Scientific Overview: Introduction to Genetics and Genomics

Genetics and the Law, PHG 523/LAW H 520

Sarah Nelson, MPH, PhD

January 6th, 2021

@blueyedgenes myopenreadingframe.com sarahcn@uw.edu

Learning objectives

Understand different types of genetic variation

Describe the relationship between genotype and phenotype

Recognize major milestones in the history of genetic research

Identify applications of genomics relevant to legal issues

Overall goal: Gain basic scientific understanding of genetics in order to apply the law and legal concepts in cases that involve genetics

Outline

I. Basic Biology

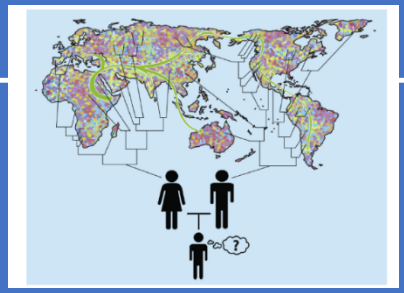
II. Genetic Research

III. Applications

“Many orientedness” of DNA

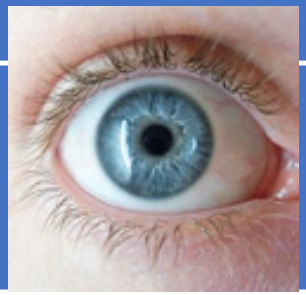
Family relationships

Ancestry/population history



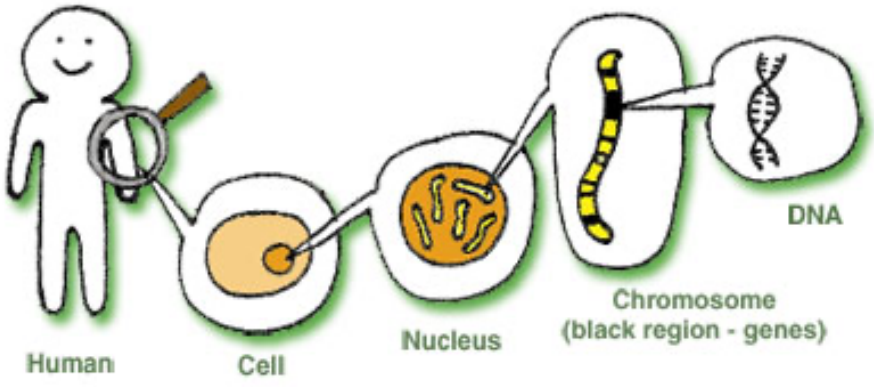
Disease risk/susceptibility

Non-medical traits



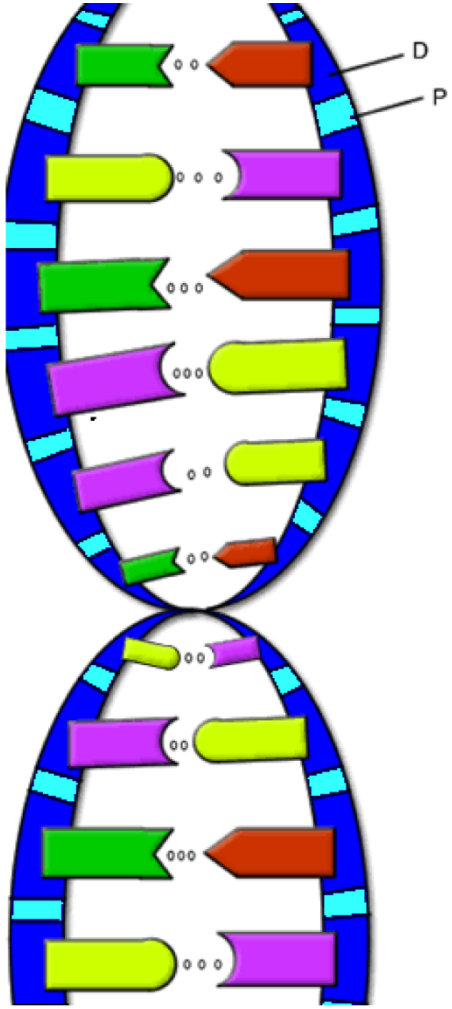
Basic Biology: overview

- DNA
  - Molecule that carries genetic information
- Genes
  - Protein-coding units of DNA
  - ~20K genes in human genome
- Chromosomes
  - Organized packages of DNA
  - Located in nucleus of the cell
- Human genome
  - Entire set of genetic instructions found in a cell



Credits: American Society of Human Genetics, https://www.ashg.org/education/everyone_1.shtml

What Does DNA Look Like?

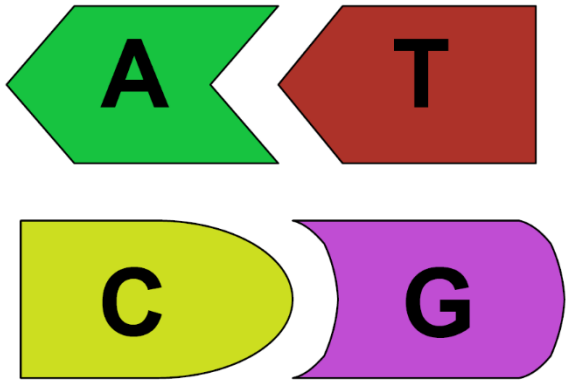


Thymine
Adenine
Guanine
Cytosine

D = Deoxyribose (sugar)

P = Phosphate

Hydrogen Bond



A, T, C, G: nucleotides, also called “base pairs”

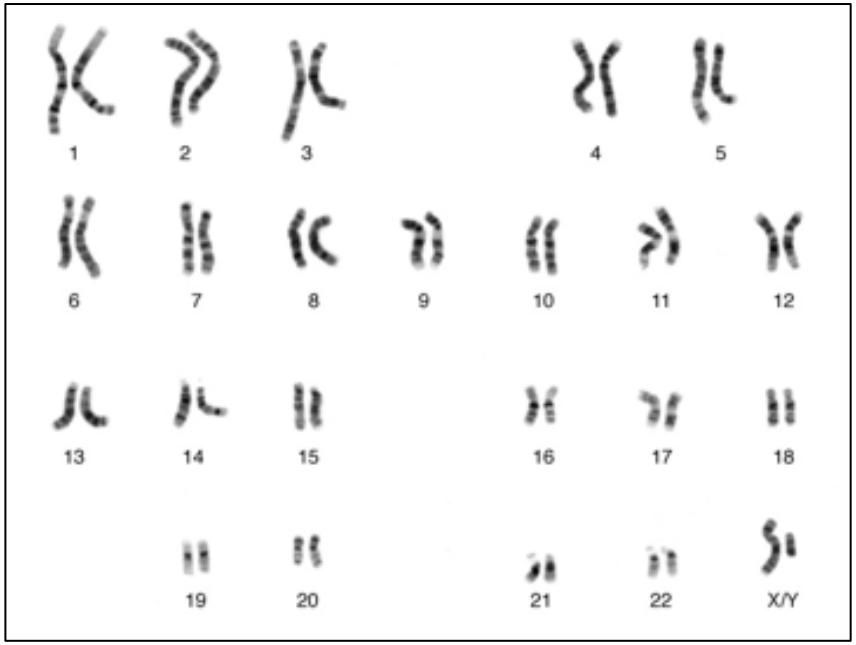
DNA is organized into chromosomes

Chromosomes

- Humans: chromosomes 1-22 (autosomes) and sex chromosomes (X, Y)
- One set of chromosomes from each parent

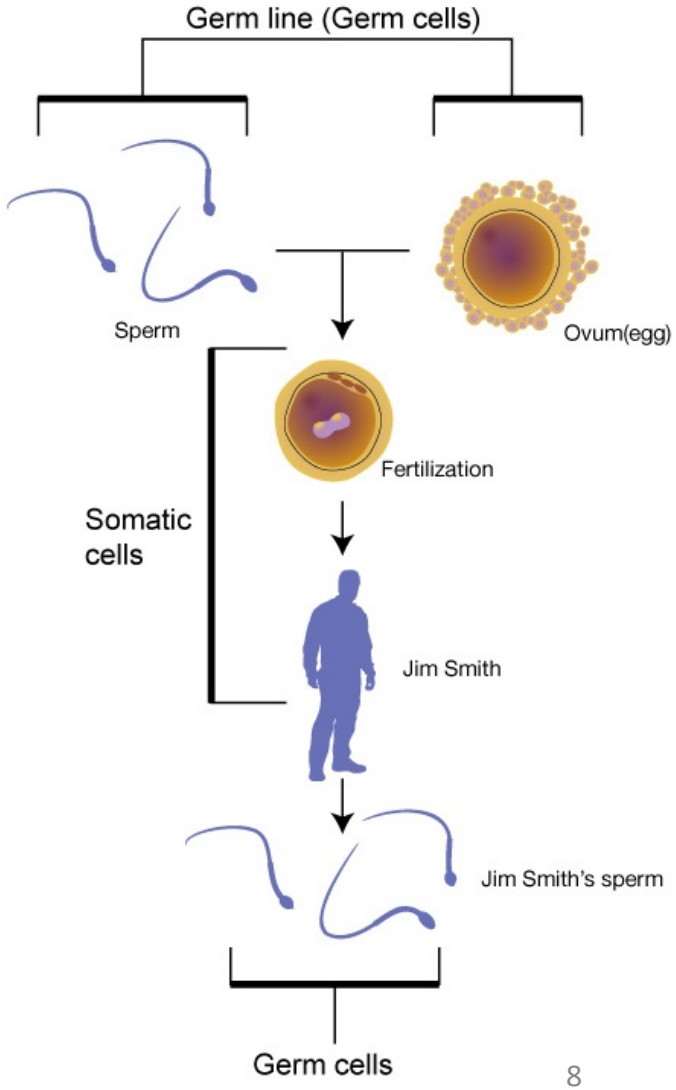
Karyotype

- An individual's collection of chromosomes
- Laboratory visualization technique

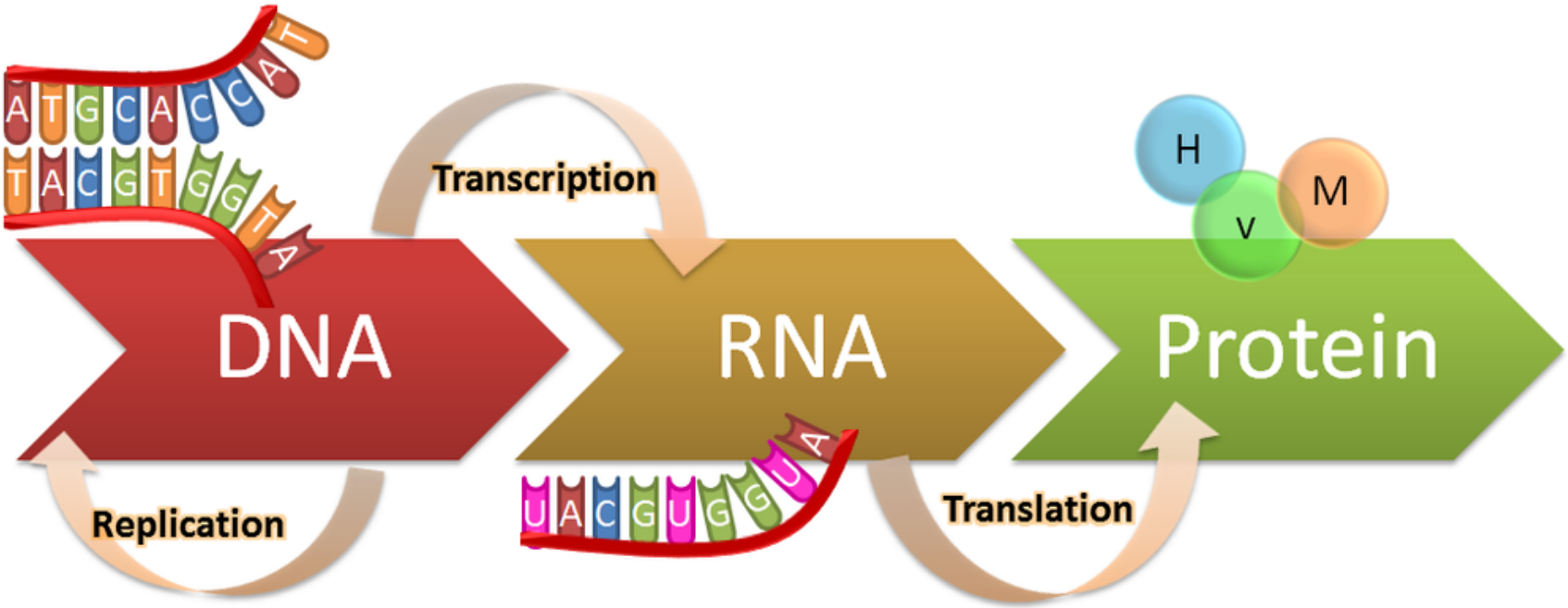


Somatic vs. Germline genome

- Germ cells
  - Reproductive cells (i.e., sperm and egg)
  - Genetic content passed onto offspring
- Somatic cells
  - Non-germline
    - Your other body tissues and parts (heart, blood, liver, etc.)
    - “soma” = body in Greek
  - Genetic content NOT passed onto offspring
    - Can be “passed on” to daughter cells during cell division, in the same person)



