Genomics in the Practice of Medicine and Public Health

Bruce R. Korf, MD, PhD

Outline

- Genetic architecture of disease
- Genetic variation
- Paradigm of genomic medicine
- Diagnosis of rare disorders
- Genomics of common disorders
- Cancer genomics
- Pharmacogenetics
- Genomics in public health

Genetic Architecture of Disease
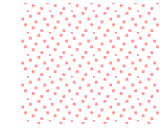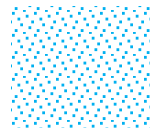
Genetic Architecture of Disease

monogenic
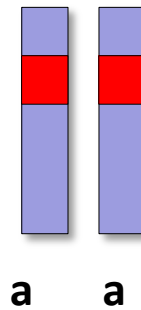large effect size
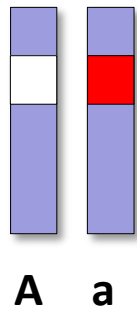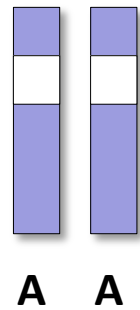
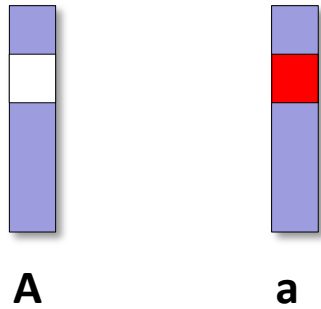polygenic
genes contributing to tall stature

genes contributing to short stature

small effect size
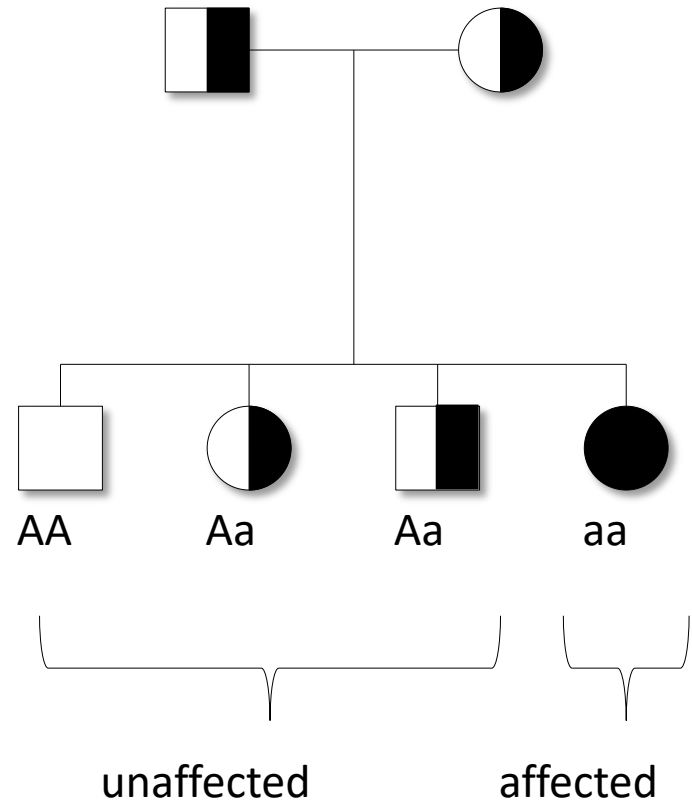
Autosomal Recessive

Autosomal Dominant

Genotype	homozygous	heterozygous	homozygous
Phenotype	unaffected	affected	affected

X-linkage

Male

A Y
unaffected

a Y
affected

Female

A A
unaffected

A a
unaffected

a a
affected
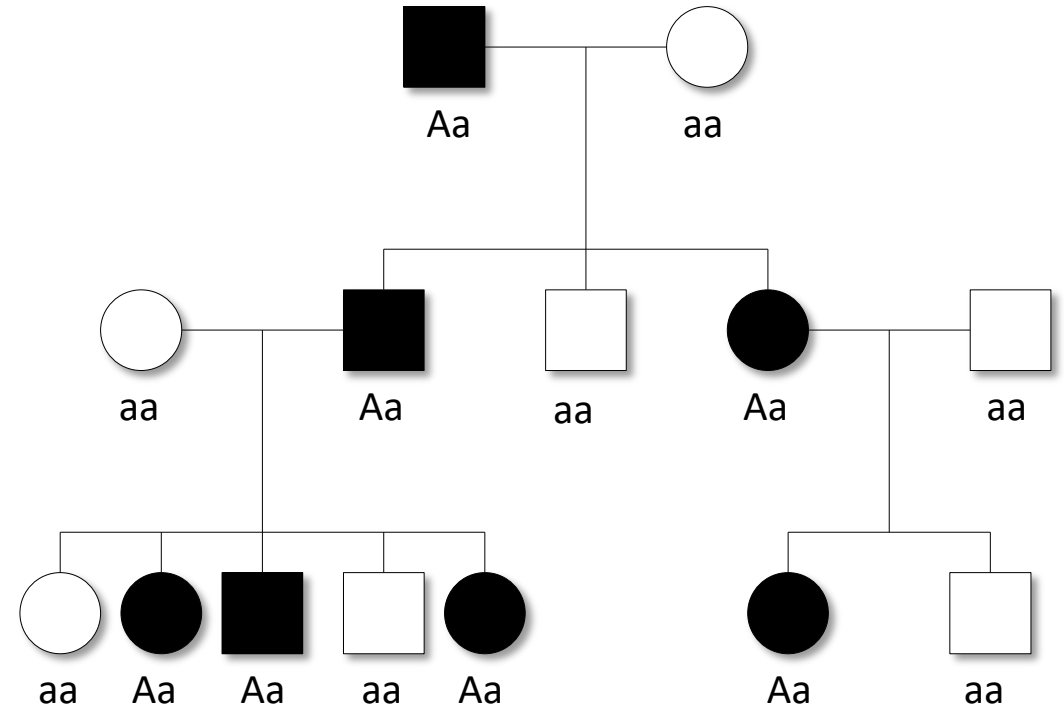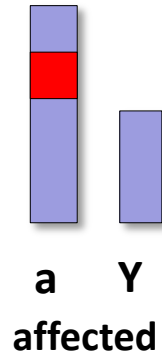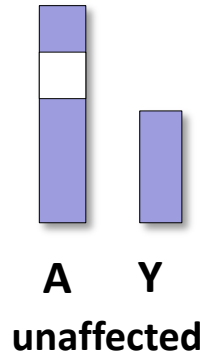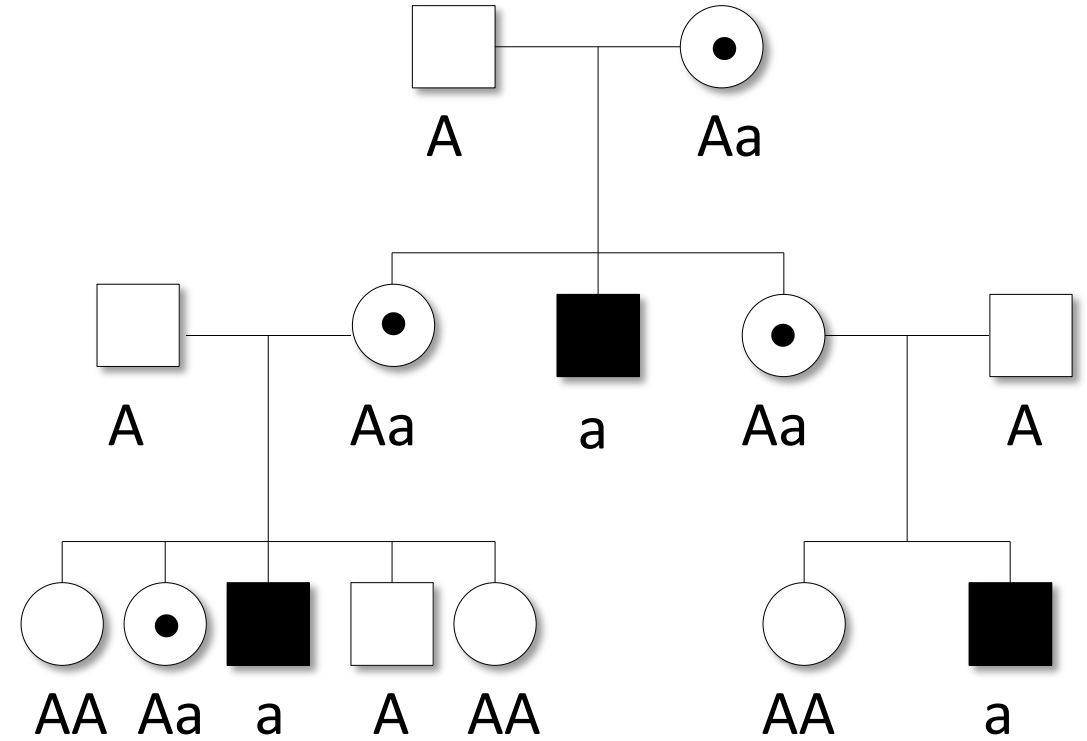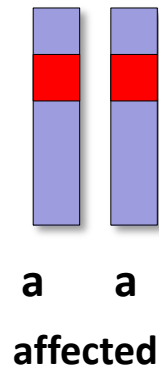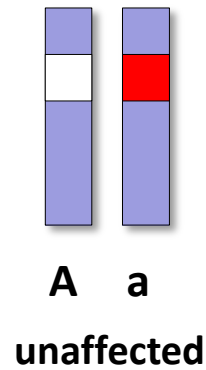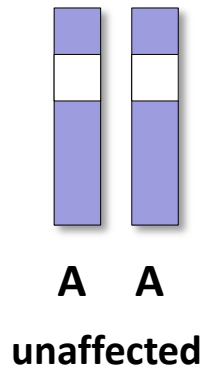
No male-to-male transmission

Penetrance

Fraction of individuals who carry a genetic variant who manifest a specified phenotype
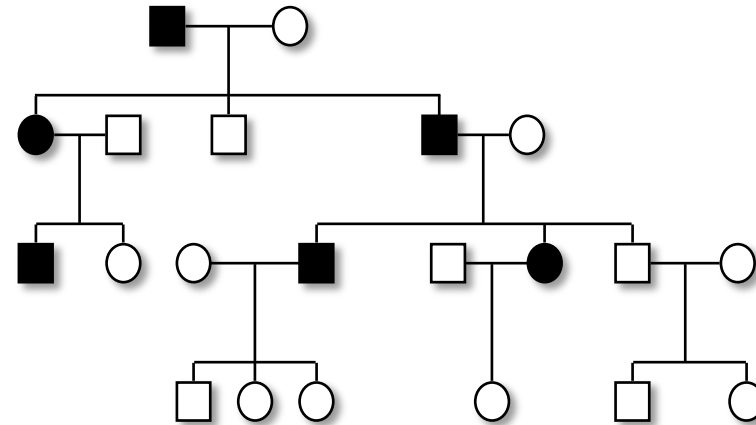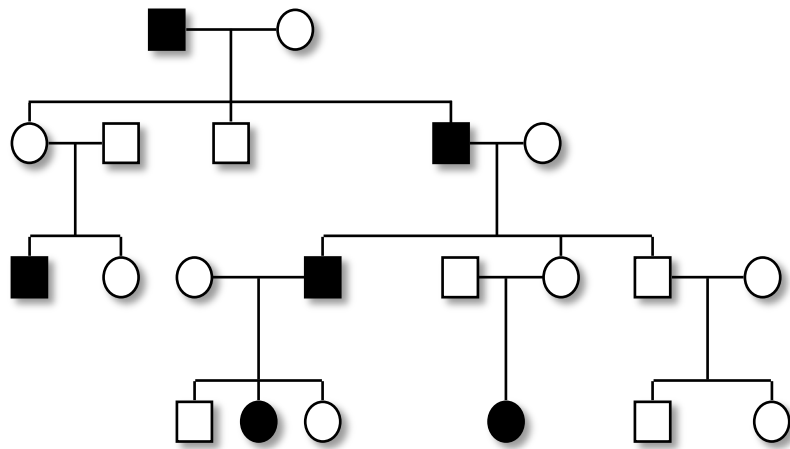
Expressivity

