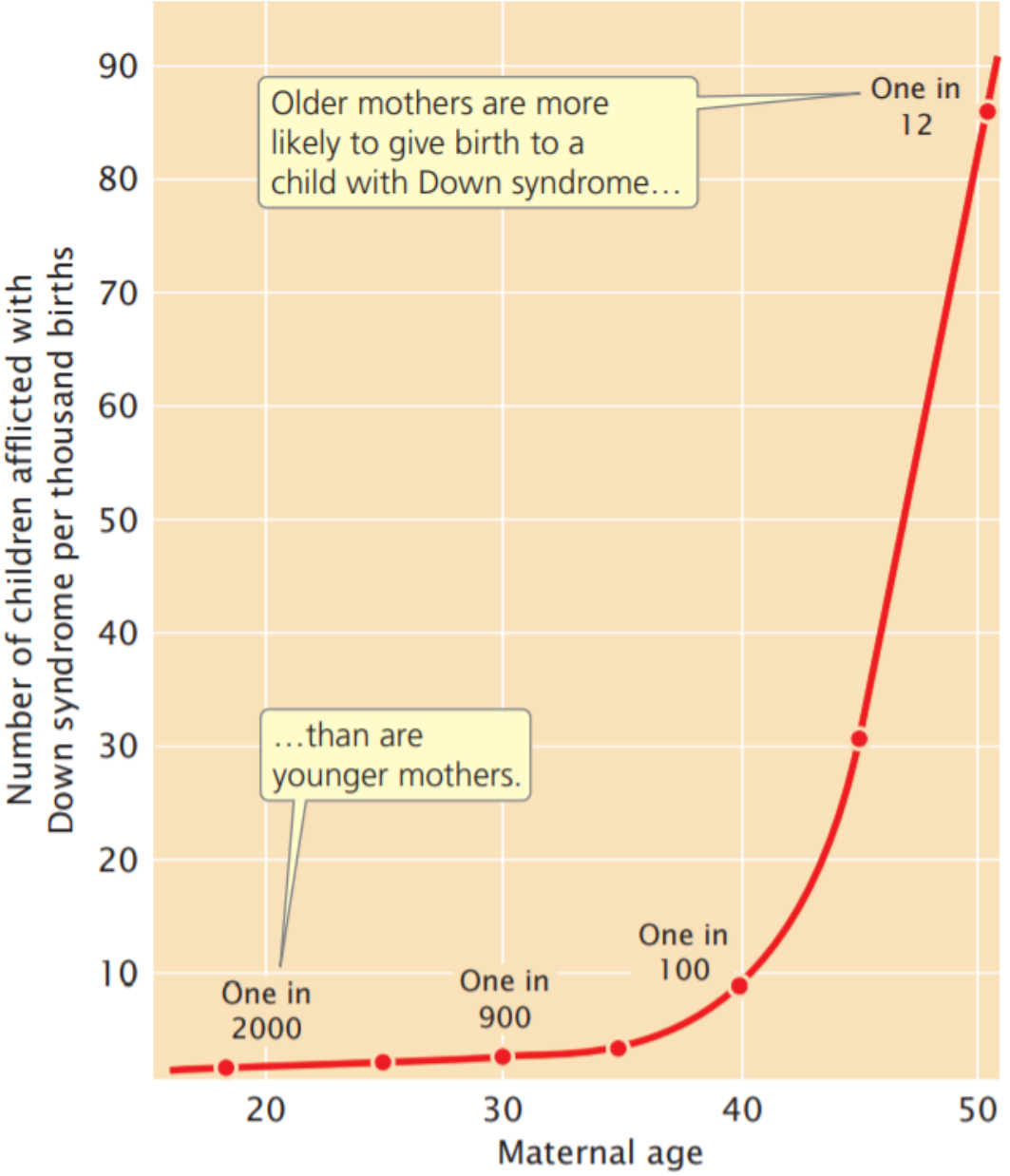




Griffiths -- *Introduction to Genetic Analysis*

Aneuploidy





Down syndrome

Pierce -- *Genetics Essentials. Concepts and Connections*

Aneuploidy

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,XYY		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