“Central Dogma” of Biology

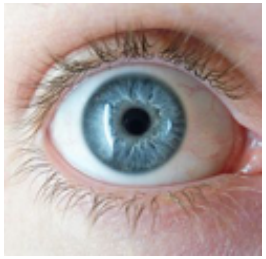


Images: https://genius.com/Biology-genius-the-central-dogma-annotated

Genotype and Phenotype

- Genotype: genetic information

- Phenotype: outward expression of a trait



- Genotype matters more for some phenotypes than others…

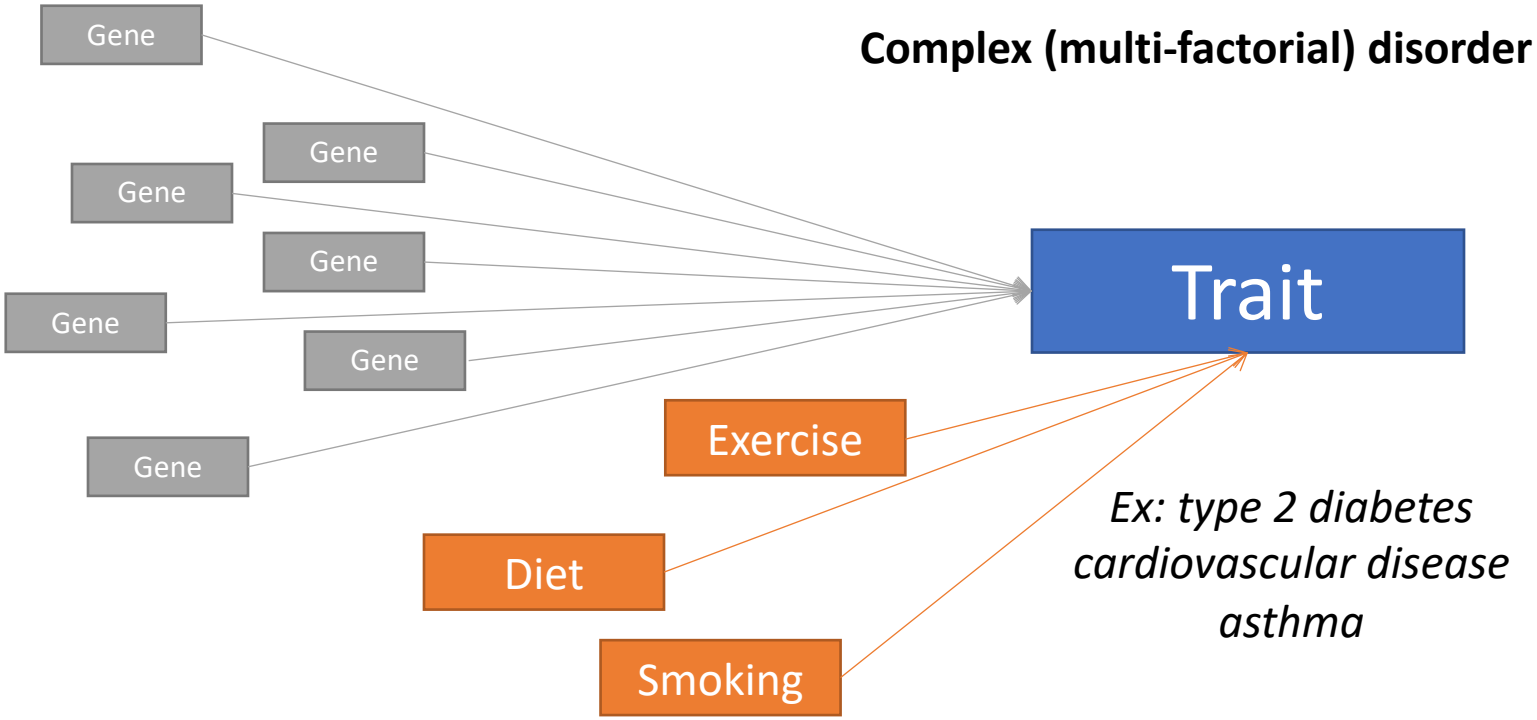
Genotype and Phenotype

Mendelian (single gene) disorder

HTT gene

Huntington's Disease

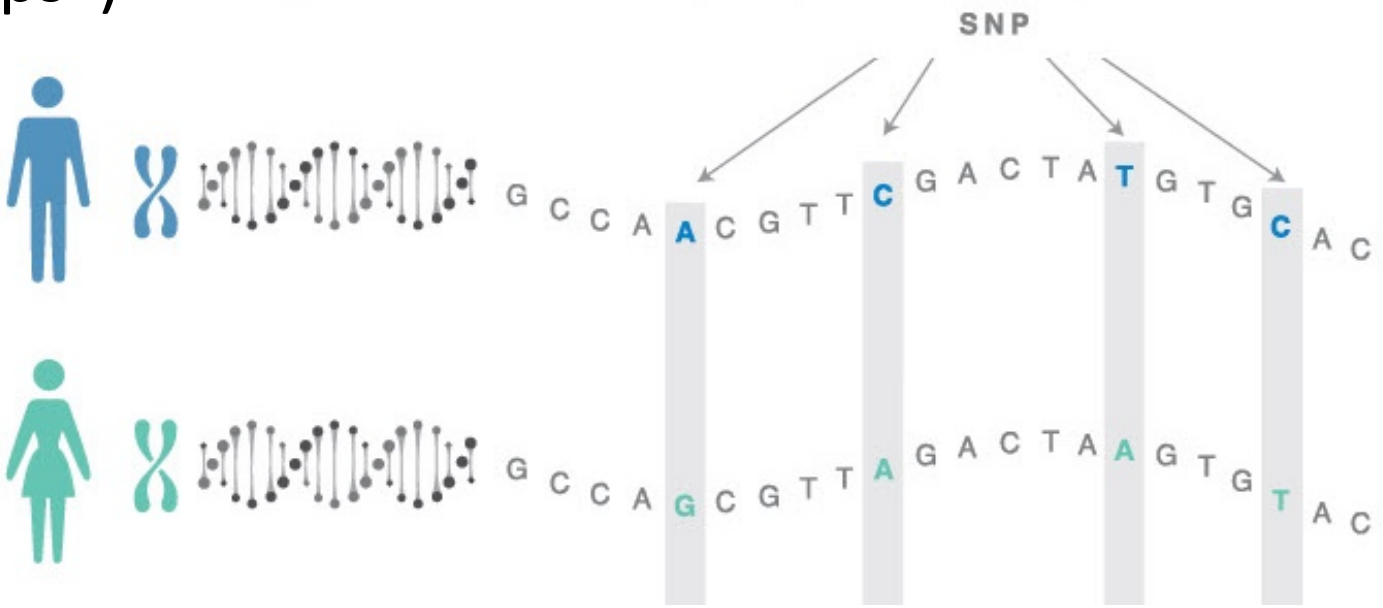
Complex (multi-factorial) disorder



Ex: type 2 diabetes
cardiovascular disease
asthma

Types of genetic variation

1) Single Nucleotide Polymorphisms (SNPs, or "snips")



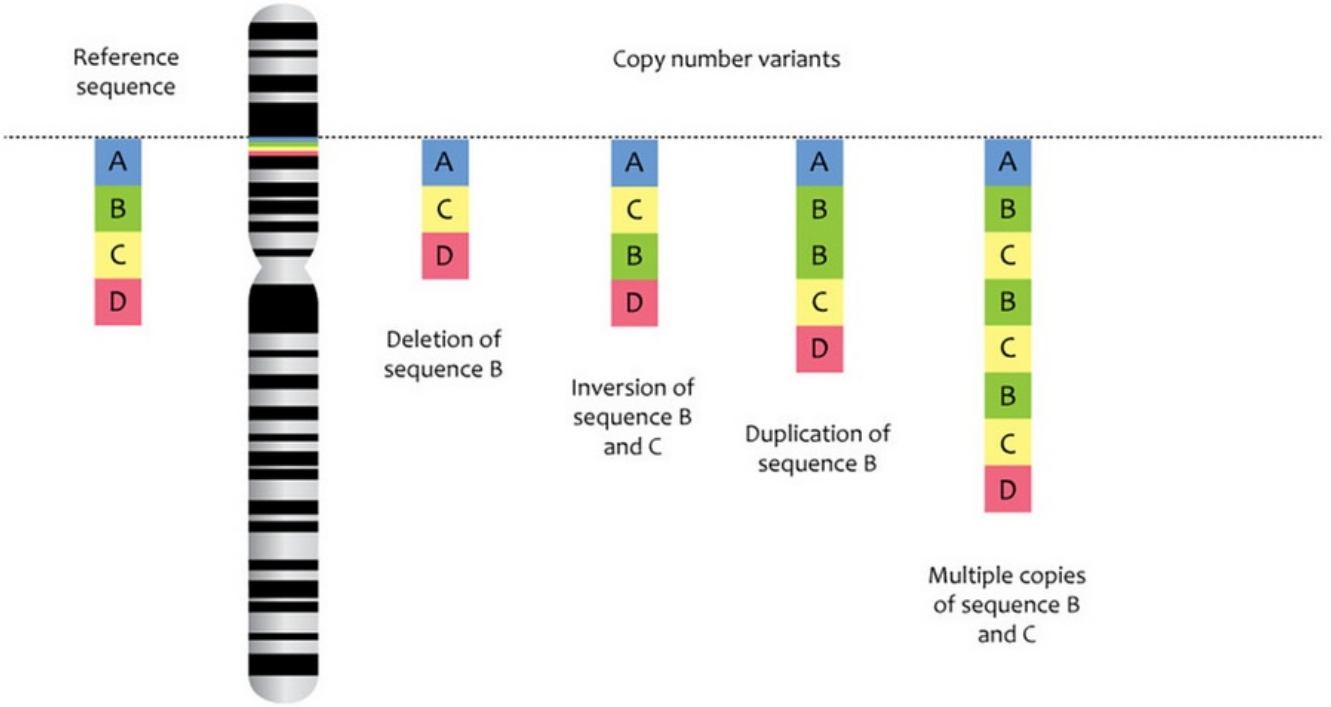
Legal application: investigative genetic genealogy

https://www.natera.com/node/1856

Types of variation

2) Copy number variation (CNV)

- also referred to as structural variation



Short Tandem Repeats (STRs)

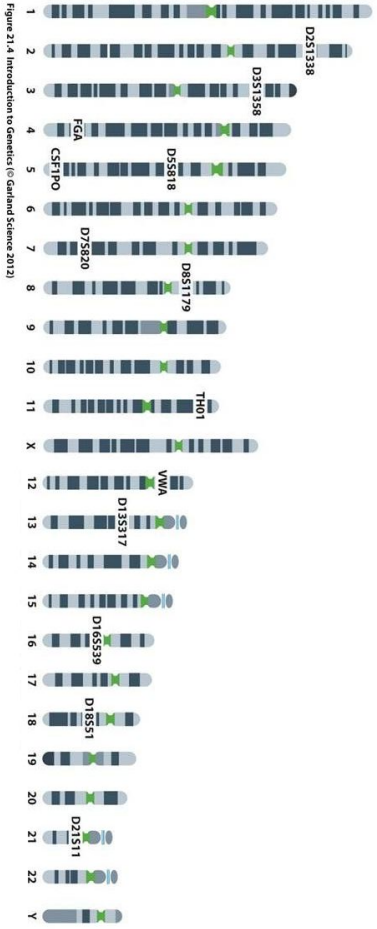
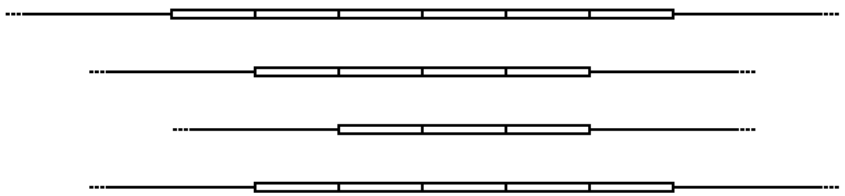
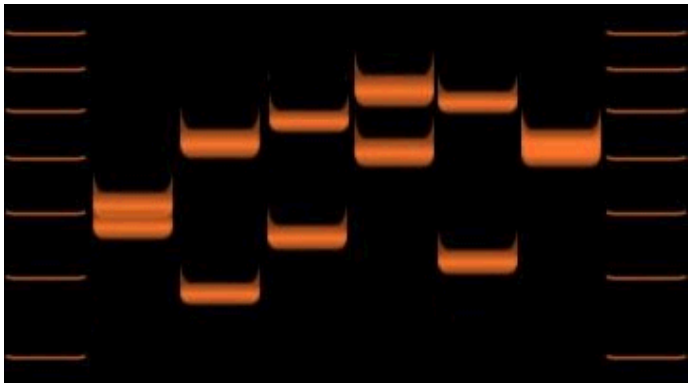


Figure 21.4 Introduction to Genetics (© Garland Science 2012)