Different modes or degrees of expression of trait in different individuals

cutaneous neurofibromas in NF1

Variability in "single gene" disorders

- Allelic heterogeneity
- Genetic modifiers
- Environmental effects
- Chance
- Mosaicism

Allelic Heterogeneity

gene

Compound
Heterozygote
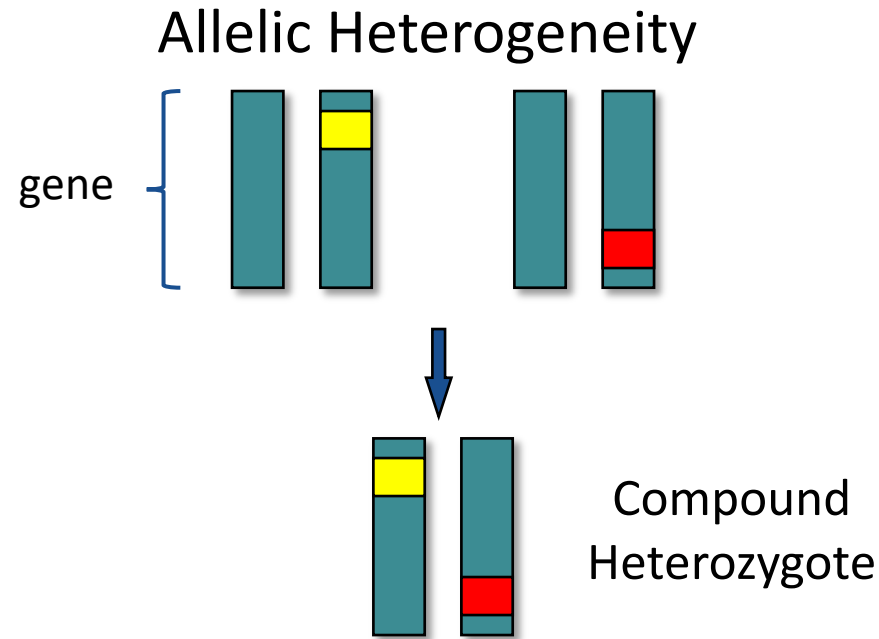
Mosaicism
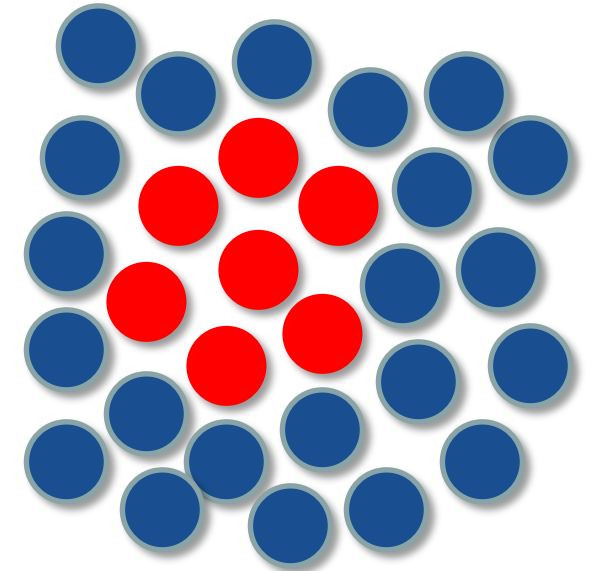
- Gonadal
- Somatic

Mitochondrial Inheritance

- 16.5 kb circular double stranded DNA

- Multiple copies per mitochondrion
  - Heteroplasmy – mixture of mitochondria with different genotypes in same cell

- 13 subunits of mitochondrial proteins, tRNAs, rRNAs

- Maternal Transmission

- Most mitochondrial proteins encoded in nucleus

Human Genetics and Genomics, Fourth Edition. Bruce R. Korf and Mira B. Irons.
© 2013 John Wiley & Sons, Ltd. Published 2013 by John Wiley & Sons, Ltd.

Human Genetics and Genomics, Fourth Edition. Bruce R. Korf and Mira B. Irons.
© 2013 John Wiley & Sons, Ltd. Published 2013 by John Wiley & Sons, Ltd.
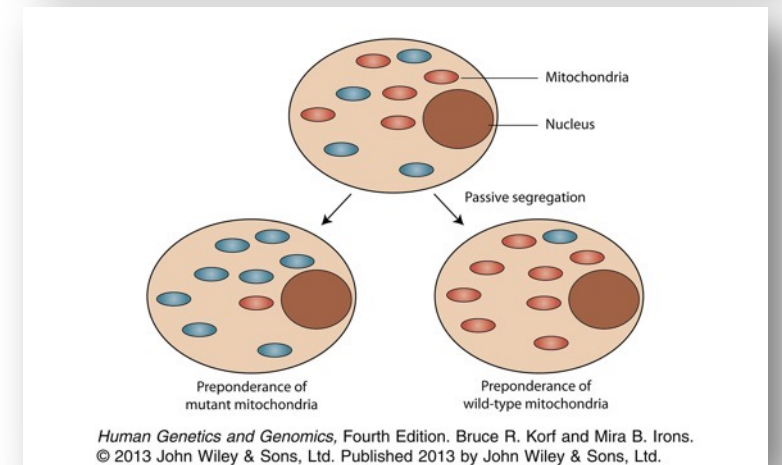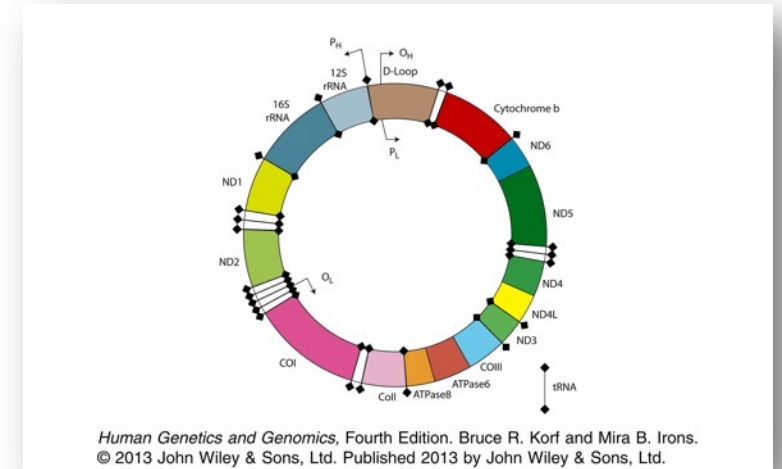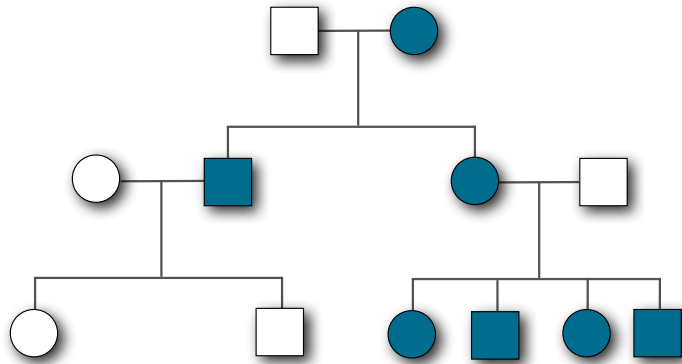
Multifactorial Inheritance

- Trait clusters in family
- Increased concordance in identical twins
- Multiple genetic and/or non-genetic factors

Genetic Variation

DNA Structure

- **DNA**: backbone of sugars (deoxyribose) and phosphate residues
  - Attached to sugar carbon at 1′ : a nitrogenous base:
    - purine (A and G)
    - pyrimidine (C and T)
    - Nucleotide: + phosphate group attached to carbon atom 5′ (mono-, di-, tri-P)
- **RNA**
  - ribose instead of deoxyribose
  - uracil instead of thymine
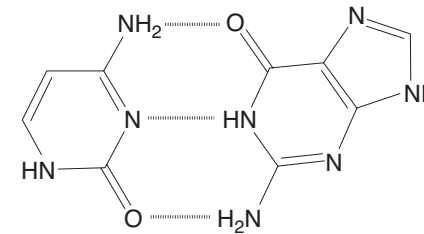
Cytosine   Guanine

Thymine   Adenine

deoxyadenosine monophosphate

3′   5′

sugar-phosphate helix

3′   5′
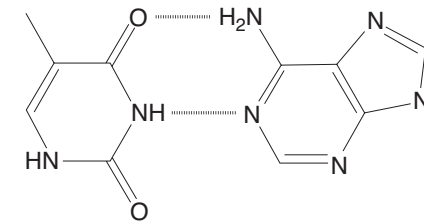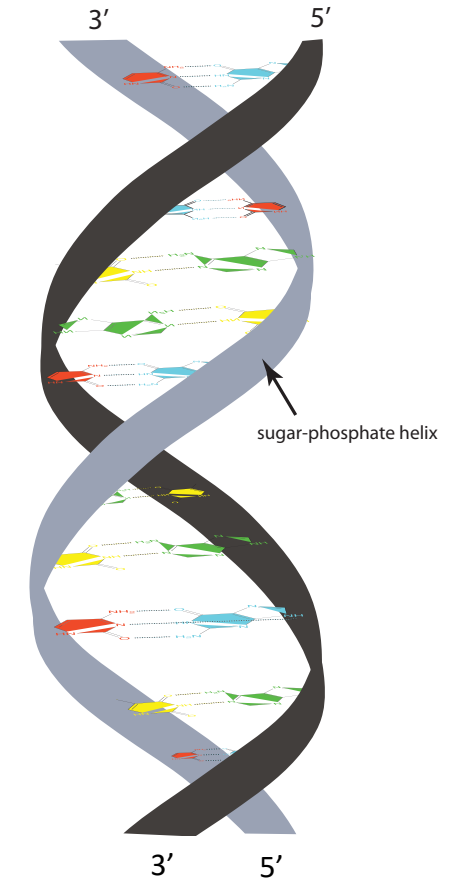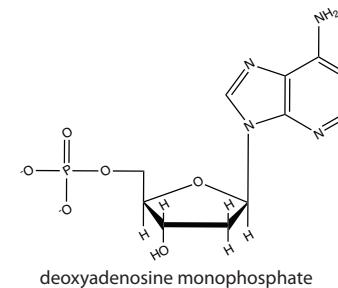
DNA Replication

DNA Repair

Nucleotide Excision Repair

thymidine dimer → cut by endonuclease → bases removed by helicase → new bases added / ligase

C G A G T T C A G C A
| | | | | | | | | | |
G C T C A A G T C G T

Base Excision Repair

aberrant DNA base → base removed by glycosylase → sugar phosphate removed by endonuclease → replace base and ligate

C G A G T T U A G C A
| | | | | | | | | | |
G C T C A A G T C G T

Mismatch Repair

MutS protein / MutL protein → remove newly synthesized segment → fill in gap and ligate

base mismatch / nick in DNA strand

newly synthesized strand

Basic Gene Structure

- Transcription factors (*trans*-acting) bind in promotor region to specific short sequence elements
- *Cis*-acting: TATA box (-25) or CCAAT box (-80); GC box (housekeeping); other enhancers and silencers
- Binding of RNA polymerase II and initiation of the transcription

Transcription

- RNA polymerases use transcription units to synthesize the primary transcripts

- Anti-sense strand is transcribed , read in 3'-5' direction

- RNA grows in 5'-3' direction

- One gene can have more than one promotor → variability

RNA polymerase

5'   3'
3'   3'   5'
                                  direction of
                                  transcription

mRNA

5'

poly A addition downstream of AAUAAA

5' - 5' triphosphate bond

7-methylguanine

5'
mRNA

mRNA cap   3'