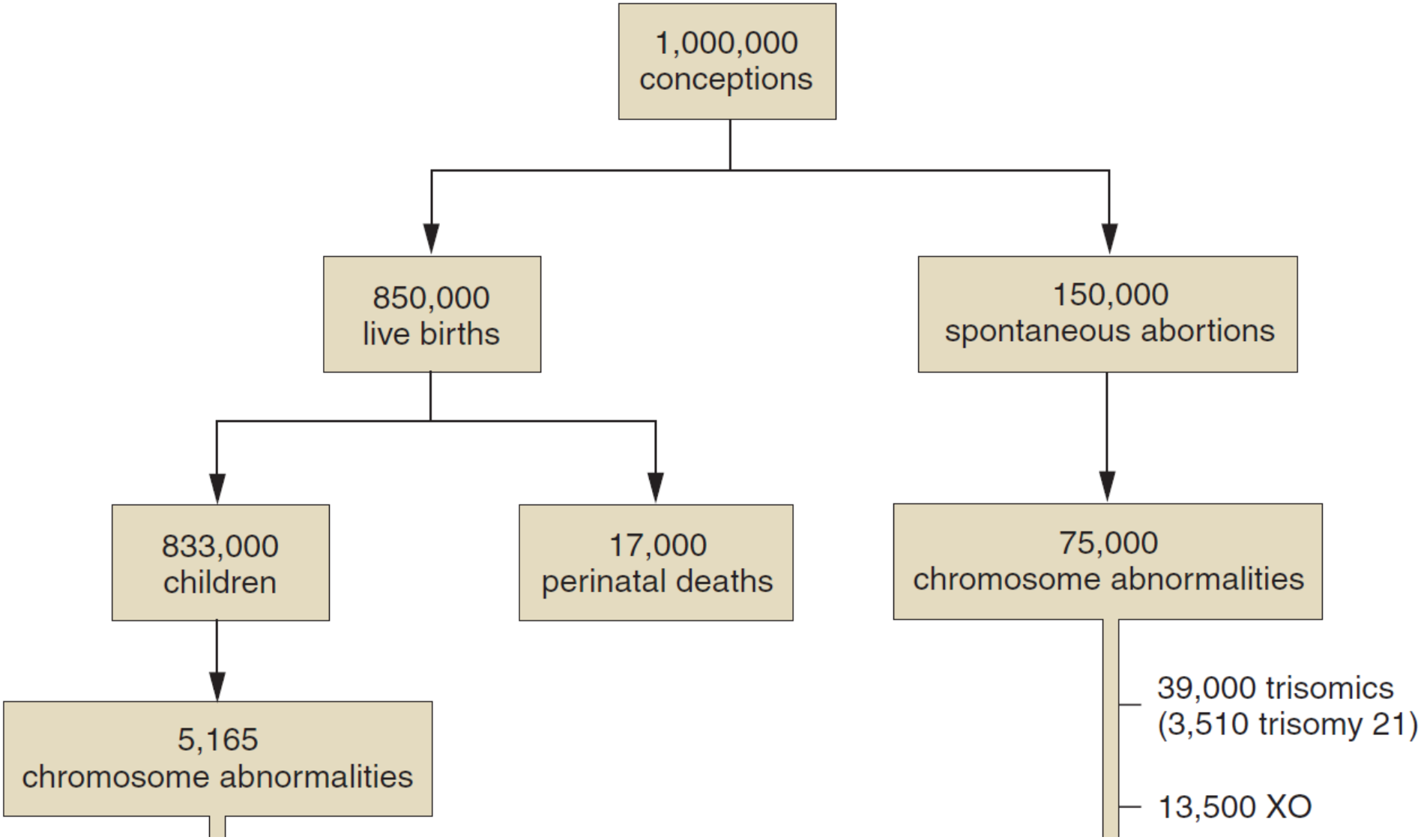
Jackson (2018) *Essays in Biochemistry*

X-inactivation

X-inactivation: in every cell in the female embryo, one of the two X chromosomes becomes inactivated and condensed.