- Repeating sequences of 2-6 base pairs of DNA
- Multi-allelic



- Specimen DNA amplified, and samples are separated by size

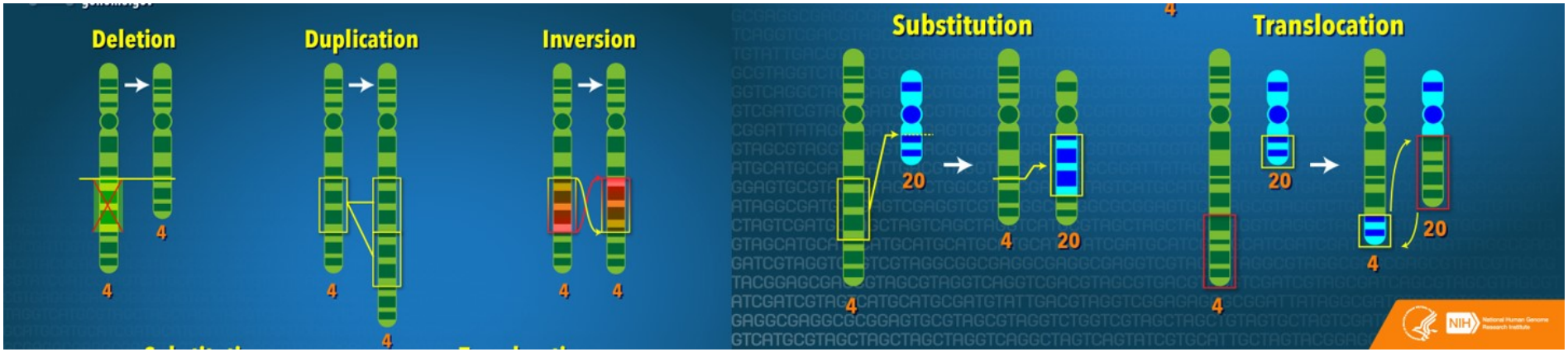


Legal application: Combined DNA Index System (CODIS)

Type of variation

3) Chromosomal anomalies

- Large pieces of chromosomes deleted, duplicated, or rearranged



- Also whole chromosome deletions or duplications
  - e.g., Trisomy 21 – 3 copies of chrom21 – Down syndrome
  - Sex chromosome aneuploidies: XXX, XXY, XYY, X0

Types of Genetic Variation

GERMLINE

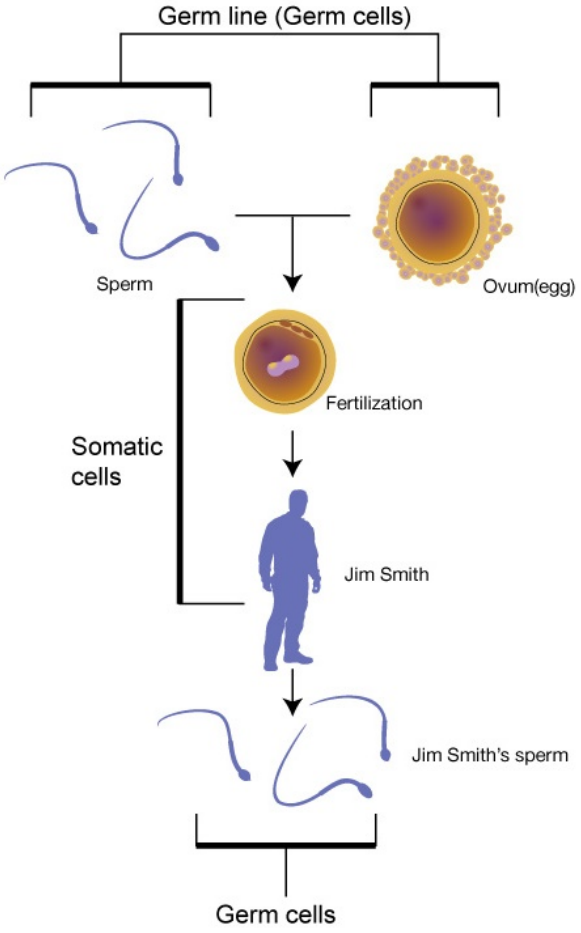
- Inherited from parents
- New (de novo) mutations in sperm or egg cells (germline)

<u>CAN be passed on to offspring</u>

SOMATIC

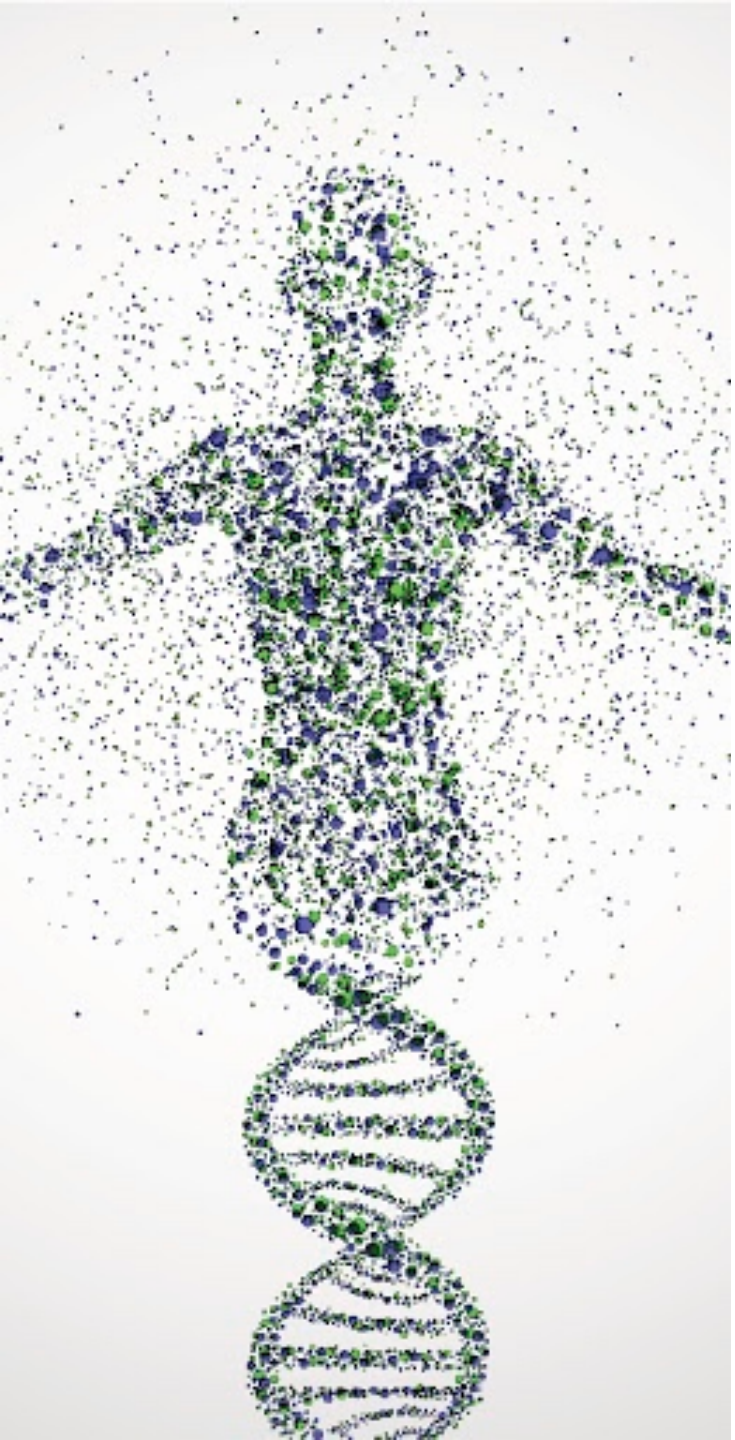
- Acquired variation in non-germ cells, e.g. from….
- Errors in DNA replication
- DNA damage from environment
- Intentional modification (gene therapy)

<u>CANNOT be passed on to offspring</u>



Variation: additional terms

- **Locus**: specific physical location of a gene or other DNA sequence on a chromosome (plural="loci")
  - Could be a single SNP or an entire gene
- **Allele**: one of two or more versions of a locus
  - E.g.: a SNP with two alleles, A and G
  - Alleles can also refer to longer stretches of a DNA
    - E.g., a neighboring set of SNPs, called a "haplotype" or a whole gene
- **Homozygous:** both alleles are the same (e.g., AA)
- **Heterozygous:** two alleles are different (e.g., AG)

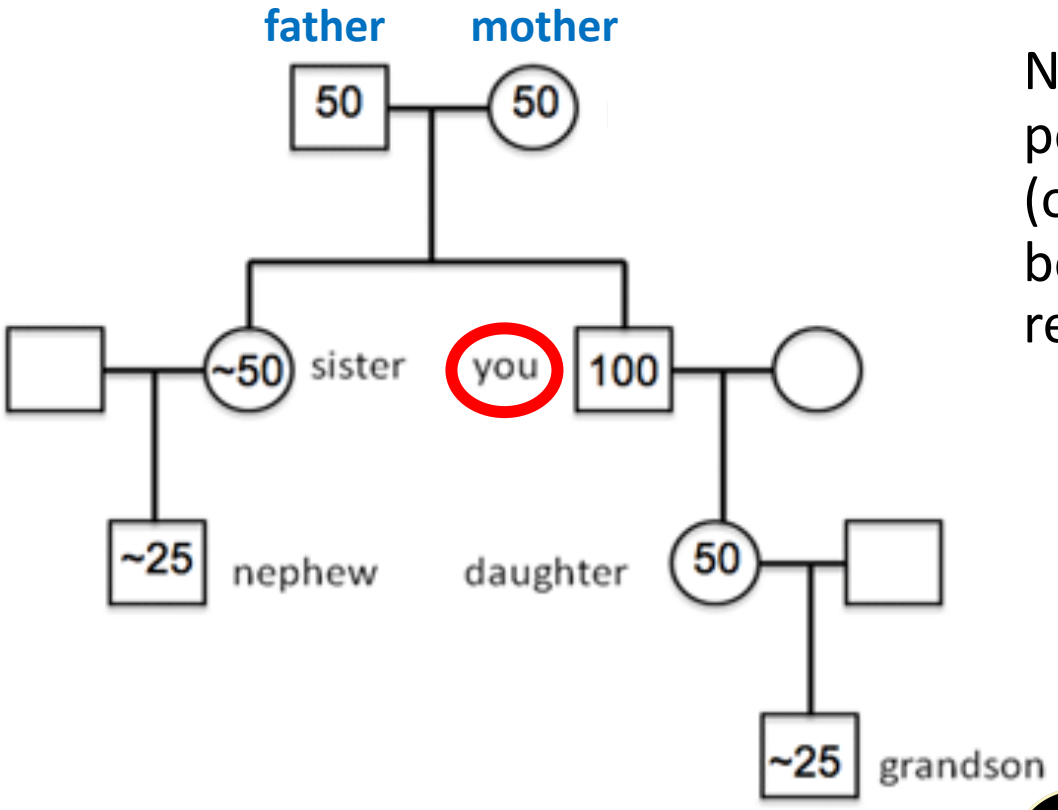


Variation on what?

- Being a DNA “variant” typically means the DNA sequence varies/differs from the **human genome reference sequence**
  - Generated during Human Genome Project (more later)
  - A consensus sequence based on ~20 anonymous donors
  - Not comprehensive representative of human genetic variation worldwide
  - But practically useful for measuring and describing DNA variation
    - E.g., at a given position, does a chromosome contain the “reference” vs an “alternative” allele

Relatedness

Pedigree: common way of depicting family relationship



Numbers given are the percent of genetic material (on average) shared between these different relative pair types

Legal applications: Family law; familial searching

Image adapted from http://genetics.thetech.org/ask-a-geneticist/grandson-nephew-relatedness

II. GENETIC RESEARCH

Learning objectives

Understand different types of genetic variation

Describe the relationship between genotype and phenotype

Recognize major milestones in the history of genetic research

Identify applications of genomics relevant to legal issues

Overall goal: Gain basic scientific understanding of genetics in order to apply the law and legal concepts in cases that involve genetics