AAUAAA
poly A tail
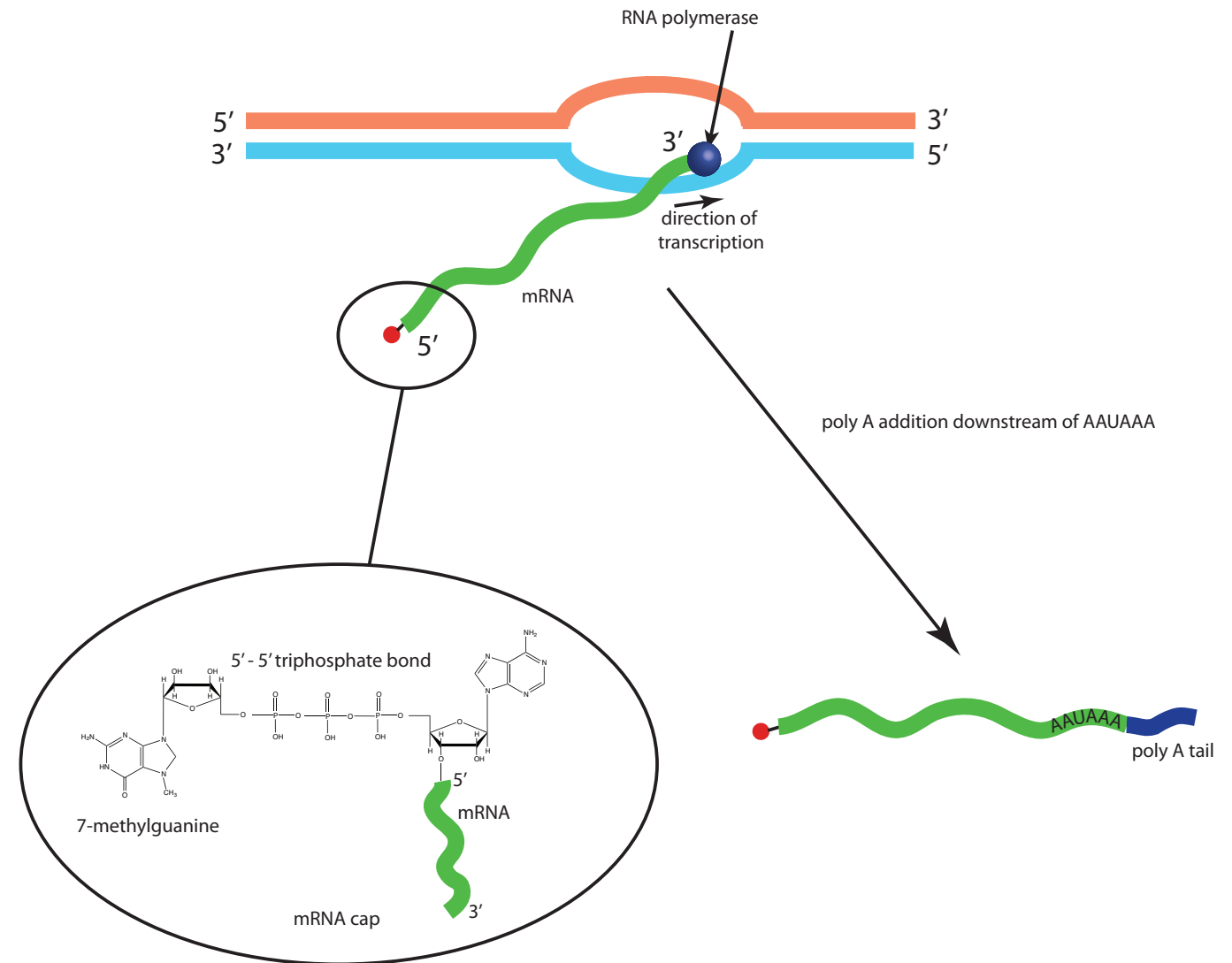
RNA Splicing
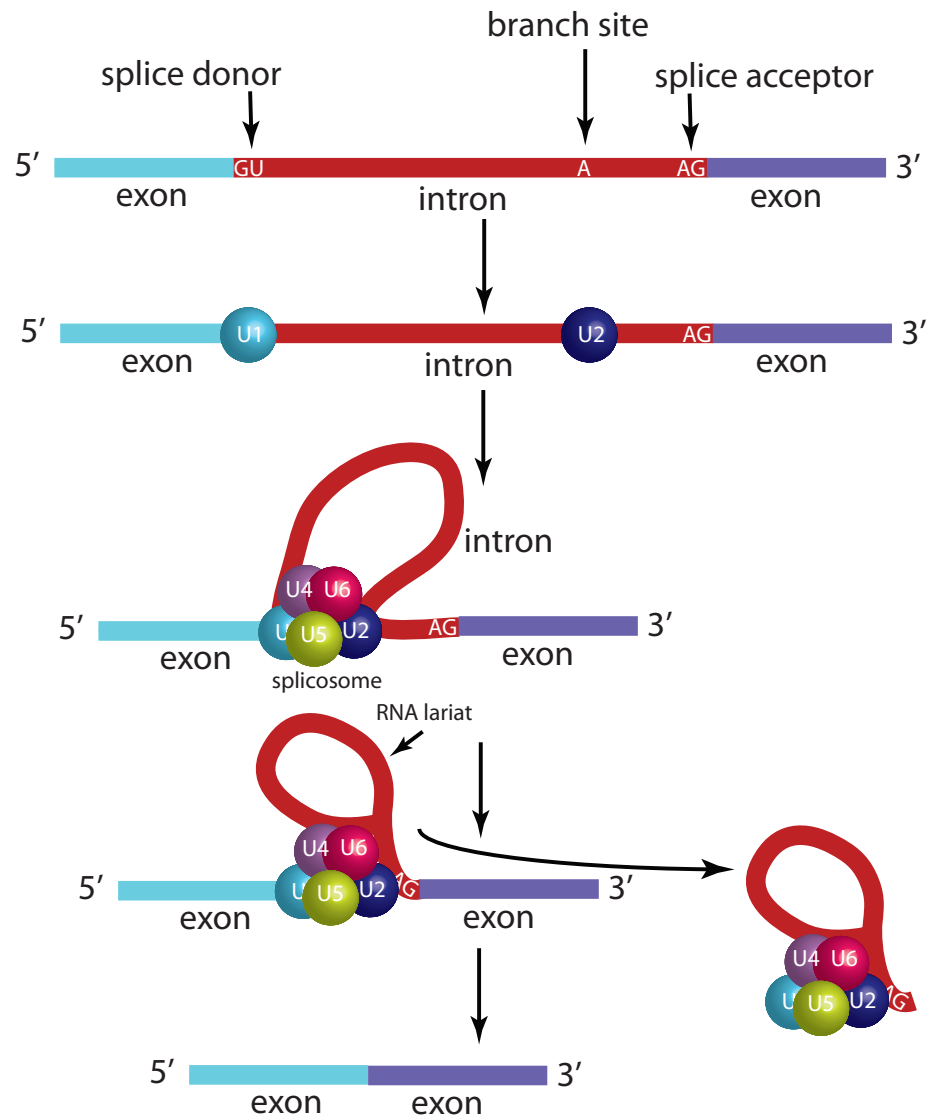
Translation

- Ribosomes read mRNA in 5'-3' direction

- Per codon (3 bases) 1 AA gets incorporated

- Codon recognized by the tRNA anti-codon

- Initiation codon: AUG (methionine
- 3 stop codons: UAA, UAG, UGA

Human Genome

98.5% 1.5%
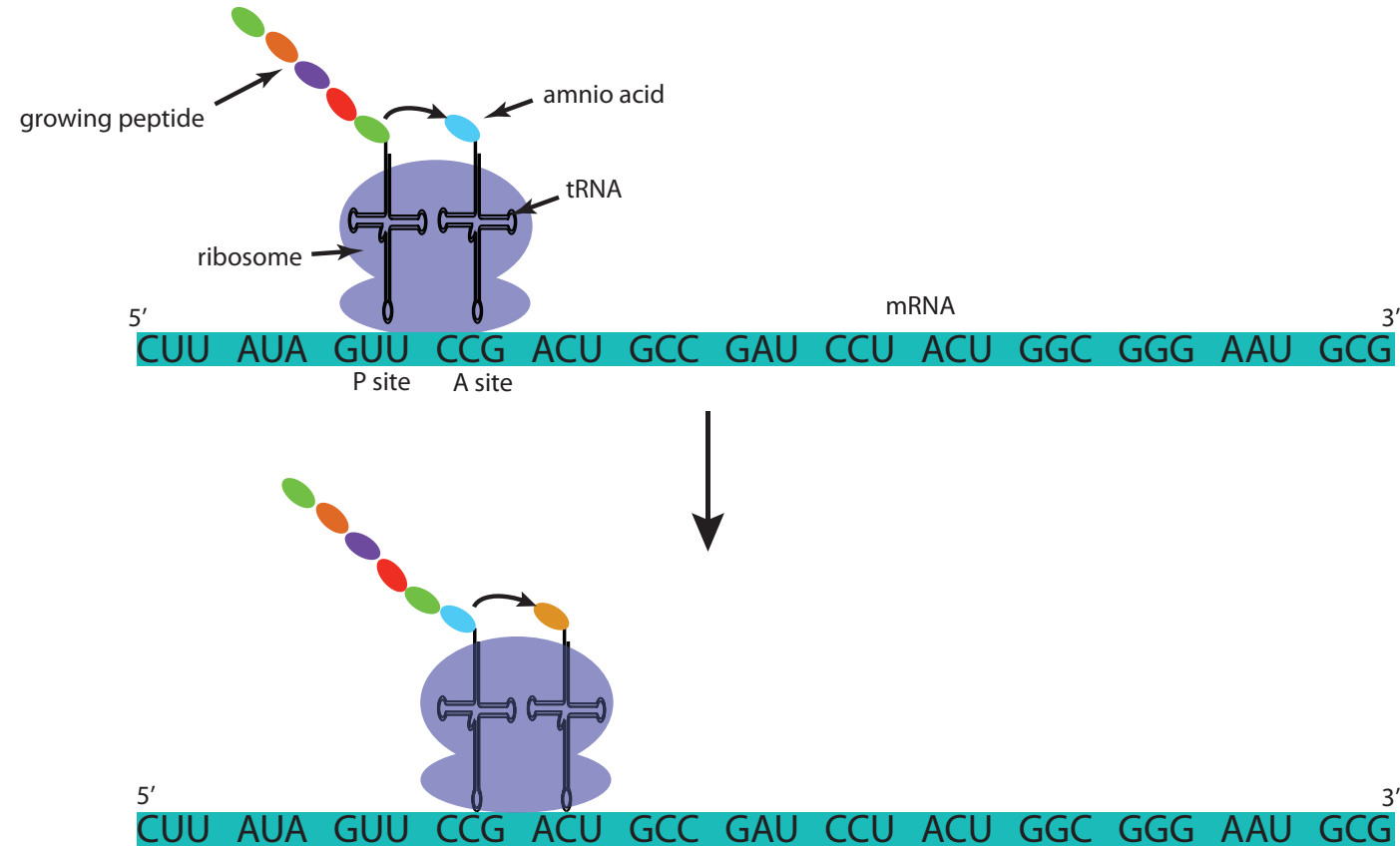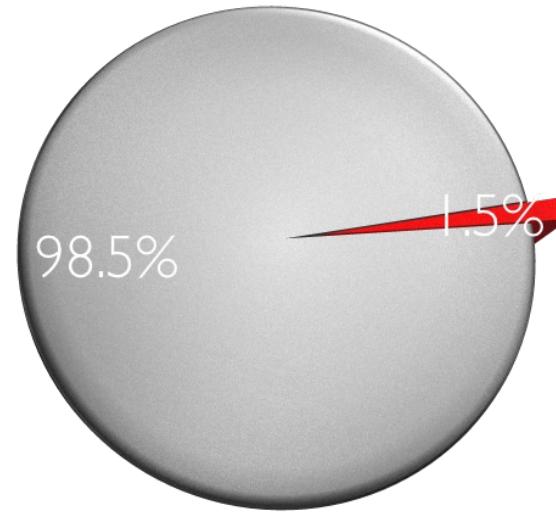
coding sequences

Repeated and non-coding sequences

- Interspersed repeats
- Pseudogenes
- Simple sequence repeats
- Segmental duplications
- Blocks of repeated sequences