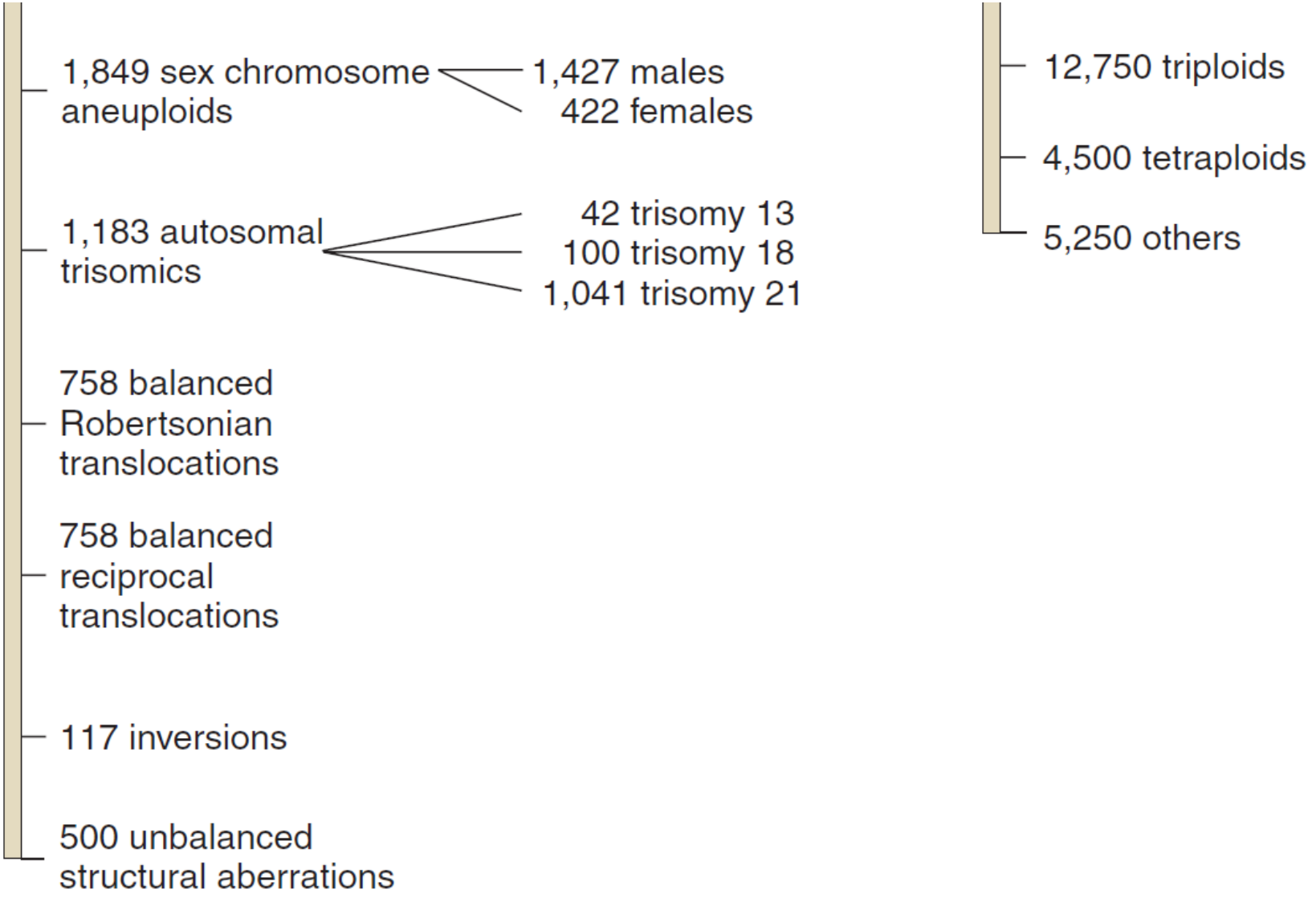
- Early in development
- Random in different cells
- Persists through subsequent cell divisions, but not generations
- Female tissues are a patchwork, with 50/50 inactivated paternal and maternal chromosomes
- X-inactivation provides **dosage compensation**: comparable levels of expresssion for ~1,500 X-chromosome genes in males and females

The fates of a 1 mln implanted human zygotes



Griffiths -- *Introduction to Genetic Analysis*

The fates of a 1 mln implanted human zygotes



Griffiths -- *Introduction to Genetic Analysis*

De novo variants rates and counts

DNM type	*Rate per generation*	*Total in an individuum*
Single nucleotide variants (SNVs)	**$1.20 \cdot 10^{-8}$** per bp **$1.66 \cdot 10^{-8}$** ‡	44..82
Dinucleotide repeats	$2.73 \cdot 10^{-4}$ per locus	N/A
Coding SNVs	N/A	1-2
Small indels (<50bp)	**$0.53\text{-}1.5 \cdot 10^{-9}$** per bp $1.26 \cdot 10^{-9}$ ‡	3-9
Large indels	0.16	0.16
Copy number variants (CNVs)	$10^{-6} – 10^{-4}$ per locus per generation	0.0154

‡ Ref: Palamara (2015) *AJHG*

De novo variants rates and counts

DNM type	*Rate per generation*	*Total in an individuum*
Single nucleotide variants (SNVs)	**$1.20 \cdot 10^{-8}$** per bp **$1.66 \cdot 10^{-8}$** ‡	44..82
Dinucleotide repeats	$2.73 \cdot 10^{-4}$ per locus	N/A
Coding SNVs	N/A	1-2
Small indels (<50bp)	**$0.53\text{-}1.5 \cdot 10^{-9}$** per bp $1.26 \cdot 10^{-9}$ ‡	3-9
Large indels	0.16	0.16
Copy number variants (CNVs)	$10^{-6} – 10^{-4}$ per locus per generation	0.0154

‡ Ref: Palamara (2015) *AJHG*

Summary

- Several mechanisms ensure the high rate of accuracy in DNA replication, including precise nucleotide selection, proofreading, and mismatch repair
- However, mutations are inevitable due to spontaneous replication errors and endogenous and exogenous DNA damage
- Human mutation rate is a trade-off between extinction and need for evolutionary change
- There is a wide spectrum of de novo mutations with varying rates and consequences: single nucleotide variants, structural variants and aneuploidies

Further reading

- 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.