II. Genetic Research

- Overall goal

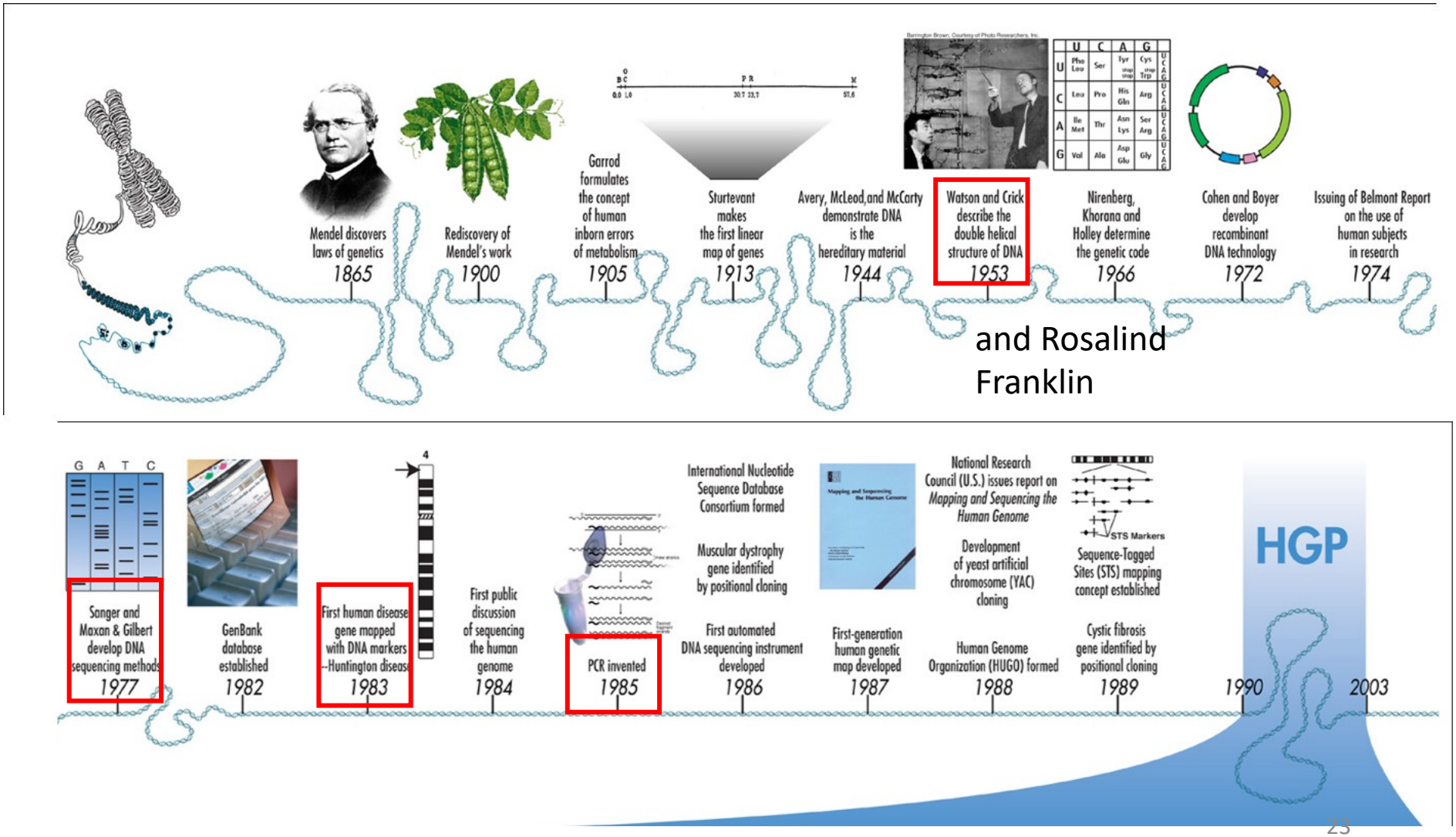
To **understand** human biology and disease in order to **improve** individual and population health

  - Basic → clinical → translational ("bench to bedside")

- Theme

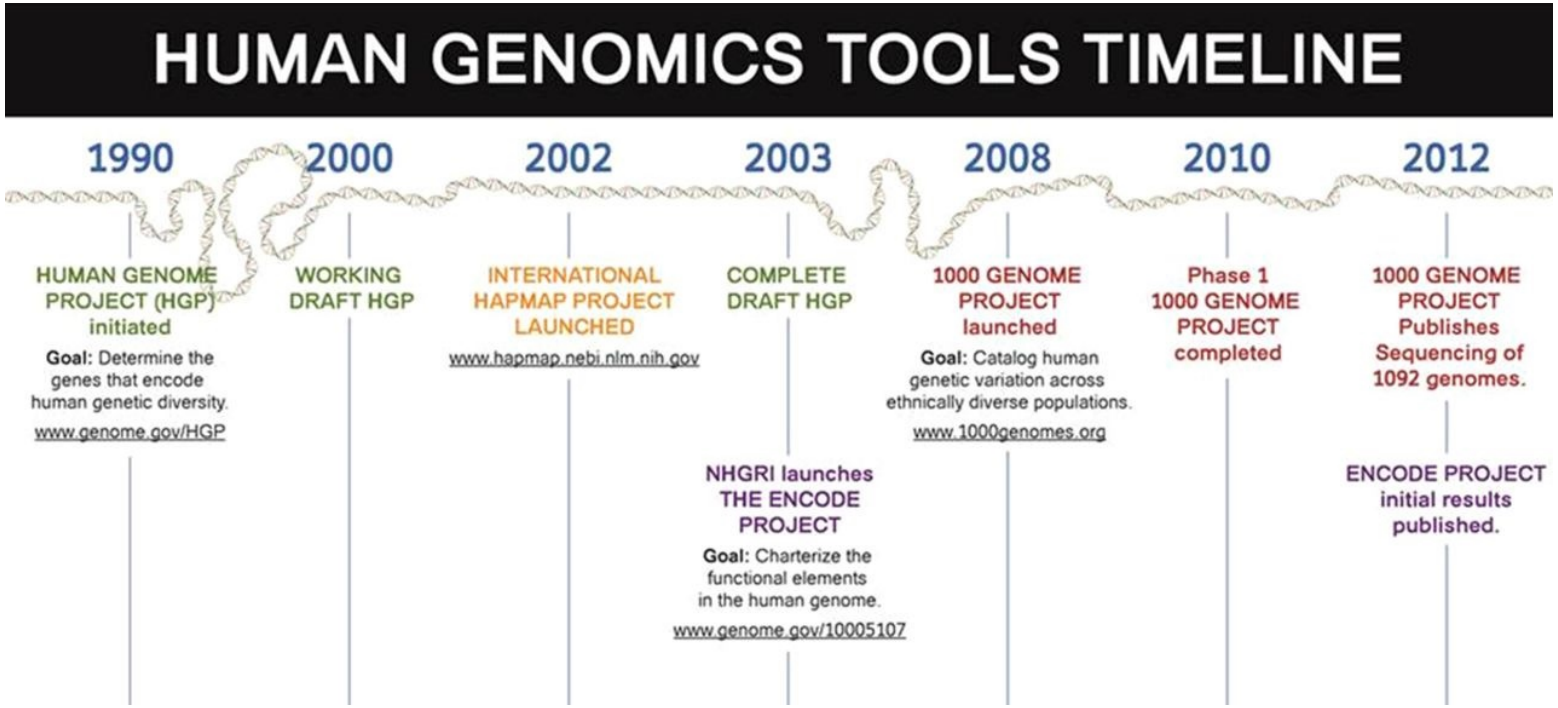
Advances in laboratory technologies, bioinformatics, and statistical techniques have allowed new research directions

Genetic research pre 21st century



Research highlights

Project	Technical side note
Human Genome Project	Sanger sequencing
HapMap Project	Genome-wide association studies (GWAS) & genotyping microarrays (“SNP chips”)
1000 Genomes Project	Next-generation sequencing
Precision Medicine Initiative	Combine genomics with other “omics”



Fox et al. Circulation 2015

Human Genome Project

- International, collaborative research program to “map” (determine the sequence of) the human genome
  - In US, led by National Institutes of Health (NIH) and Department of Energy (DOE)
- Launched in 1990, completed in April 2003
  - National Human Genome Research Institute (NHGRI) created in 1997
- Total cost: $2.7 billion in FY 1991 dollars
  - Cost to US taxpayers
- 5% of NHGRI budget dedicated to examining ethical, legal, and social implications (ELSI) of genetics

HGP, cont.

- "human genome reference sequence"
  - Composite sequence of a few (~20) anonymous individuals
- Sanger sequencing technology
  - Slow
  - Expensive

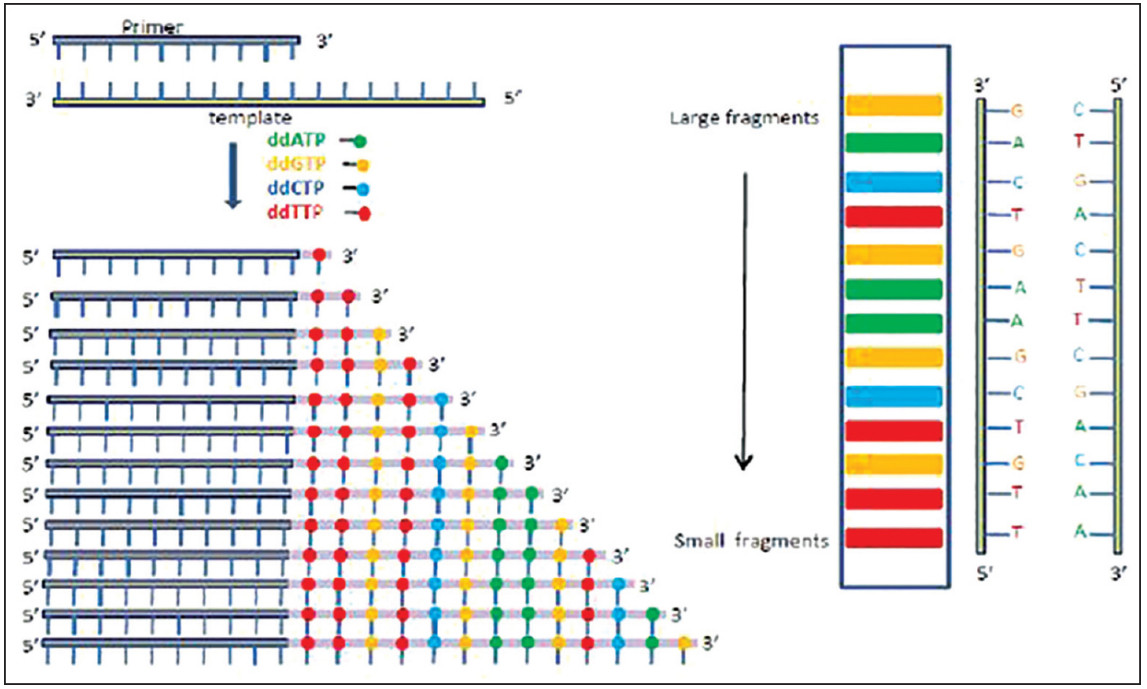
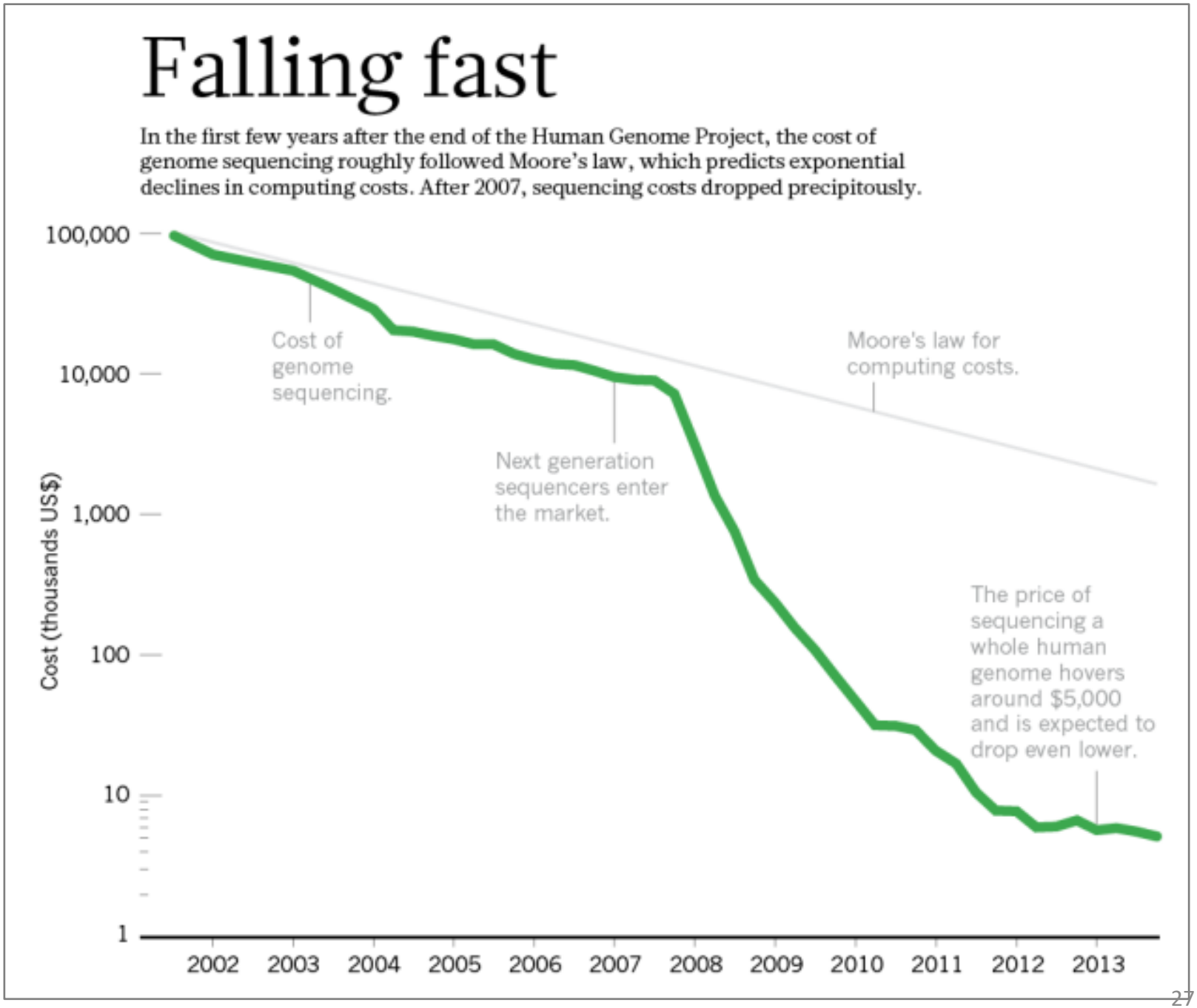