- non-coding RNA's
- tRNA
- rRNA
- snRNA
- protein-encoding RNA's

Sickle Cell Variant

Normal Globin

Sickle Globin

α α

ß ß

glutamic acid
GAG

α α

ß S ß S
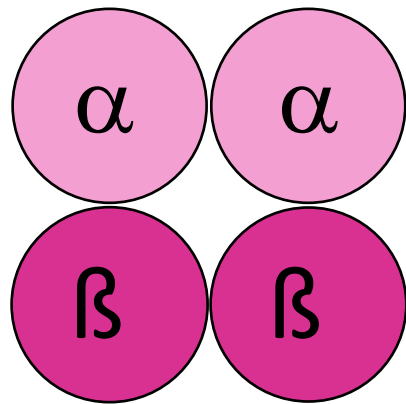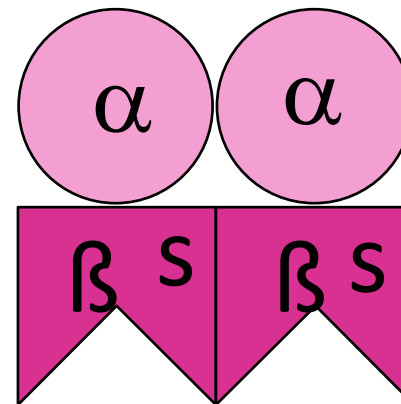
valine
GTG

Single Nucleotide Variants

TCC CAA ATC GTC CCT CGA GTT   wild type sequence
 ser gln ile val pro arg val

TCC CAG ATC GTC CCT CGA GTT   silent mutation
 ser gln ile val pro arg val

TCC CAA ATC CTC CCT CGA GTT   conservative mutation
 ser gln ile leu pro arg val

TCC CAA ATC GTC GCT CGA GTT   non-conservative mutation
 ser gln ile val ala arg val

TCC CAA ATC GTC CCT TGA GTT   stop mutation
 ser gln ile val pro stop

TCC CAG AAT CGT CCC TCG AGT T   frameshift mutation
 ser gln asn arg pro ser ser

Splicing Variant

splice acceptor mutation

cryptic splice acceptor

truncated exon

splice acceptor mutation

mutation creates new splice acceptor

cryptic splice acceptor

inclusion of intron in processed mRNA
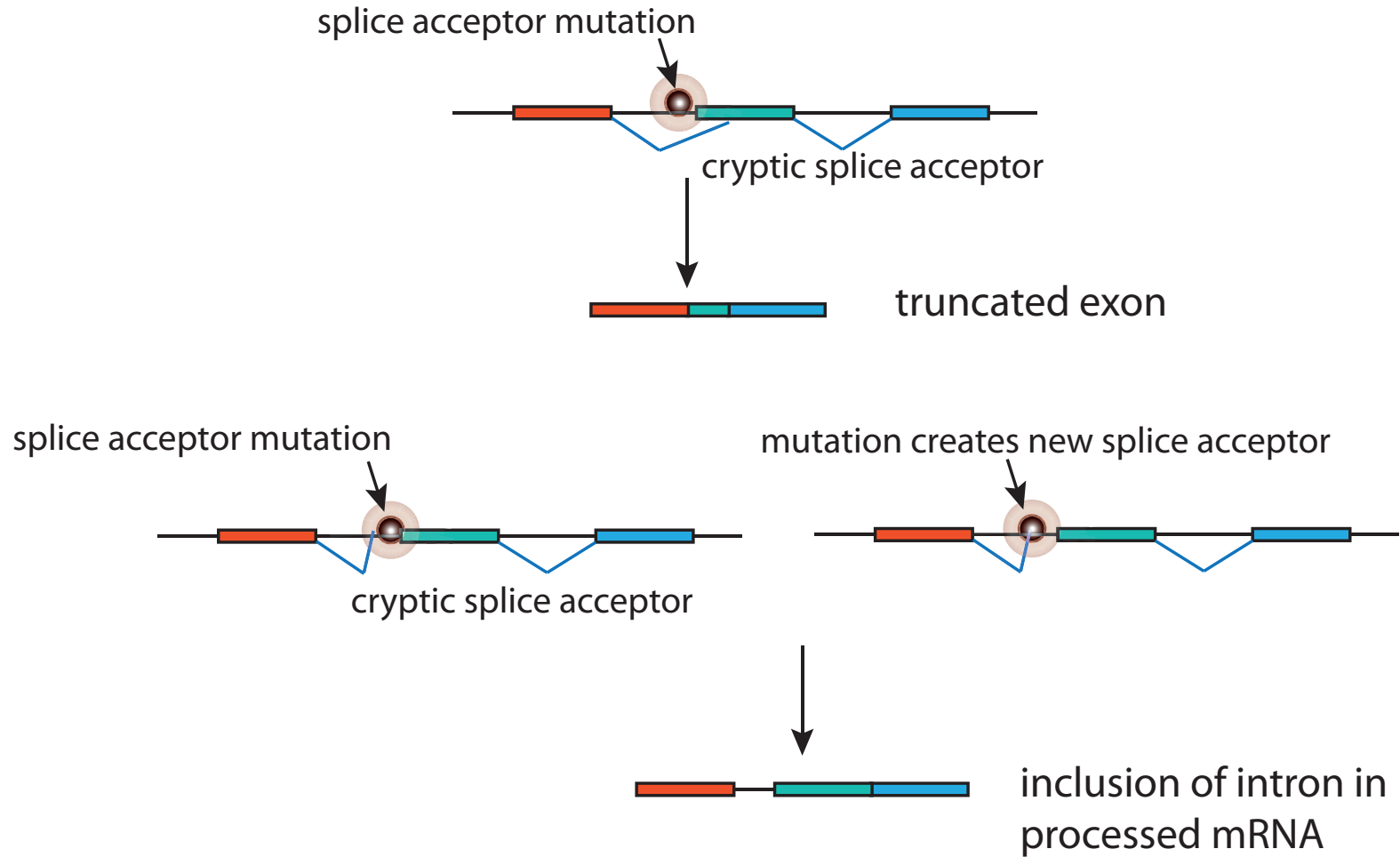
Copy Number Variant
e.g., Red-Green Color Blindness

Normal color vision

normal*

Red-green color blind
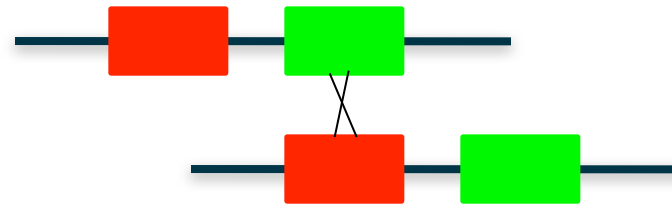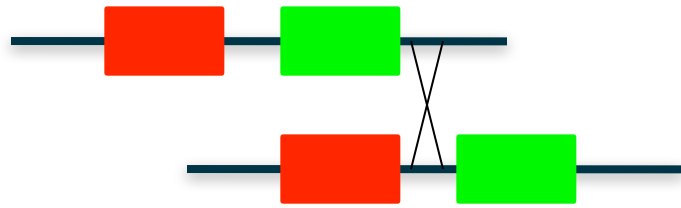
* Color vision may be abnormal if green gene not expressed

Chromosomal Variant
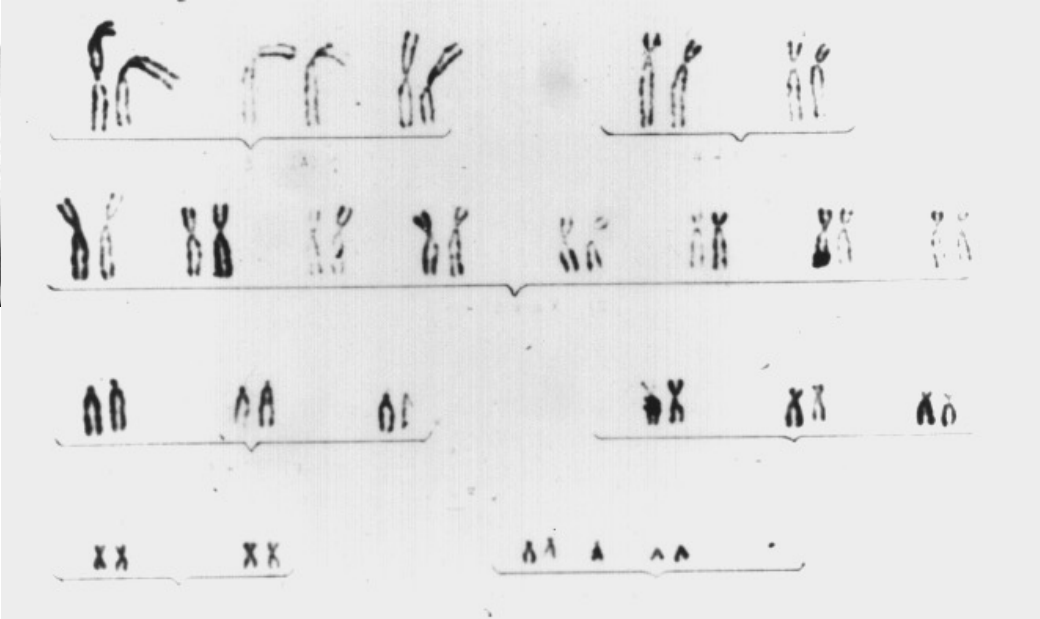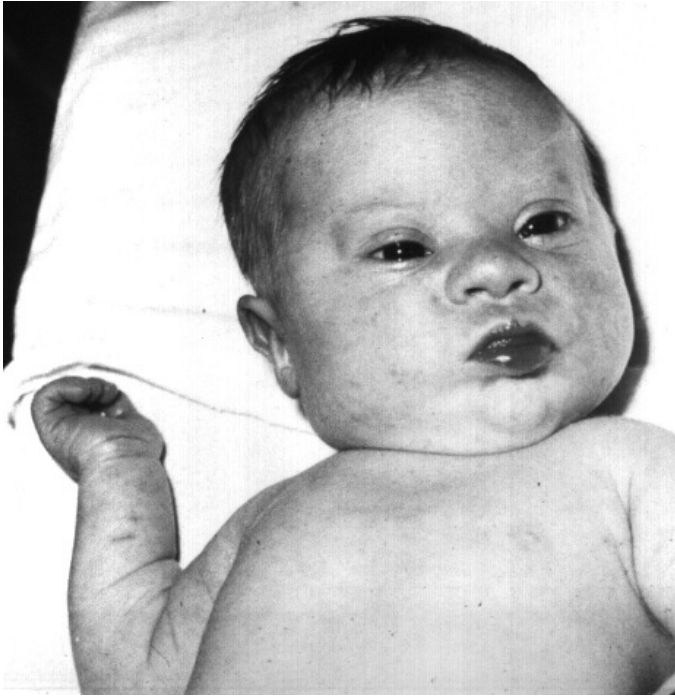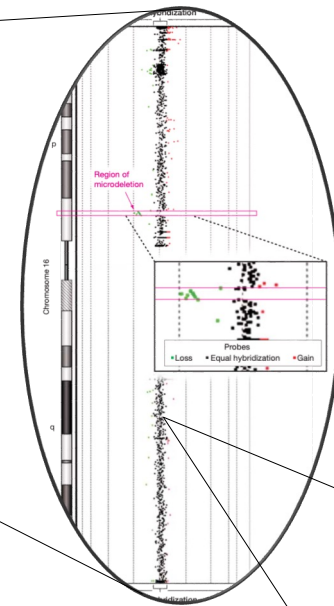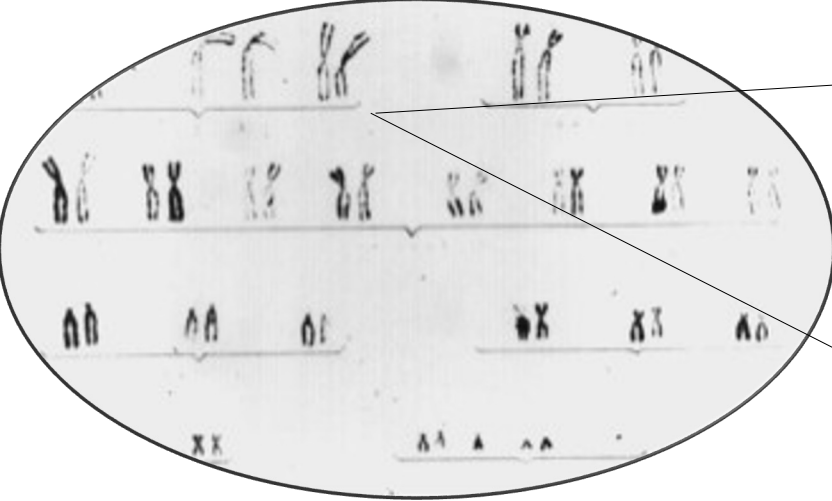
Cytogenomic Analysis
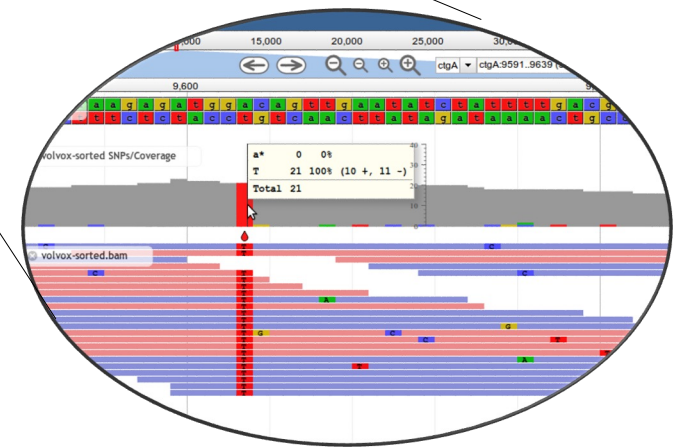
Genetic Variation is Common

- ~1 in 1,000 bases
- Hence ~6 million variants in ~6 billion nucleotide genome
- Variant types
  - Benign
  - Affect phenotype but do not result in disease
  - Disease-associated (major or minor effect)
- Variant frequencies may differ in different populations

Genomic Medicine

Genomics

"For the newly developing discipline of mapping/sequencing (including analysis of the information) we have adopted the term GENOMICS. We are indebted to T. H. Roderick of the Jackson Laboratory, Bar Harbor, Maine, for suggesting the term. The new discipline is born from a marriage of molecular and cell biology with classical genetics and is fostered by computational science."