Image: Aliyu. ANM 2014



Hayden Nature News 19 Mar 2014

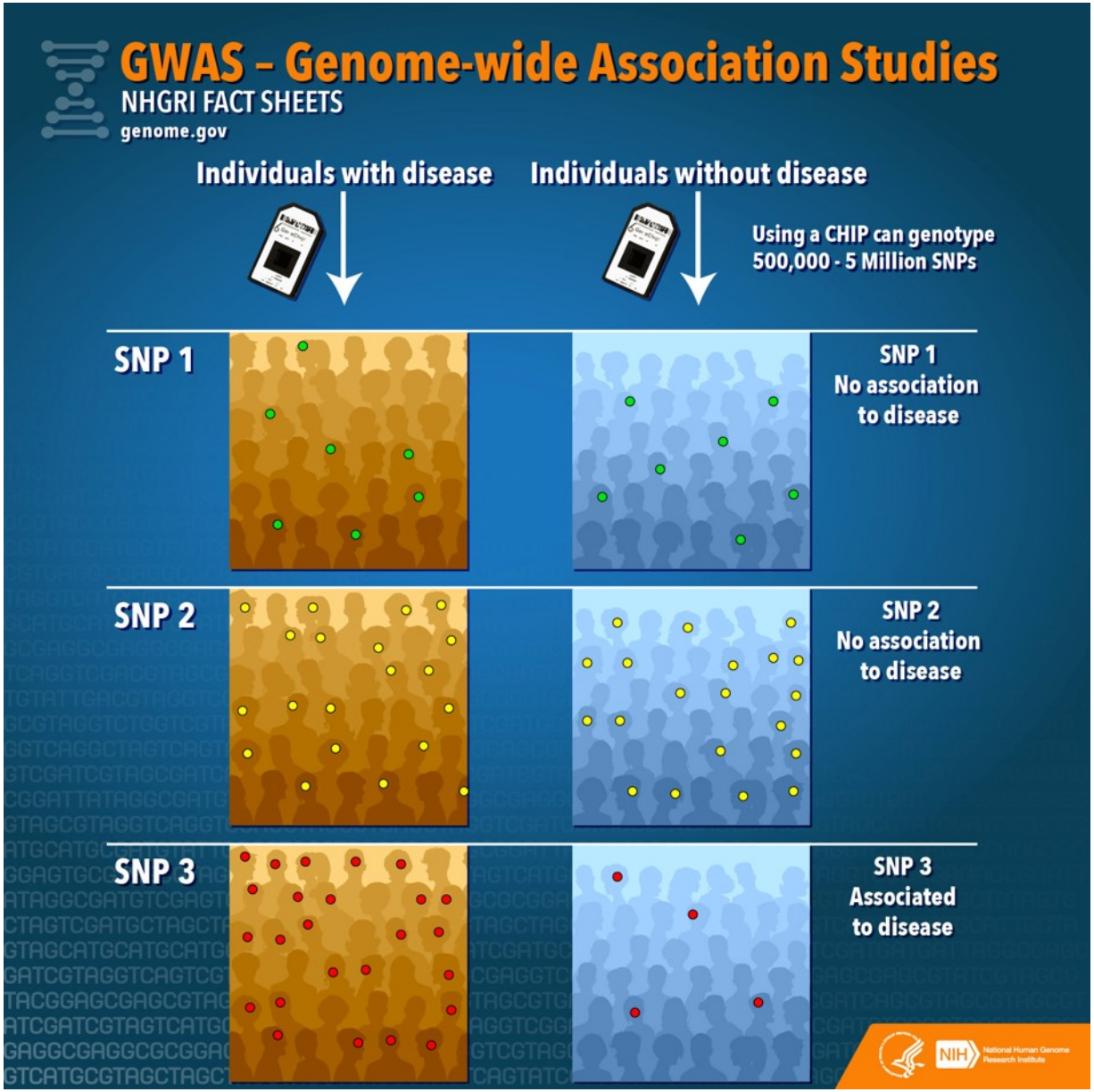
Post-HGP Research Initiatives

- Multi-national project 2002-2009
- Goal of identifying "tag" SNPs that are informative about genotype for other polymorphic loci
  - i.e. catalog **common variants**
- Initially 4 populations but expanded to 11 in final phase

www.hapmap.ncbi.nlm.nih.gov

Technical development:

- SNP microarray or "chip"
  - High throughput
  - Genotype hundreds of thousands to millions of SNPs in one experiment
  - Initially content focused on HapMap "tag" SNPs



HapMap Project helps bring us SNP chips which in turn usher in the age of GWAS (2007 – present, though waning)

NHGRI-EBI GWAS Catalog

This diagram shows all SNP-trait associations with p-value ≤ 5.0×10^{-8}, published in the GWAS Catalog.

http://www.ebi.ac.uk/gwas/diagram

- Digestive system disease
- Cardiovascular disease
- Metabolic disease
- Immune system disease
- Nervous system disease
- Liver enzyme measurement
- Lipid or lipoprotein measurement
- Inflammatory marker measurement
- Hematological measurement
- Body measurement
- Cardiovascular measurement
- Other measurement
- Response to drug
- Biological process
- Cancer
- Other disease
- Other trait

Post-HGP Research Initiatives, cont.

1000 Genomes
A Deep Catalog of Human Genetic Variation

- Multi-national project, 2008-2012
- Sequence the genomes of 2,504 individuals
- Goal of identifying >95% of genetic variation
  - Structural variation
  - Rare variation in genes
  - Population allele frequencies
  - Haplotypes and linkage disequilibrium patterns
- Samples from 26 populations worldwide

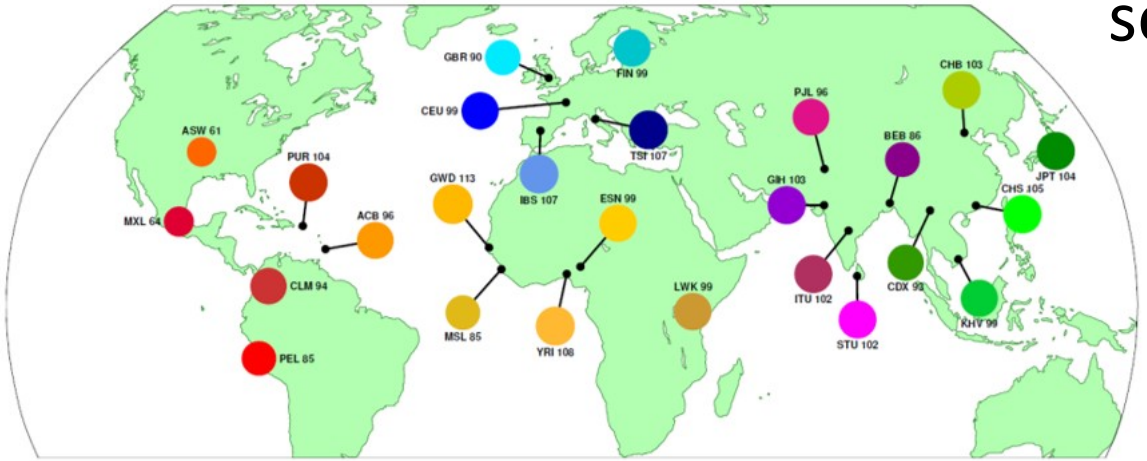
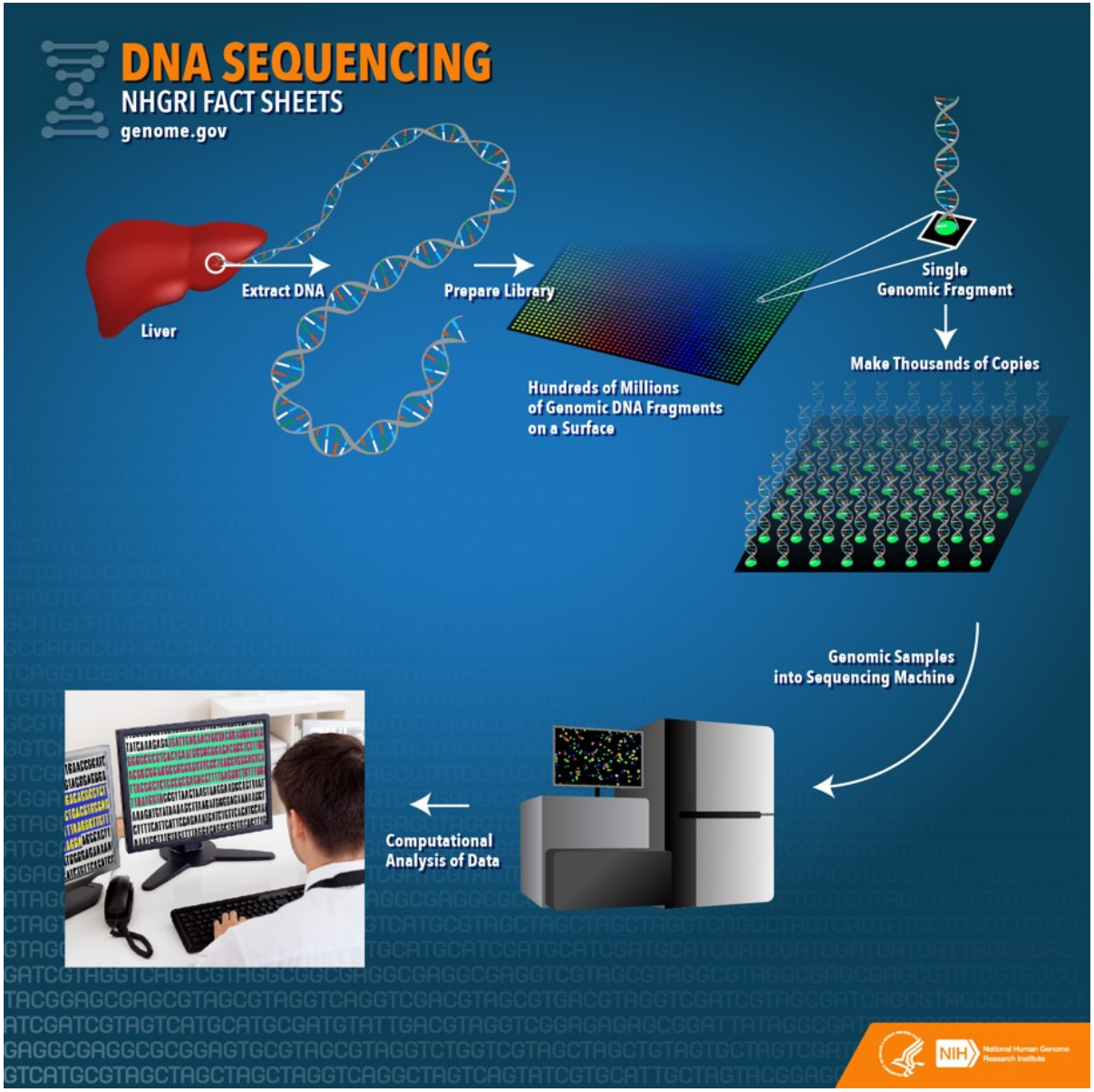


Image: G. Abecasis slides, ASHG 2014.
bubble size indicates sample size

Technical development:

- Next-generation sequencing (NGS)
- Much quicker and cheaper than Sanger and other earlier sequencing methods

Next generation DNA sequencing

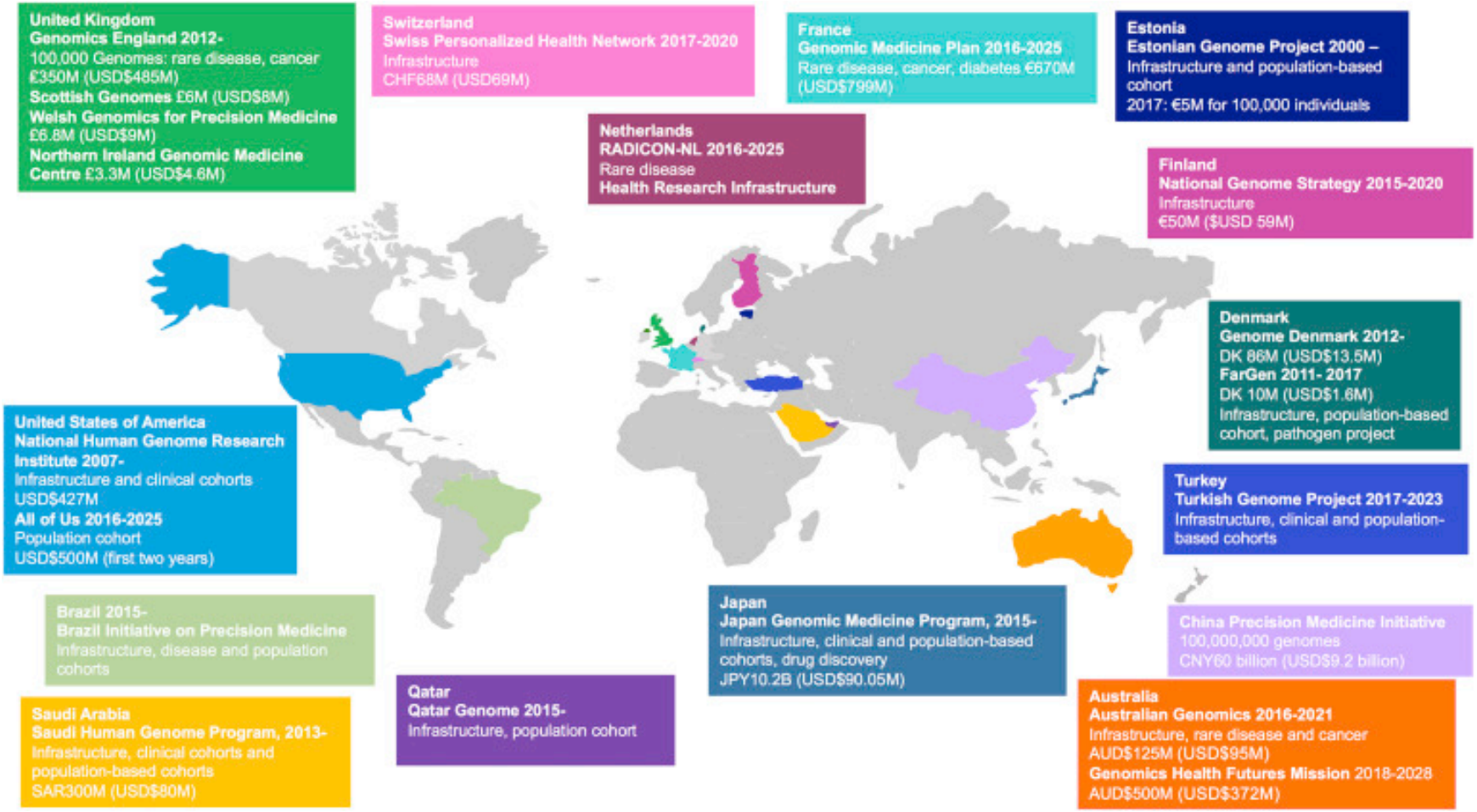


Precision Medicine Initiative

- Launched by President Obama during 2015 State of the Union
- Long-term focus: “bringing precision medicine to all areas of health and healthcare on a large scale”
- Now recruiting 1M person **research cohort**, “All of Us”

https://ghr.nlm.nih.gov/primer/precisionmedicine/initiative; https://obamawhitehouse.archives.gov/node/333101