(Victor A. McKusick and Frank H. Ruddle. A new discipline, a new name, a new journal [editorial]. Genomics 1987 Sep;1:1-2.)

Genomic Medicine

Scope	Focus	Target
• Rare Disease • Common Disease • Cancer	• Prevention • Diagnosis • Treatment	• Individual • Family • Population

Genomics of Rare Disorders
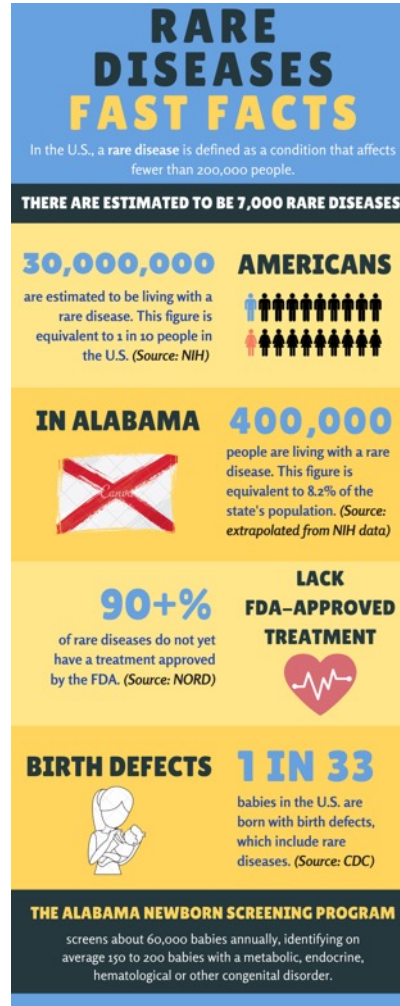
Genome Sequencing: Diagnosis of Rare Disorders
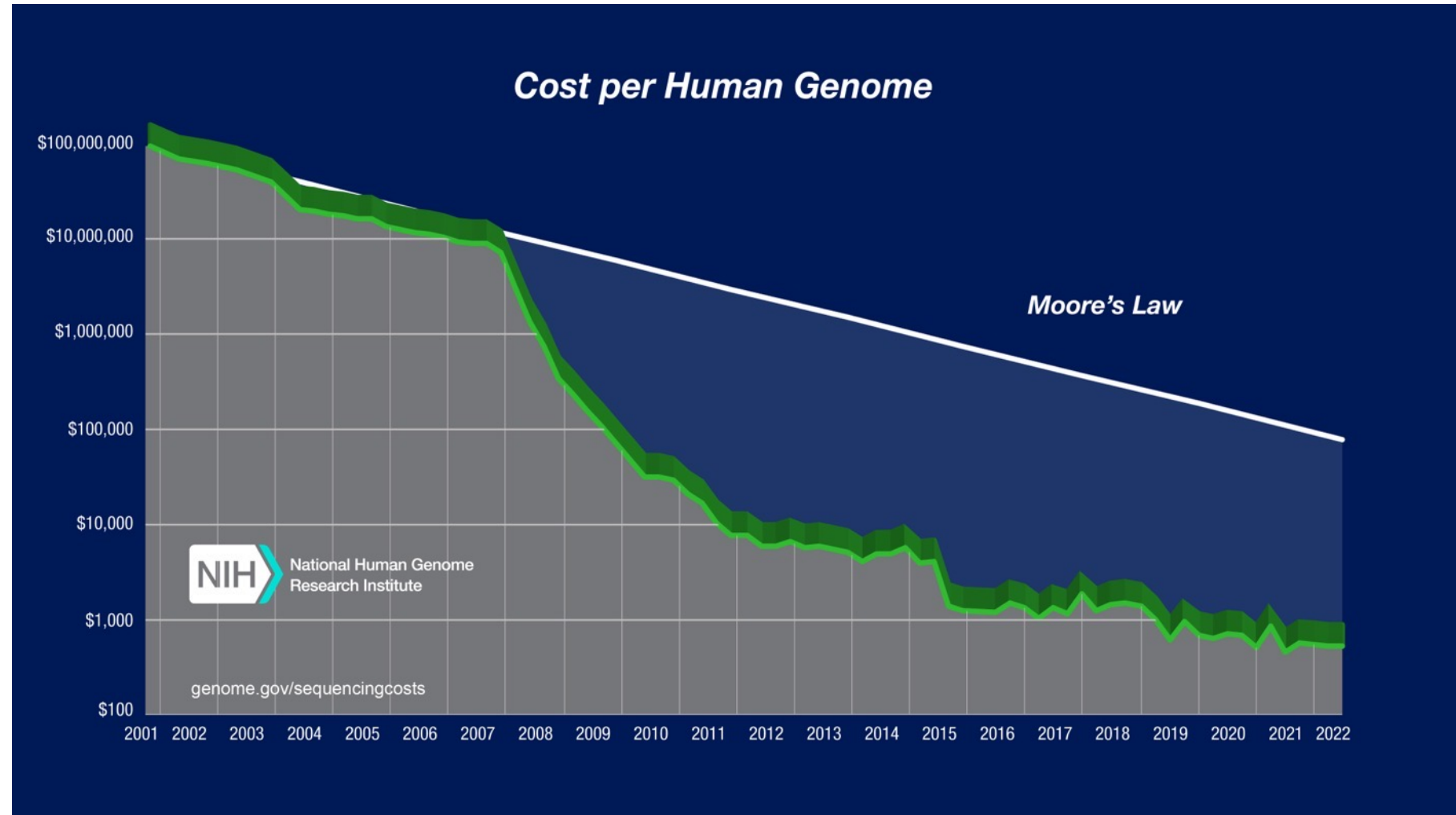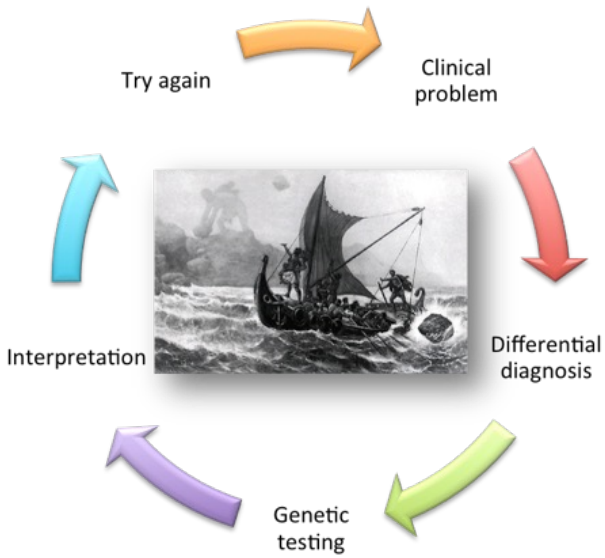
Bioinformatic Analysis

Function

Frequency

In silico Prediction

Candidate Variants

pathogenic

Likely pathogenic

Variant of unknown significance

Likely benign

Benign
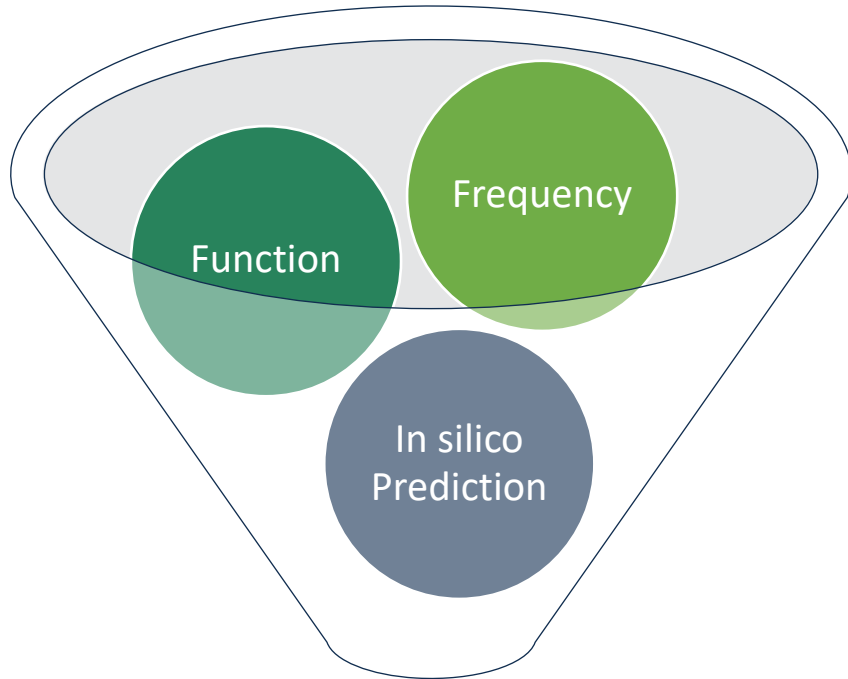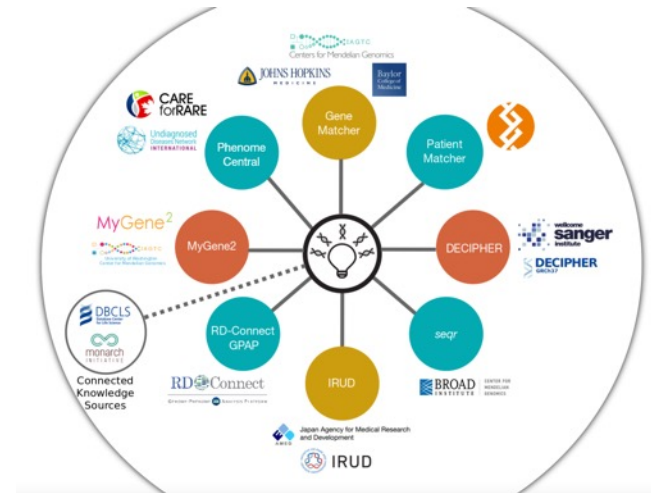
Treatment of Rare Genetic Disease (e.g., Tuberous Sclerosis Complex)