Research initiatives for translational genomics: global context



Global Alliance for Genomics & Health (GA4GH), https://www.sciencedirect.com/science/article/pii/S0002929718304221

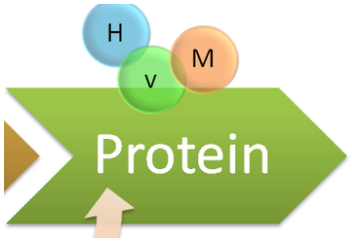
Large-scale human genetic research: Ethical and legal questions

- Can participants give **broad consent** or should they have input on **specific proposed uses** of their biobanked samples and associated clinical info?
- Do researchers have ethical or legal responsibility to **return results** (clinically relevant or otherwise) to participants?
  - What about when participant is deceased? Or notifying family members who may also be implicated?
- Can participants request that their sample and/or data be destroyed?

Mehlman, Rothstein, and Suter, *Genetics: Ethics, Law and Policy* 5th ed. Ch 4

Beyond the genome: "-Omics"

- Other areas of research interest
  - Transcriptomics (RNA)
  - Proteomics (protein)
  - Metabolomics
    - Study of small molecules ("metabolites")
  - Epigenomics
    - Study of chemical marks that regulate gene expression
  - Microbiome
    - Study of genomes of the many microorganisms that live in/on our bodies (e.g., gut, skin)



III. APPLICATIONS

Medicine and Public Health

Consumer Genomics

Forensics

Biotechnology (CRISPR)

Learning objectives

Understand different types of genetic variation

Describe the relationship between genotype and phenotype

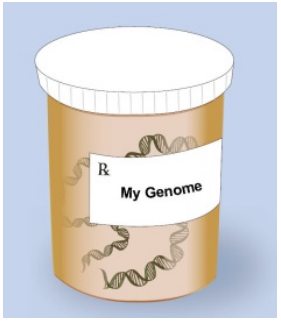
Recognize major milestones in the history of genetic research

Identify applications of genomics relevant to legal issues

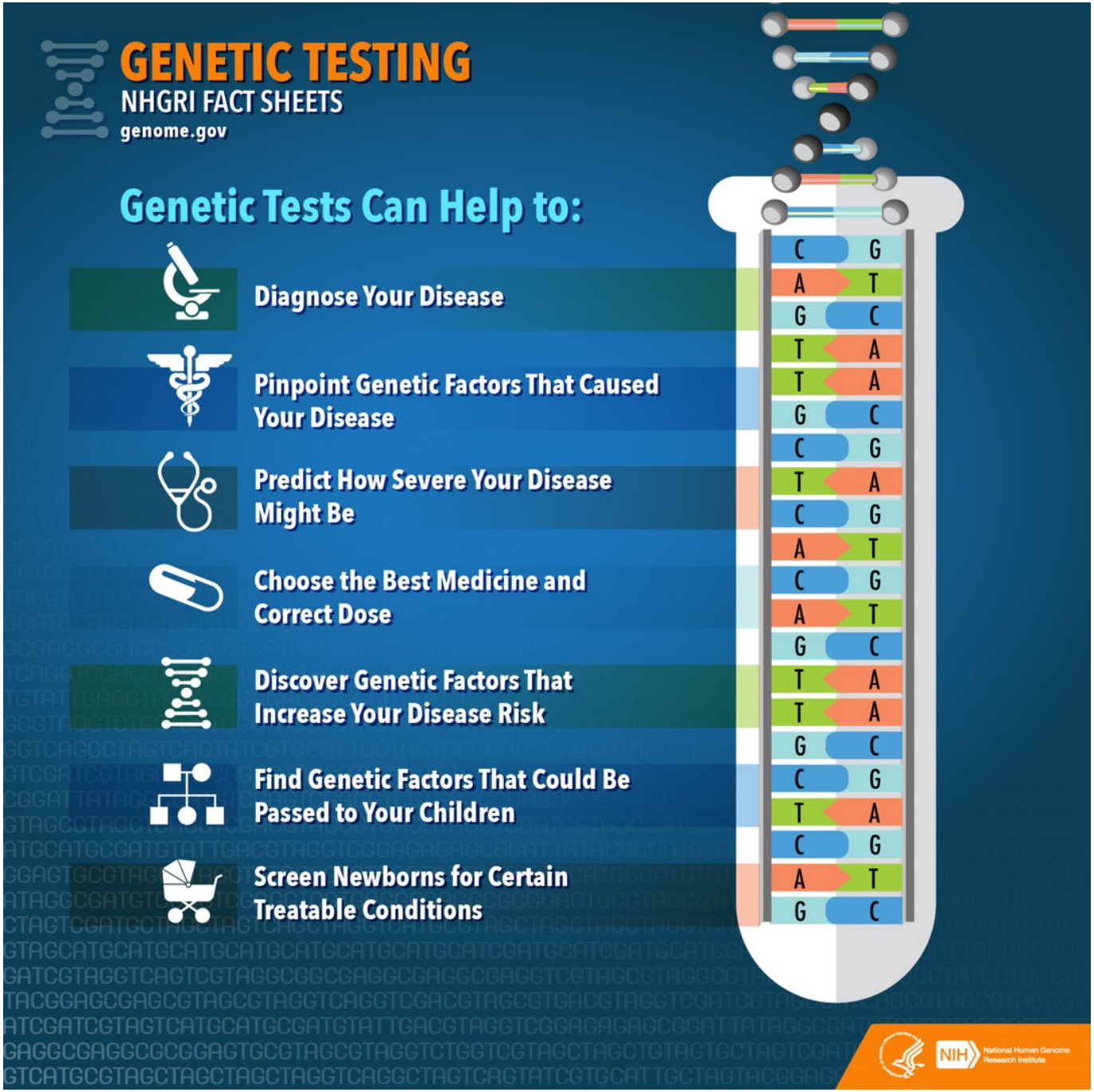
Overall goal: Gain basic scientific understanding of genetics in order to apply the law and legal concepts in cases that involve genetics

III-A. Medicine and Public Health

- Genetic information can be used to guide the **prevention**, **diagnosis**, and **treatment** of disease



- **Personalized/precision medicine**
  - Alternative to "one size fits all" health care that takes into account individual differences in people's genes, environments, and lifestyles
    - Goal of the Precision Medicine Initiative and "All of Us" Research cohort: to lay the groundwork for clinical implementation of genetics and precision medicine


GENETIC TESTING
NHGRI FACT SHEETS
genome.gov
Genetic Tests Can Help to:
Diagnose Your Disease
Pinpoint Genetic Factors That Caused Your Disease
Predict How Severe Your Disease Might Be
Choose the Best Medicine and Correct Dose
Discover Genetic Factors That Increase Your Disease Risk
Find Genetic Factors That Could Be Passed to Your Children
Screen Newborns for Certain Treatable Conditions
NIH National Human Genome Research Institute

Variant of Uncertain Significance

- Clinical interpretation of a genetic variant:
  - Pathogenic
  - Likely pathogenic
  - VUS
  - Likely benign
  - Benign

- VUS: Common problem in clinical sequencing
  - e.g. 5% of results for BRCA1/2 breast cancer; can go up to 20% for other diseases
  - Sequencing DNA is now easy compared to interpreting

Legal application: liability for labs that don't update a VUS in a clinical report once more info is available?

https://www.ashg.org/education/csertoolkit/uncertainresults.html

Public Health

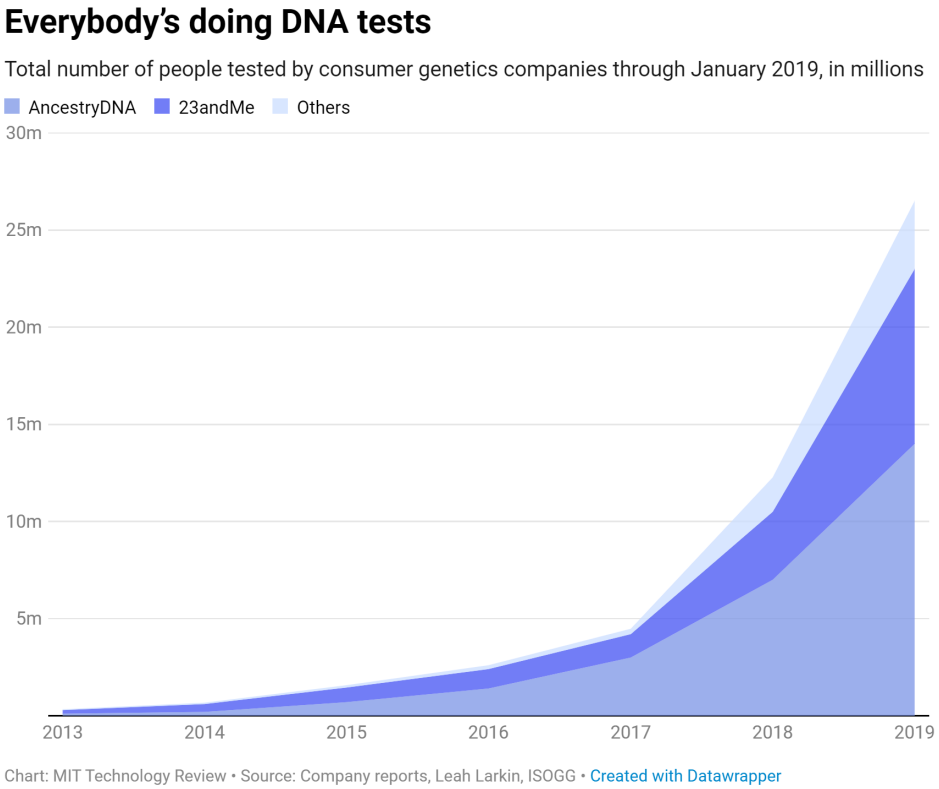
- Newborn screening
  - Best example of large scale public health genetic screening in US
  - State run
  - Blood test w/in 1-2 days of birth
  - Began in 1960's
  - Recommended Uniform Screening Panel: 32 conditions

http://www.biopoliticaltimes.org/article.php?id=5477

- Other applications
  - Infectious disease surveillance
  - Characterizing and tracking outbreaks

III-B. Consumer Genomics

- Genetic testing available direct-to-consumer (DTC) since ~2007
- Industry estimates: > 26M people tested



Vayana et al. 2012; Regalado 2019

DTC genetic testing process

Customer → **DTC company**

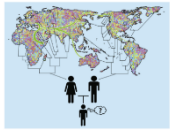
(1) Mail saliva sample using company-provided spit kit

(2) Extracts customer's DNA

(3) Measures (genotypes) DNA at ~1M variants
Using SNP microarray technology

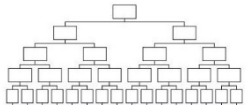
(4) Analyzes **subset** of genotyped variants to provide **interpreted reports**:

Genetic ancestry



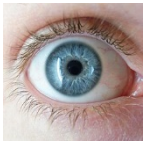
ancestryDNA, FamilyTreeDNA, 23andMe, MyHeritage

Genealogy



ancestryDNA, FamilyTreeDNA, 23andMe, MyHeritage

Traits



23andMe

Health

23andMe

Diet & fitness

DNAFit, ATHLETIGEN

(5) Provides to costumer:

- **interpreted reports**
- **raw data** file of all ~1M measured variants

POLICY

Hey, soldiers and spies — think twice about that home genetic ancestry test

Lawmakers appear to be concerned that China could access genetic and health data of U.S. soldiers and secret agents through home ancestry tests

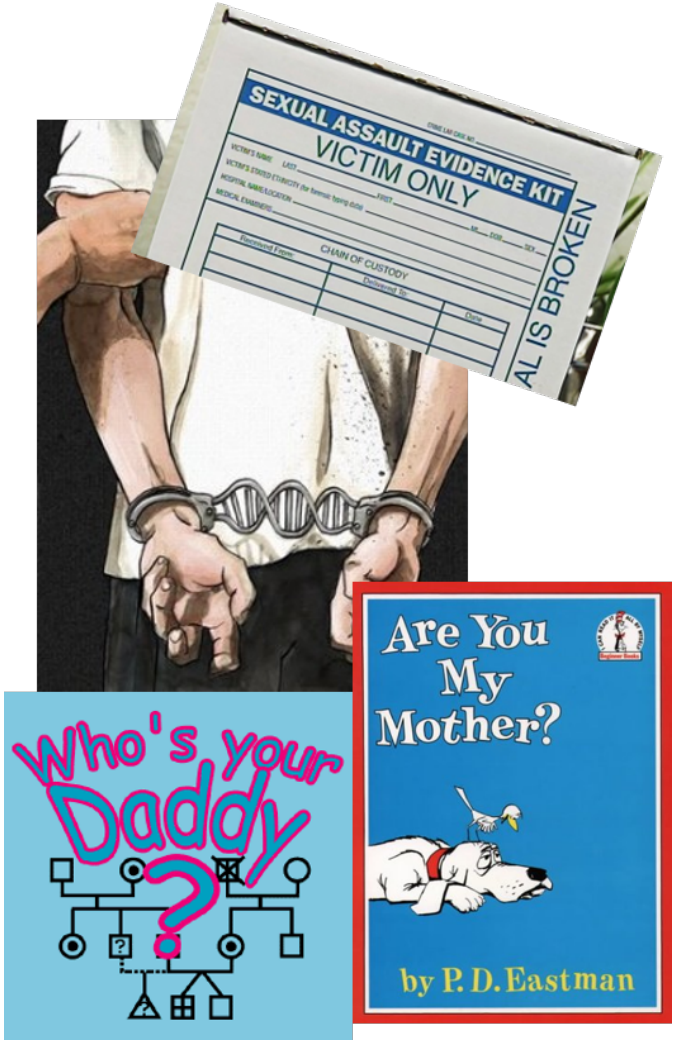
In this photo illustration, an Ancestry.com logo is displayed on a smartphone. (SOPA Images/LightRocket via Getty Images)

- Omnibus spending package passed by Congress end of Dec 2020
- Asks GAO to examine risks to intelligence community and military from use of DTC tests

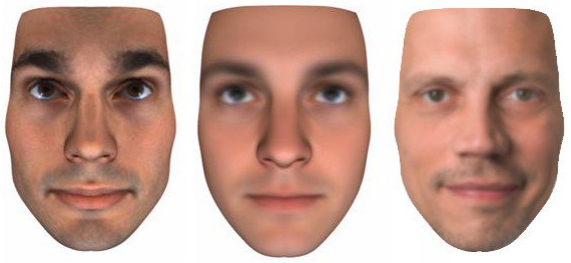
https://www.rollcall.com/2020/12/24/hey-soldiers-and-spies-think-twice-about-that-home-genetic-ancestry-test/

III-C. Forensic Uses

- **Relatedness**
  - Determine kinship, including paternity
- **Comparative DNA Profiling**
  - Evaluation of the extent of genetic similarity between a known individual and an unidentified sample (eg. Blood stain from crime scene)
  - Match probability based on population allele frequencies
- **ELSI Concerns**
  - Familial searches?
  - Use of consumer-facing databases (GEDmatch)?
  - DNA samples taken at arrest?
  - Overrepresentation of racial and ethnic minorities in criminal databases?



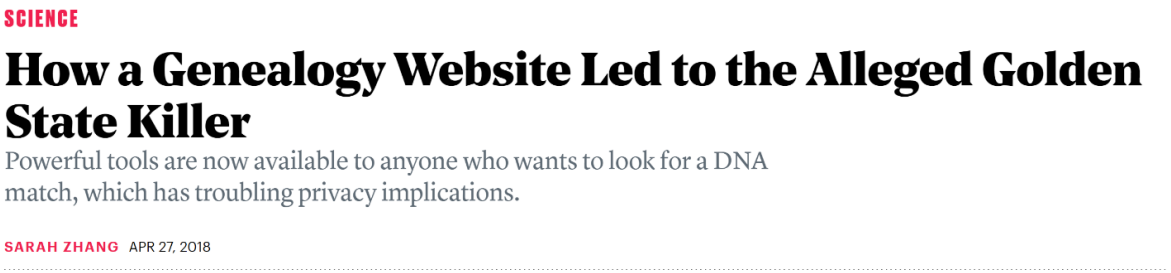
Phenotypic profiling



- Predicting facial features based on DNA
- Controversial “face prediction” paper in 2017
  - J. Craig Venter’s company, Longevity
  - Heavily critiqued by scientific community
- Implications for DNA forensics in courts:
  - Admissibility as evidence?
  - Standards for evaluating evidence?

https://www.technologyreview.com/s/608813/does-your-genome-predict-your-face-not-quite-yet/

Consumer genomics meets forensics



SCIENCE

How a Genealogy Website Led to the Alleged Golden State Killer

Powerful tools are now available to anyone who wants to look for a DNA match, which has troubling privacy implications.

SARAH ZHANG APR 27, 2018

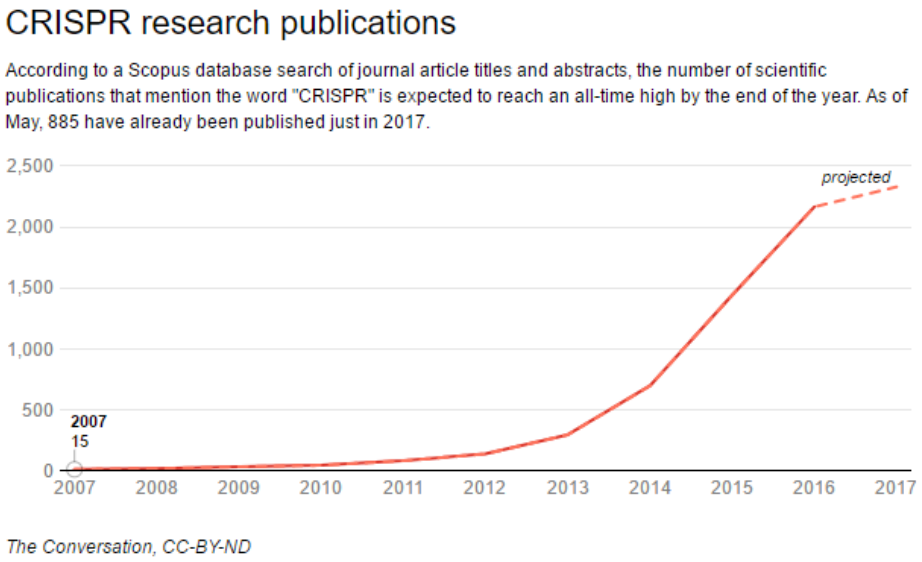
- April 2018: law enforcement announce use of consumer-facing, publicly available genealogy database (GEDMatch) to lead to arrest in Golden State Killer case
  - Via “spoofing” the data format used by DTC companies
- Opened floodgates on use of this technique (vs. government DNA databases such as CODIS)
- Over 50 “cold cases” solved

Erlich et al. Science 2018

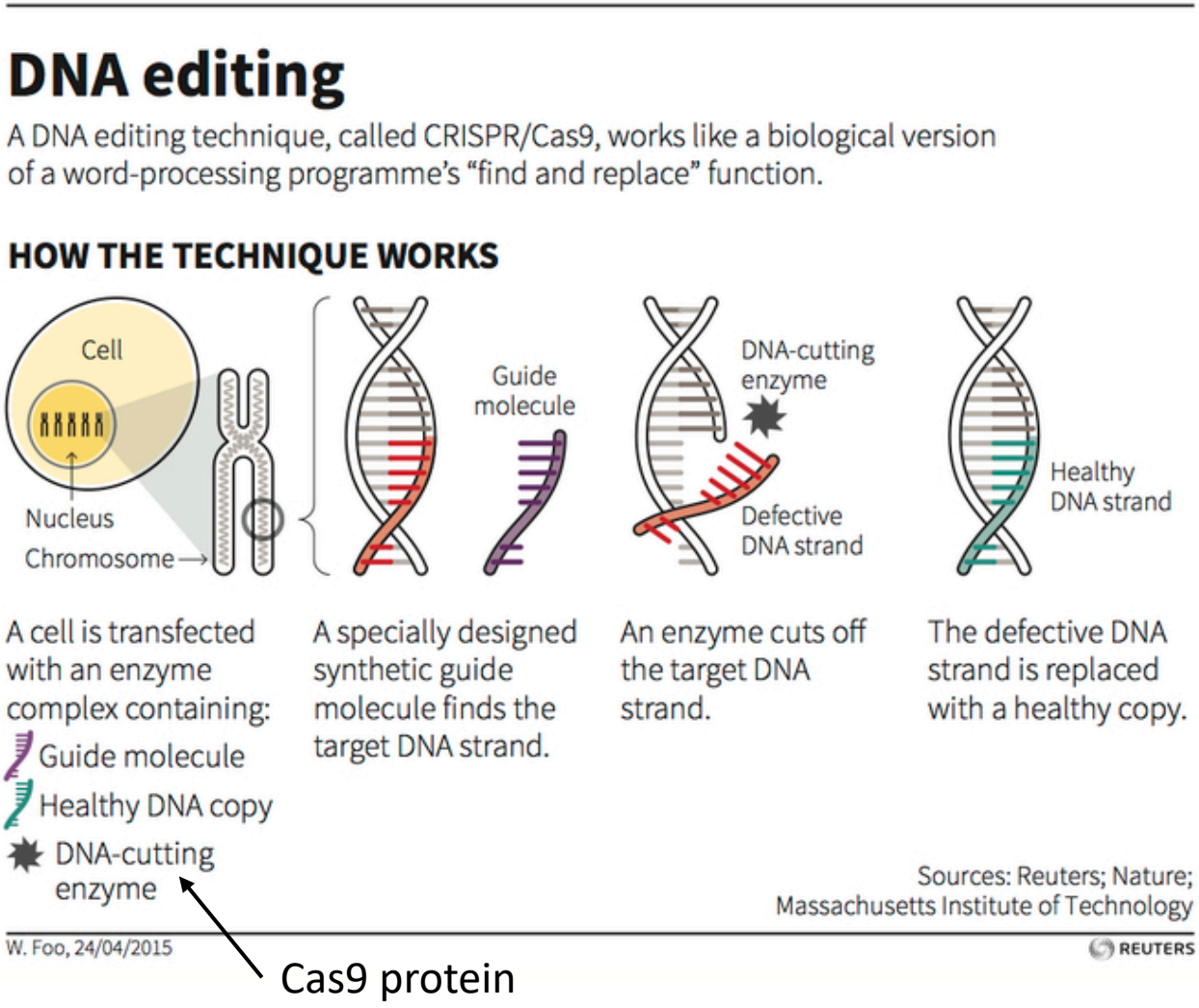
III-D. Biotechnology

Focusing on a BIG game changer…

- **CRISPR**
  - ="clustered regularly interspaced short palindromic repeats"
  - Based on natural bacterial defense system
  - Gene-editing technique harnessed in past decade



CRISPR: how does it work?



Cas9 protein

Compared to the original DNA, the inserted DNA sequence can be a deletion, insertion, OR single base pair change

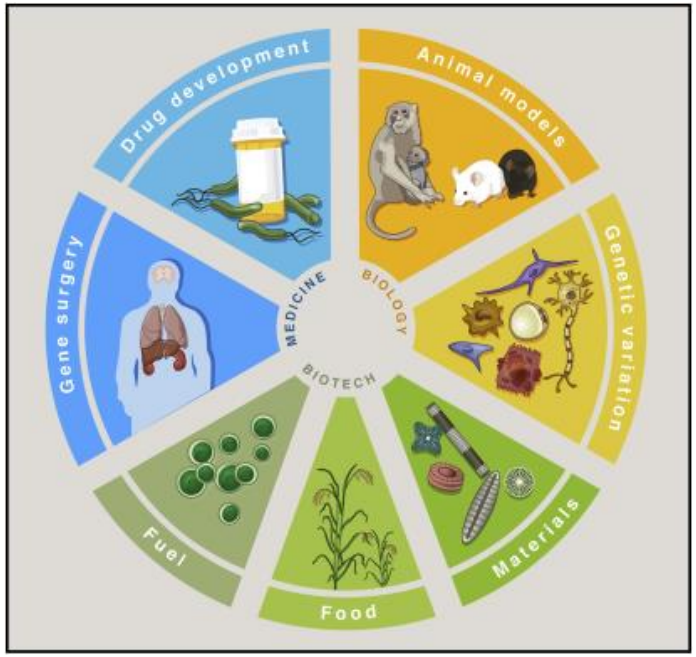
Incredibly precise compared to previous gene editing techniques "Molecular scissors"

Relatively "easy" (cost, time, materials required)

"transfection" = getting cell to express the Cas9 protein and guide RNA molecule. May involve viral vectors

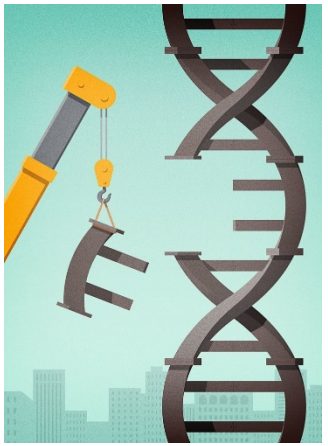
CRISPR: how can it be used?