hamartin tuberin

Rheb GDP

mTOR

growth & proliferation

SEGA

Everolimus

Renal AML
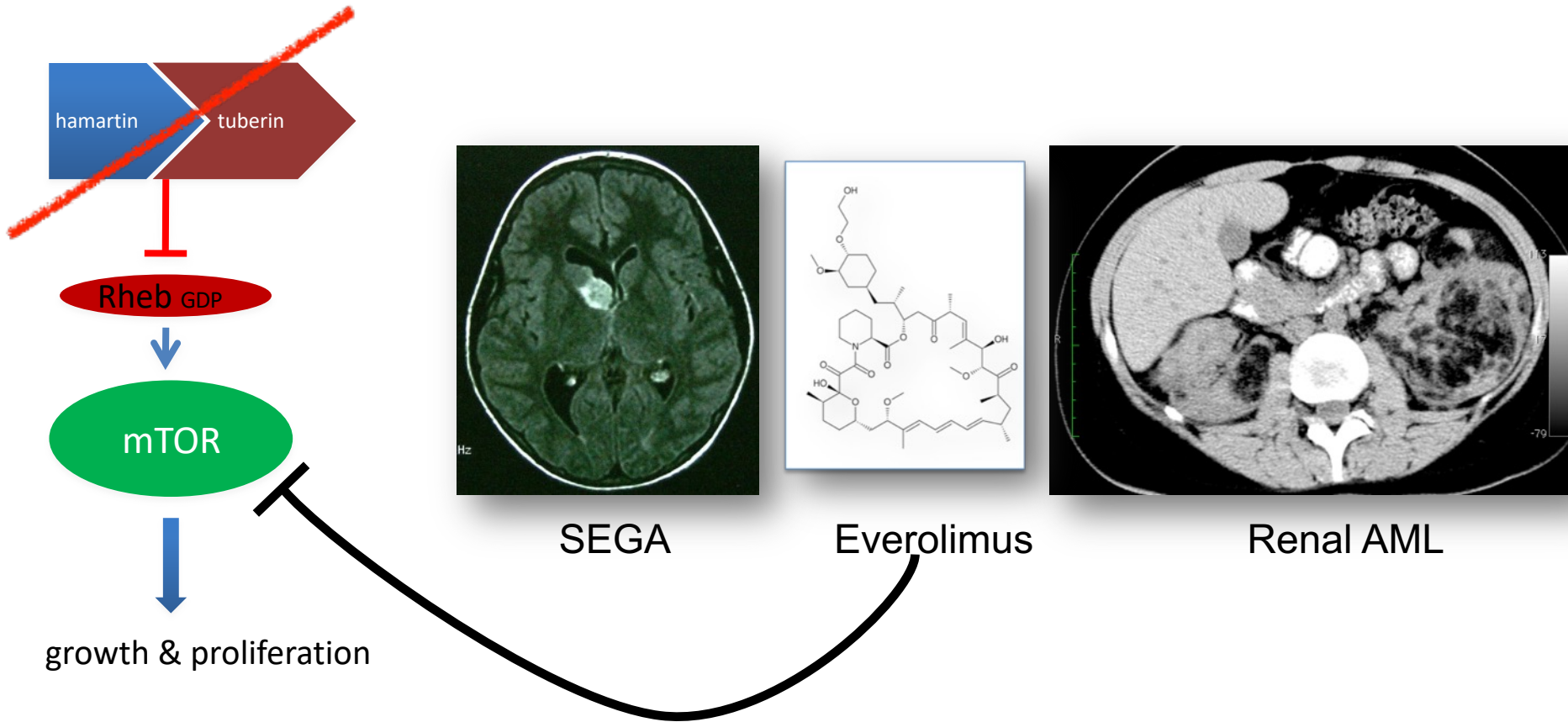
CRISPR/Cas9 Genome Editing

Guide RNA

Cas9

PAM sequence

Non-homologous
end joining

Homology-directed
repair

Mutation at site of
repair

Repair with
Donor DNA
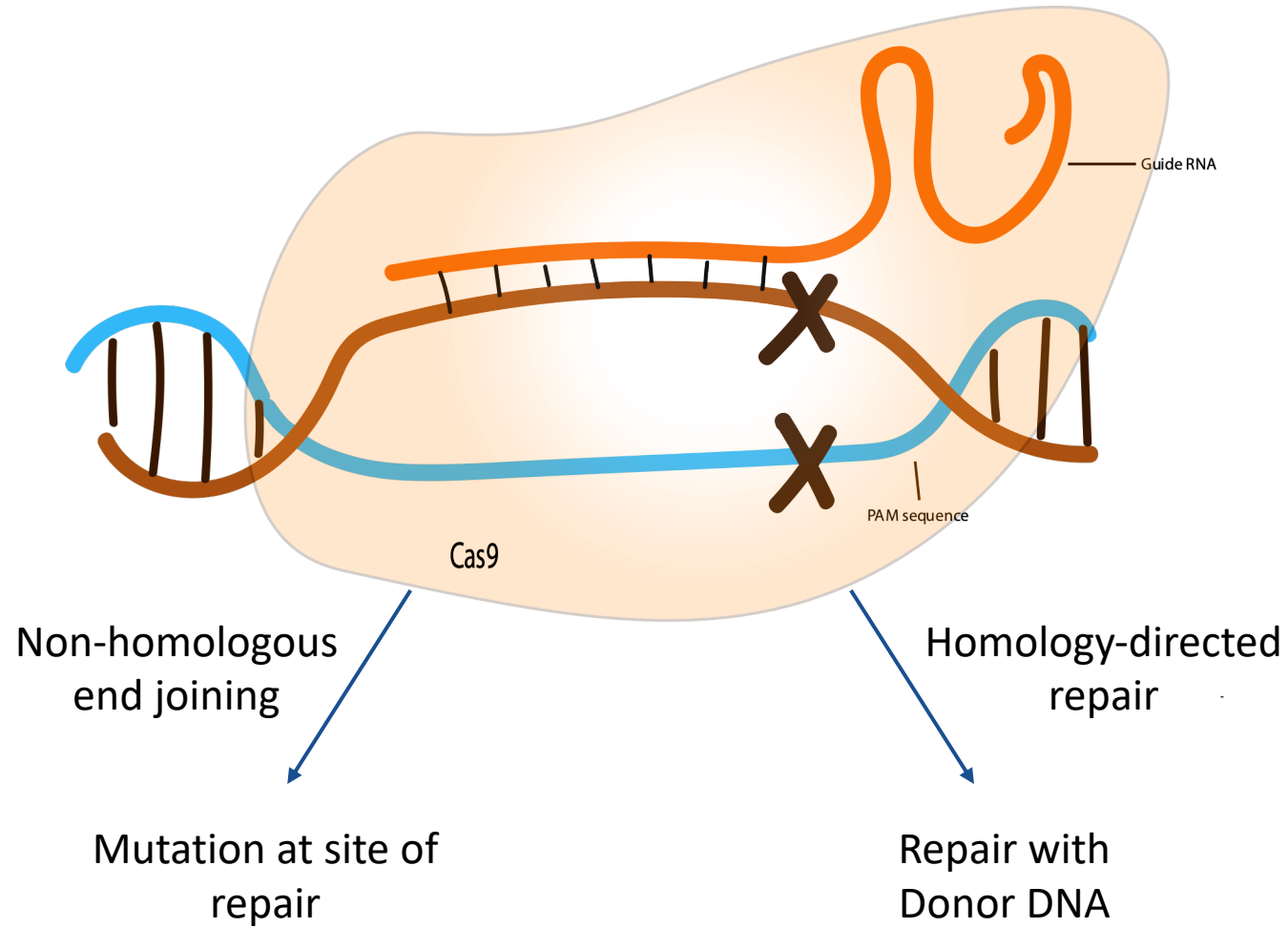
Genomics of Common Disorders

Common Disease: Case-Control Study
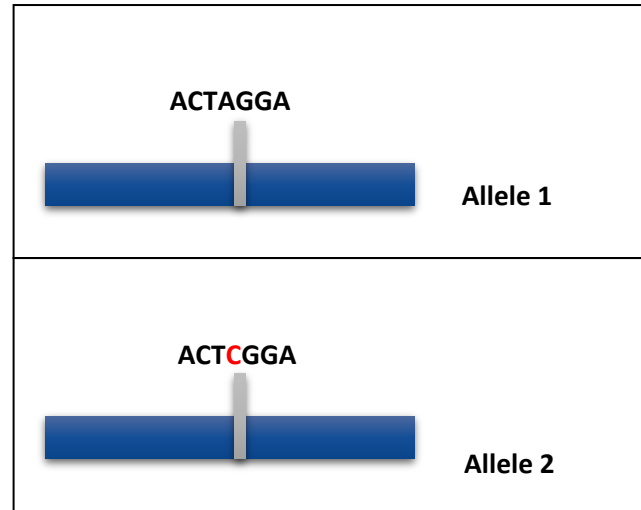
	Asthma	No Asthma
Allele 2 Present	3000	1000
Allele 2 Not Present	7000	9000

Hypothesis: Allele 2 is associated with an increased risk of asthma

Odds Ratio

	Asthma	No Asthma
Allele 2 Present	3000	1000
Allele 2 Not Present	7000	9000

Odds of disease given allele = 3000/1000 = 3
Odds of disease given not allele 2 = 7000/9000 = 0.78

Odds ratio = 3/0.78 = 3.85

Genome-Wide Association Study

GWAS

https://www.ebi.ac.uk/gwas/diagram
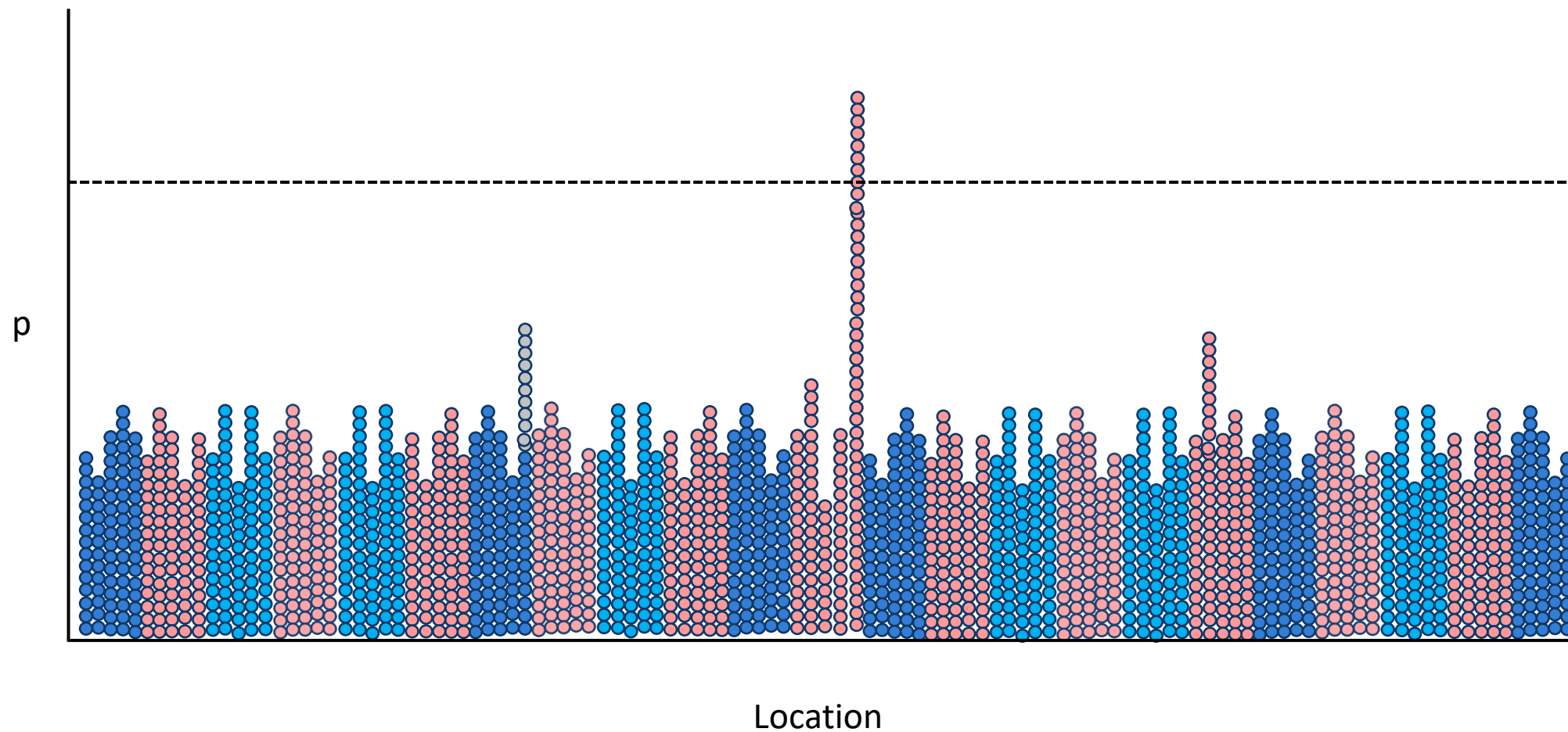
Missing Heritability
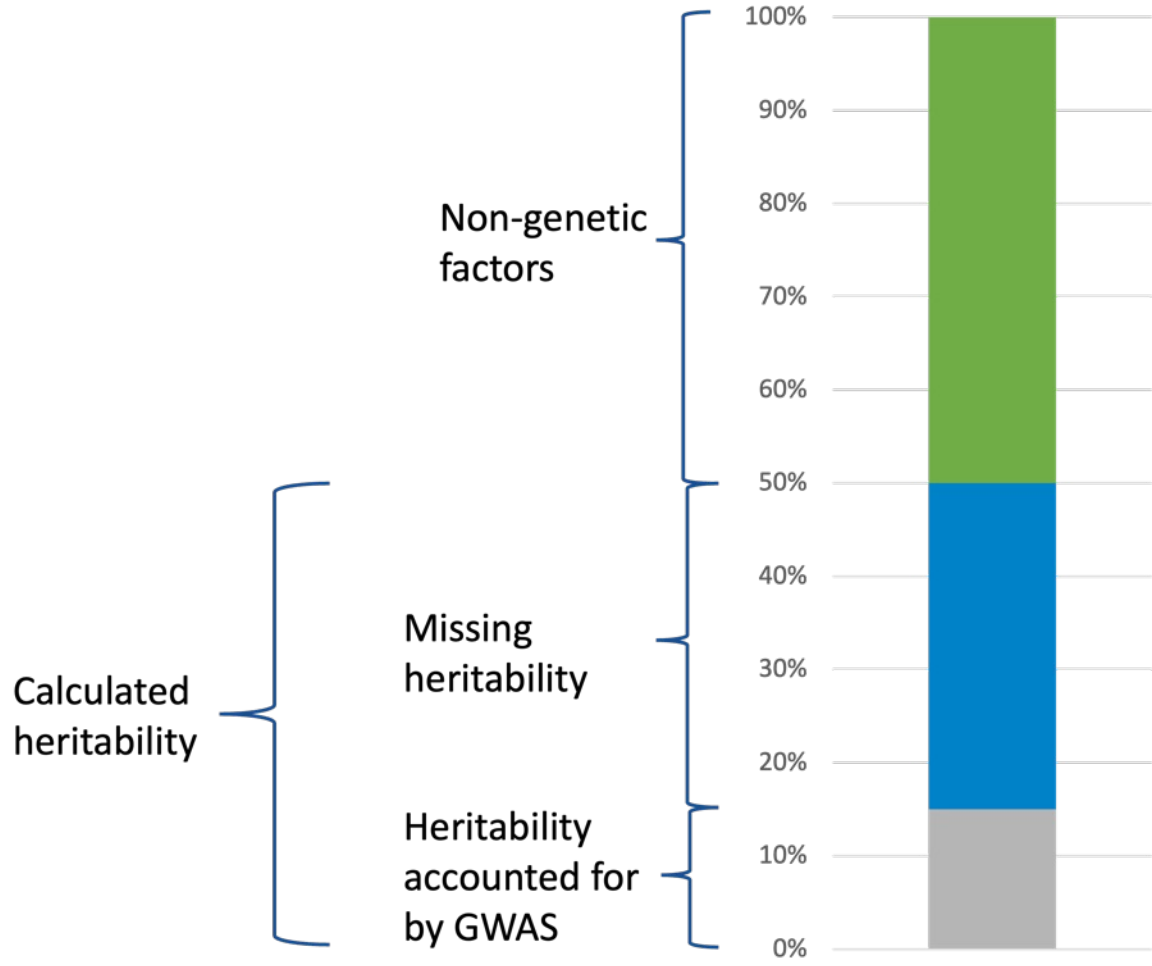
Possible Explanations

- Heritability overestimated
- Rare variants of small effect size
- Non-detected variants (e.g., CNVs)

Genetic Architecture
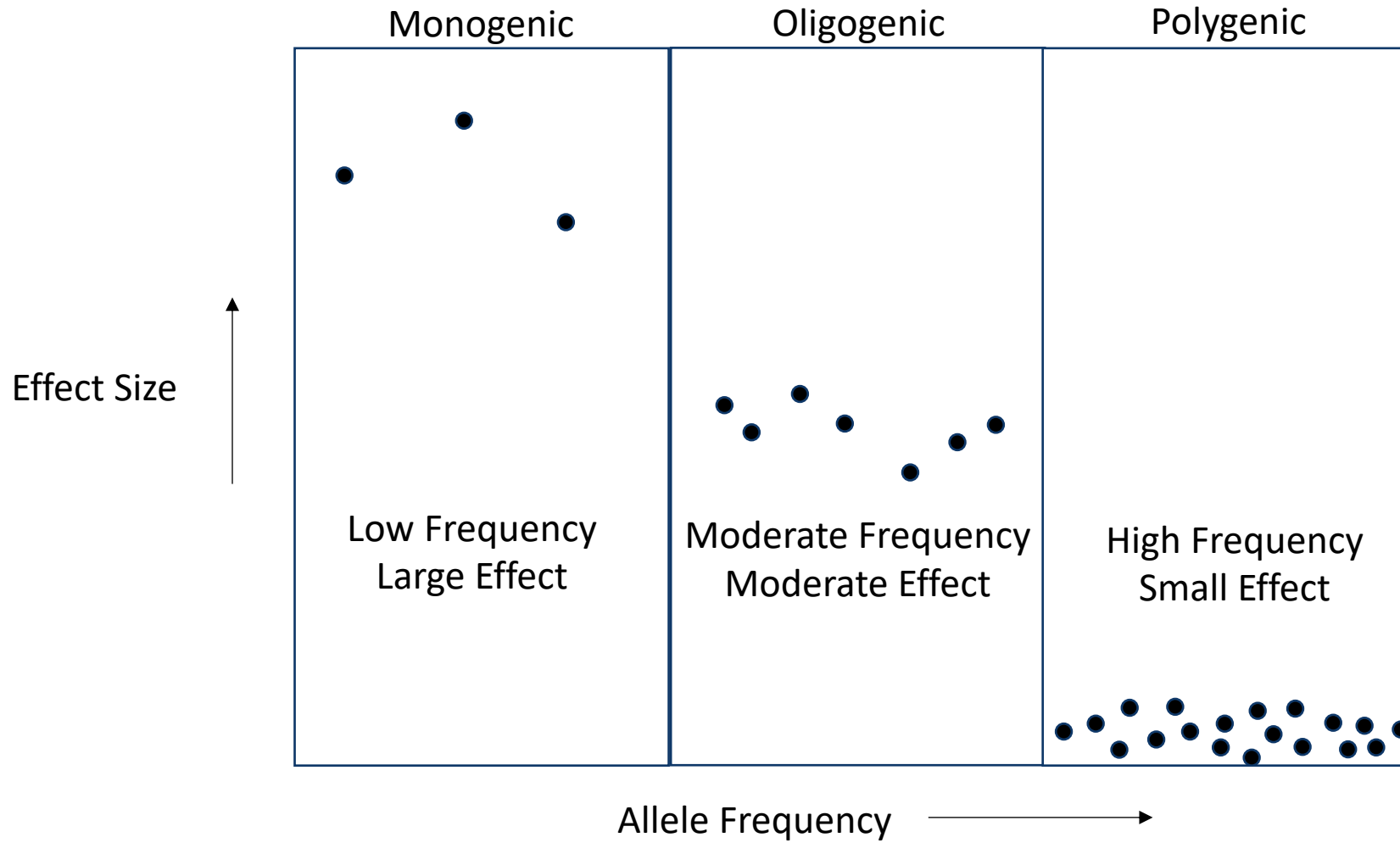
Polygenic Risk Score

- Sum of individual risks for specific SNVs weighted by effect size
  - PRS = $\beta_1X_1 + \beta_2X_2 + \beta_3X_3 + \ldots$
  - β_n = effect size of nth SNV; X_n = number of risk alleles at nth SNV
- Expressed in terms of standard deviations from mean phenotype
- Sensitive to variation in specific populations
- Identify people at extremes who need additional f/u or treatment
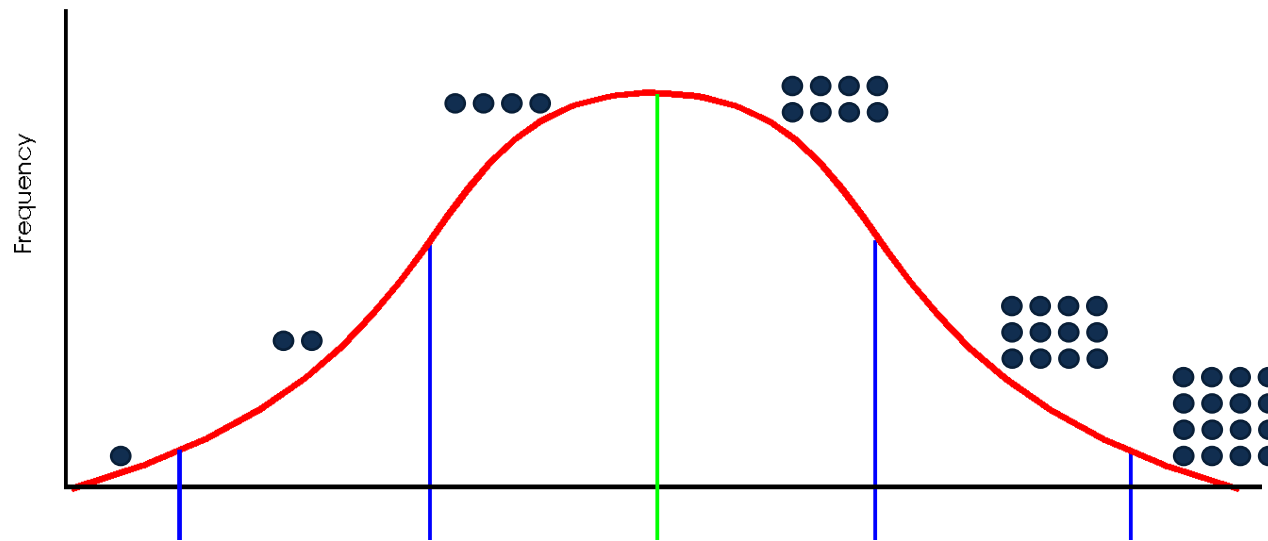
Race and Polygenic Risk Scores

Article | Open Access | Published: 25 July 2019

Analysis of polygenic risk score usage and performance in diverse human populations

L. Duncan, H. Shen, B. Gelaye, J. Meijsen, K. Ressler, M. Feldman, R. Peterson & B. Domingue

Nature Communications 10, Article number: 3328 (2019) | Download Citation
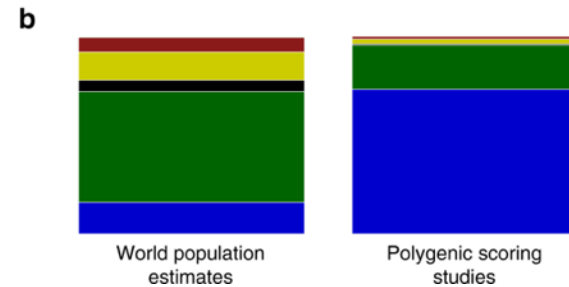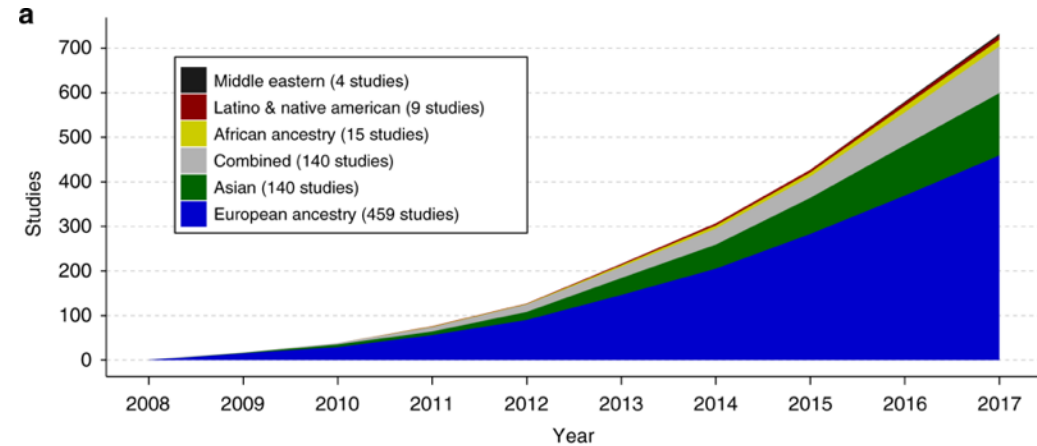
Integrating Genomics into Healthcare: A Global Responsibility

Zornitza Stark,[1,2,3] Lena Dolman,[4,5] Teri A. Manolio,[6] Brad Ozenberger,[7] Sue L. Hill,[8] Mark J. Caulfied,[9] Yves Levy,[10] David Glazer,[11] Julia Wilson,[12] Mark Lawler,[13] Tiffany Boughtwood,[1,2] Jeffrey Braithwaite,[1,14] Peter Goodhand,[4,5] Ewan Birney,[4,15] and Kathryn N. North[1,2,3,4,*]

United Kingdom
Genomics England 2012-
100,000 Genomes: rare disease, cancer
£350M (USD$485M)
Scottish Genomes £6M (USD$8M)
Welsh Genomics for Precision Medicine
£6.8M (USD$9M)
Northern Ireland Genomic Medicine
Centre £3.3M (USD$4.6M)

Switzerland
Swiss Personalized Health Network 2017-2020
Infrastructure
CHF68M (USD69M)

France
Genomic Medicine Plan 2016-2025
Rare disease, cancer, diabetes €670M
(USD$799M)

Estonia
Estonian Genome Project 2000 –
Infrastructure and population-based
cohort
2017: €5M for 100,000 individuals

Netherlands
RADICON-NL 2016-2025
Rare disease
Health Research Infrastructure

Finland
National Genome Strategy 2015-2020
Infrastructure
€50M ($USD 59M)

Denmark
Genome Denmark 2012-
DK 86M (USD$13.5M)
FarGen 2011- 2017
DK 10M (USD$1.6M)
Infrastructure, population-based
cohort, pathogen project

United States of America
National Human Genome Research
Institute 2007-
Infrastructure and clinical cohorts
USD$427M
All of Us 2016-2025
Population cohort
USD$500M (first two years)

Turkey
Turkish Genome Project 2017-2023
Infrastructure, clinical and population-
based cohorts

Brazil 2015-
Brazil Initiative on Precision Medicine
Infrastructure, disease and population
cohorts

Japan
Japan Genomic Medicine Program, 2015-
Infrastructure, clinical and population-based
cohorts, drug discovery
JPY10.2B (USD$90.05M)

China Precision Medicine Initiative
100,000,000 genomes
CNY60 billion (USD$9.2 billion)

Saudi Arabia
Saudi Human Genome Program, 2013-
Infrastructure, clinical cohorts and
population-based cohorts
SAR300M (USD$80M)

Qatar
Qatar Genome 2015-
Infrastructure, population cohort

Australia
Australian Genomics 2016-2021
Infrastructure, rare disease and cancer
AUD$125M (USD$95M)
Genomics Health Futures Mission 2018-2028
AUD$500M (USD$372M)

Cancer Genomics

Cancer is a genetic disorder

Normal cell → Increased growth → Abnormal function → Spread → Metastatic Cancer

Genetic Change 1, Genetic Change 2, Genetic Change 3, Genetic Change 4

Cancer Genomes

Tumor

Sequence

Targeted treatment

cancer-specific genetic changes

Pharmacogenetics

Pharmacogenetics: Drug Metabolism

- Absorption
  - GI
  - Tissue spaces

- Metabolism
  - Activation
  - Inactivation

- Excretion
  - Kidney
  - Liver

Phase I

Phase II

Warfarin Pharmacogenetics

CPIC
Clinical Pharmacogenetics
Implementation Consortium

CPIC level A or B:
Prescribing action recommended; alternative therapies or dosing are highly likely to be effective and safe

CPIC level C:
No prescribing change based on genetics; alternatives are unclear or evidence is weak but testing is common or gene is CPIC level A or B for other drugs

CPIC level D:
PharmGKB annotation only; no prescribing action recommended; alternatives unclear or evidence is weak; testing is rare

Genomics and Public Health

Preconceptional Screening
Newborn Screening
Population Screening

Carrier Screening

ACMG Guideline

Tier 4[¥]	<1/200 carrier frequency (includes Tier 3) genes/condition will vary by lab
Tier 3[§]	≥ 1/200 carrier frequency (includes Tier 2) includes X-linked conditions
Tier 2[±]	≥1/100 carrier frequency (includes Tier 1)
Tier 1*	CF + SMA + Risk Based Screening

ACMG recommends:
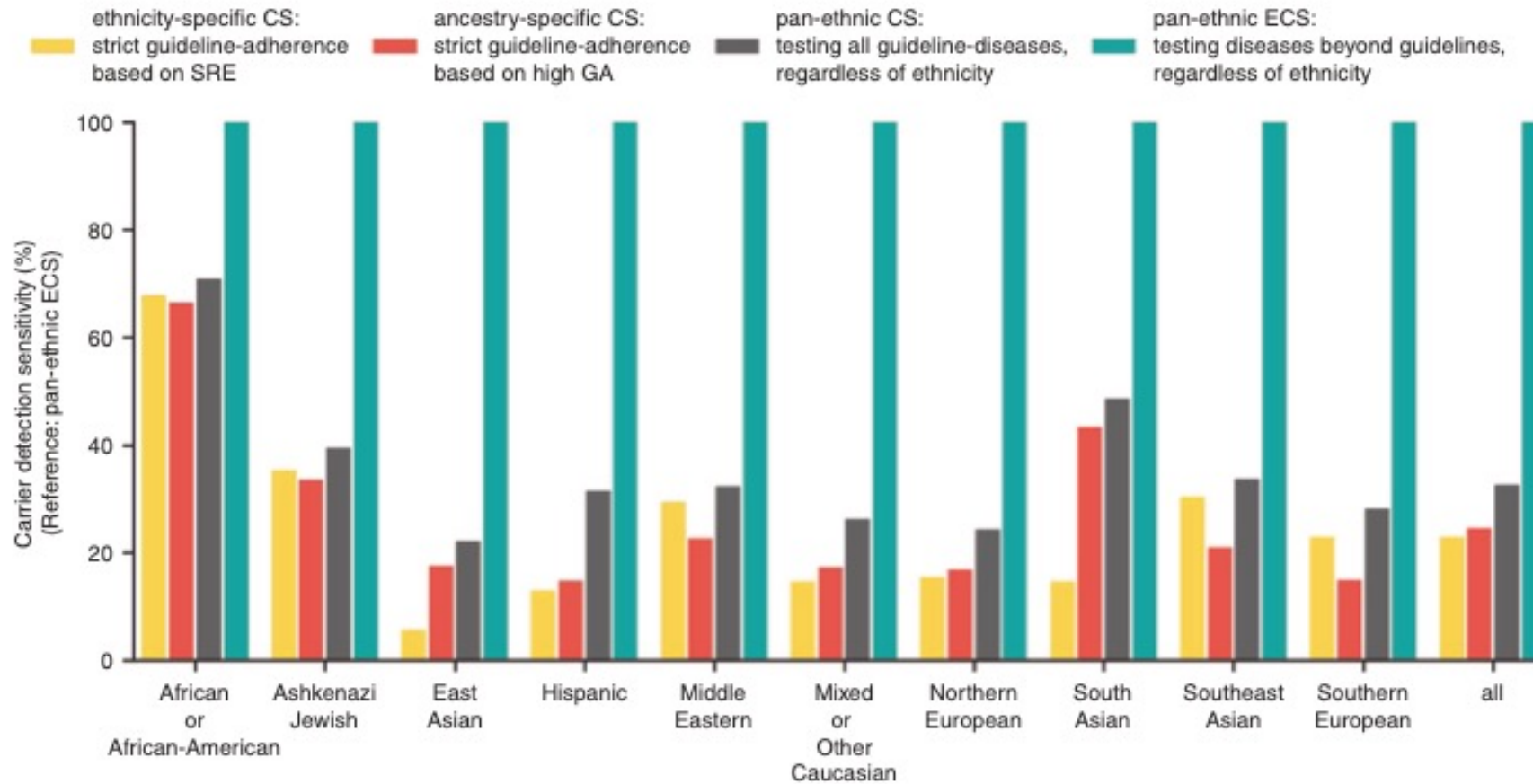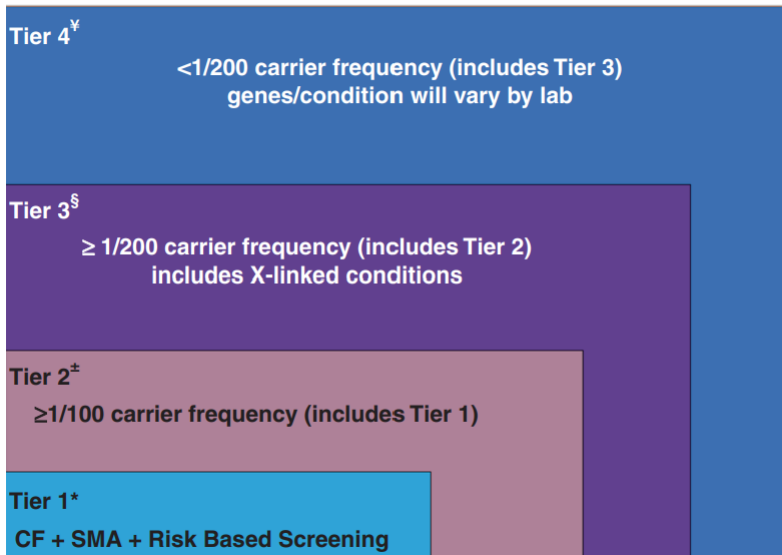
- All pregnant patients and those planning a pregnancy should be offered Tier 3 carrier screening.
- Tier 4 screening should be considered:
    - When a pregnancy stems from a known or possible consanguineous relationship (second cousins or closer);
    - When a family or personal medical history warrants.

ACMG does not recommend:

- Offering Tier 1 and/or Tier 2 screening, because these do not provide equitable evaluation of all racial/ethnic groups.
- Routine offering of Tier 4 panels.

Genetics in Medicine (2021) 23:1793–1806; https://doi.org/10.1038/s41436-021-01203-z

Genomic Newborn Screening

Challenges

- Not all conditions currently subject to screening can be detected
- Variants of unknown significance
- Adult-onset variants
- Parental consent

Population Screening

Challenges

- Non-penetrance
- Variants of unknown significance
- Actionability
- Access to care

Prioritizing Genomic Applications for Action by Level of Evidence: A Horizon-Scanning Method

WD Dotson[1], MP Douglas[1,2], K Kolor[1], AC Stewart[1,2], MS Bowen[1], M Gwinn[1,2], A Wulf[1,3], HM Anders[1,2], CQ Chang[4], M Clyne[4,5], TK Lam[4], SD Schully[4], M Marrone[6], WG Feero[7] and MJ Khoury[1,4]

- Tier 1/Green genomic applications have a base of synthesized evidence that supports implementation in practice.
- Tier 2/Yellow genomic applications have synthesized evidence that is insufficient to support their implementation in routine practice. Nevertheless, the evidence may be useful for informing selective use strategies (such as in clinical trials) through individual clinical, or public health policy, decision making.
- Tier 3/Red applications either (i) have synthesized evidence that supports recommendations against or discourages use, or (ii) no relevant synthesized evidence is available.
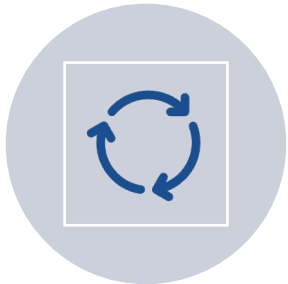
CDC

OAK RIDGE INSTITUTE FOR SCIENCE AND EDUCATION

Into the Future

Technological Innovation

Clinical Utility

Implementation

Education

We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run.

Amara's Law