- Many possible applications in research and medicine, in both humans and non-humans e.g.:
  - Generate “gene knockout” animal models for research
  - Treat infectious diseases
  - GMO/GM crops



Hsu et al. Cell (2014); Yang Military Medical Research (2015); https://phys.org/news/2017-05-crispr-lab-today.html

CRISPR: how can it be used? (cont)

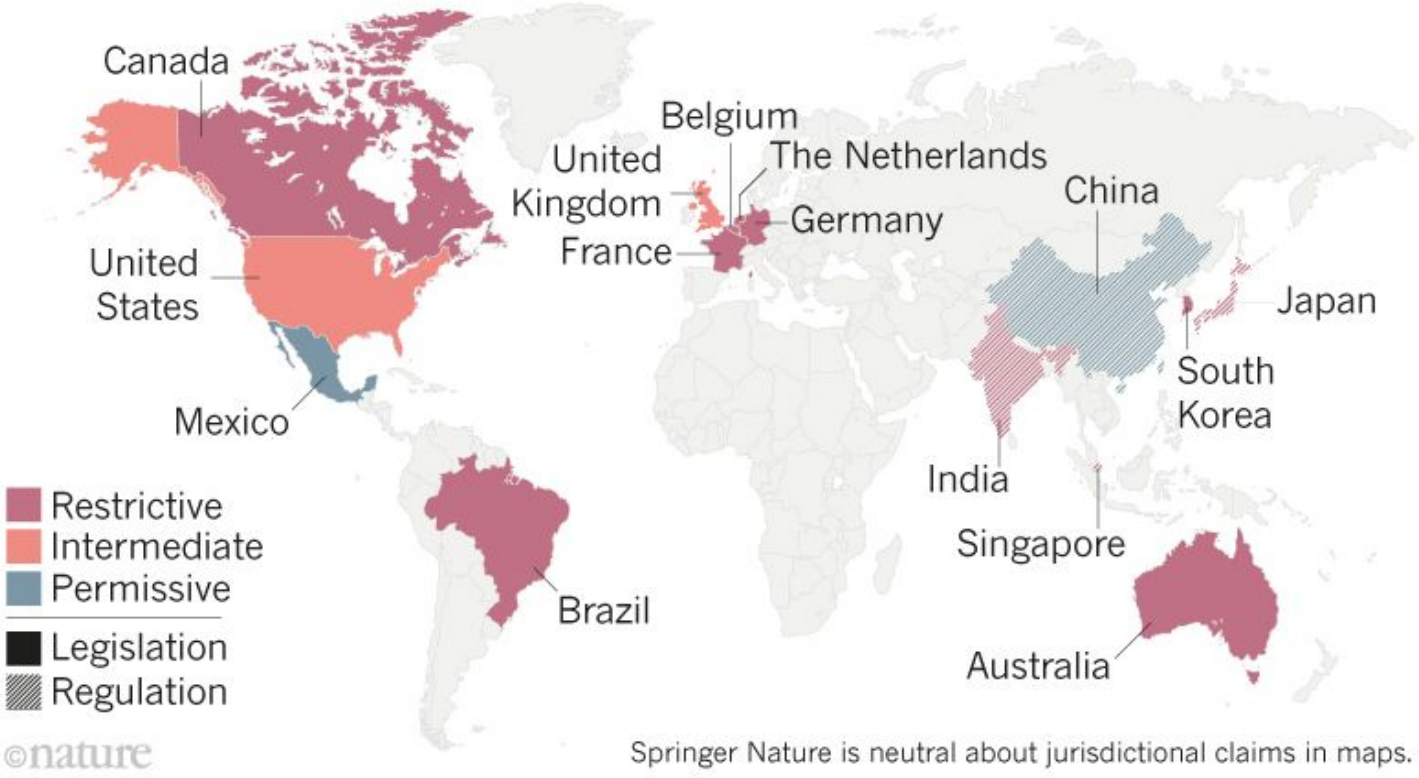


- Gene therapy/gene transfer
  - Concept not new: first clinical trials > 20 yrs ago (not much success)
- "Correcting" genetic diseases
  - Somatic gene therapy *changes NOT passed down
  - Germline gene therapy *changes passed down
    - Nov 2018: claim of first "CRISPR" babies in China
- First FDA approvals for CRISPR drugs started in 2017
  - E.g., Dec 2017: Luxturna, the first gene therapy for an inherited disease, a form of blindness
    - Expected to cost $850k/tx

https://www.technologyreview.com/s/612458/exclusive-chinese-scientists-are-creating-crispr-babies/

THE LEGAL LANDSCAPE

A 2016 survey in *Science* examined existing laws (legislation) and documented policies (regulation) that explicitly govern gene editing or might be applied to such activities. The survey labelled countries as restrictive, permissive or something in between. But specialists disagree over whether rules in some nations might be intepreted to permit gene editing.



Source: R. Isasi *et al. Science* **351**, 337–339 (2016).

General references

National Human Genome Research Institute (NHGRI) resources:

- Talking Glossary of Genetic Terms, https://www.genome.gov/glossary/
- Genetics, DNA, and Heredity: The Basics, https://www.genome.gov/pages/education/modules/basicspresentation_vs2.pdf
- Issues in Genetics, https://www.genome.gov/10000006/issues-in-genetics/
- Fact Sheets: https://www.genome.gov/10000202/fact-sheets/

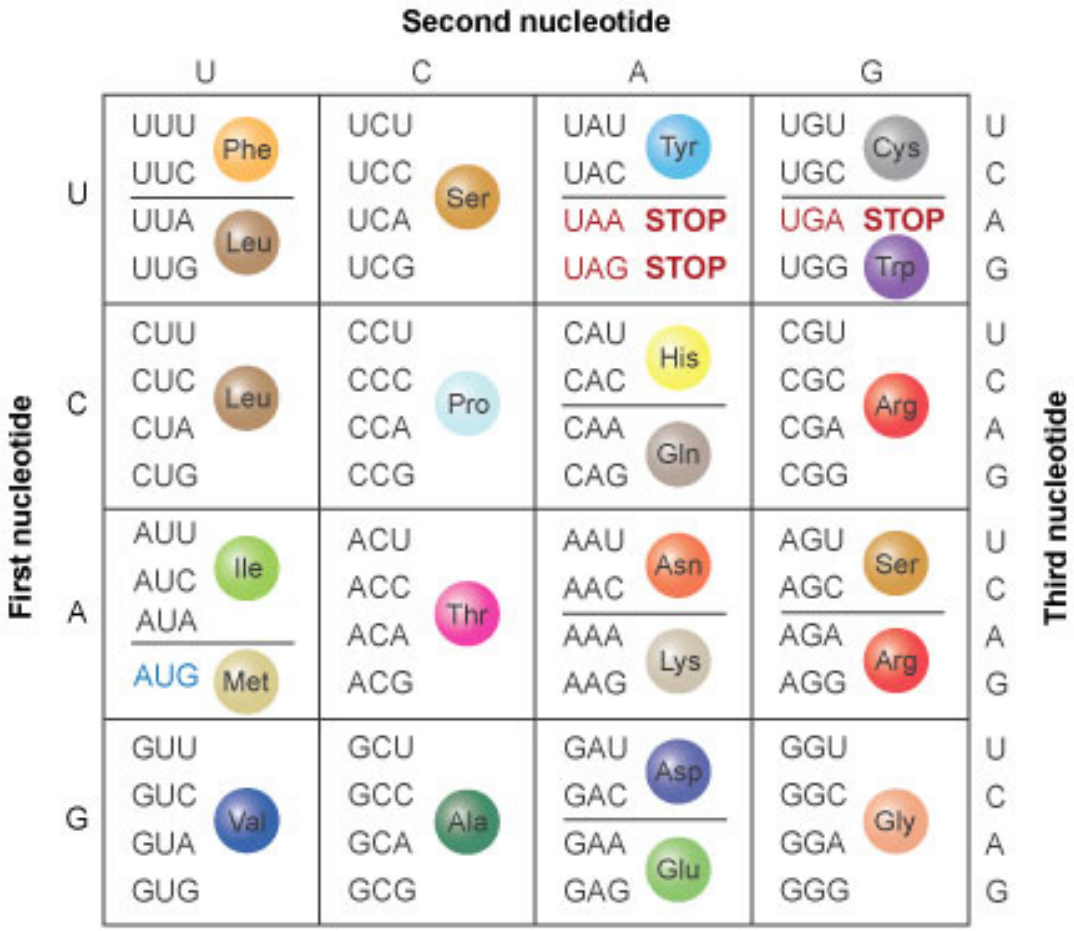
Genetics Home Reference, https://ghr.nlm.nih.gov/

University of Utah, Learn Genetics, http://learn.genetics.utah.edu

Supplementary Slides

Codons

Base pair triplets ("**codons**") specify different **amino acids**, the building blocks of all **proteins**



Images; https://www.nature.com/scitable/topicpage/nucleic-acids-to-amino-acids-dna-specifies-935/

Cell Division

Mitosis

one cell divides once to form **two** identical, **diploid** daughter cells

Meiosis

one cell divides twice to form **four** non-identical **haploid** cells

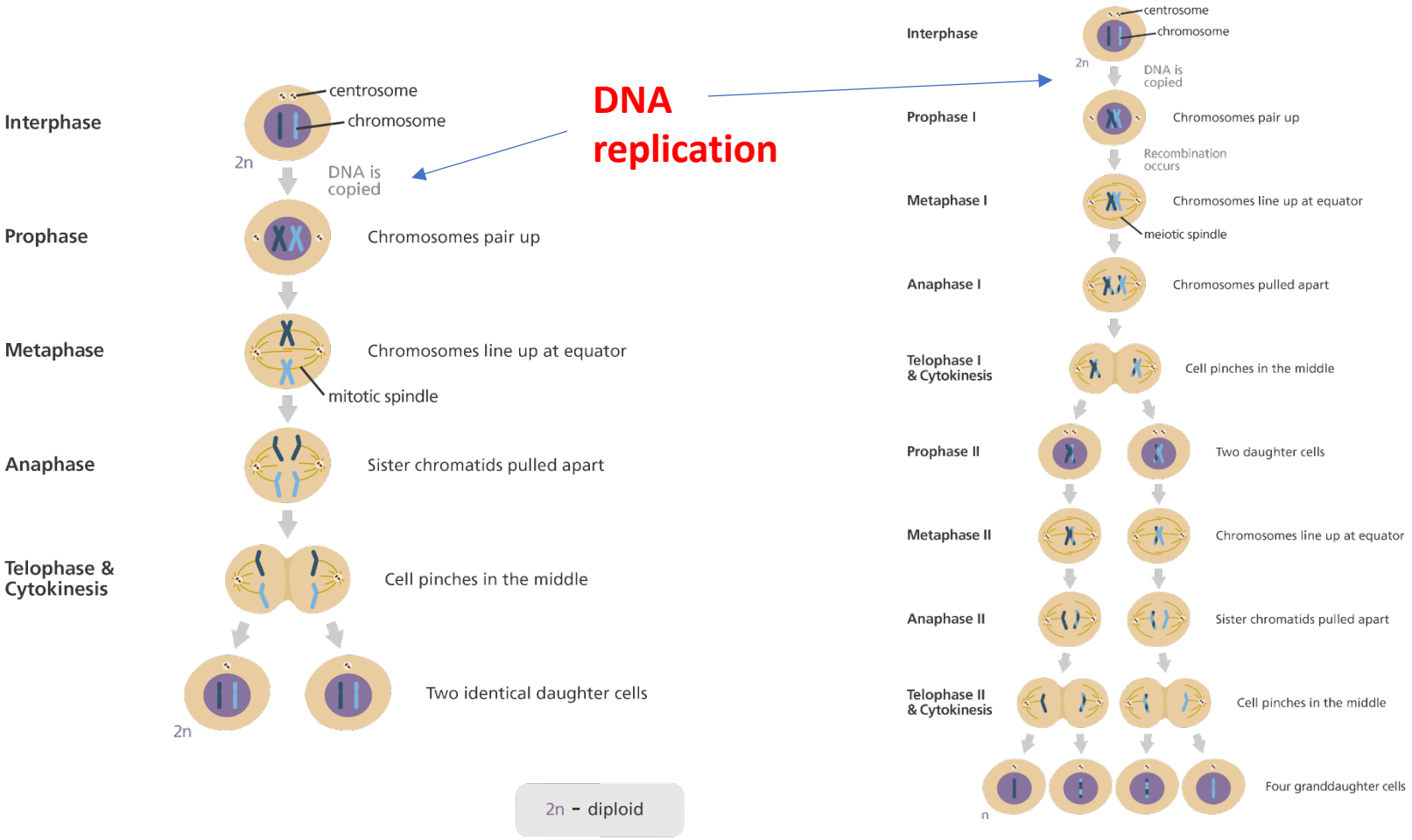
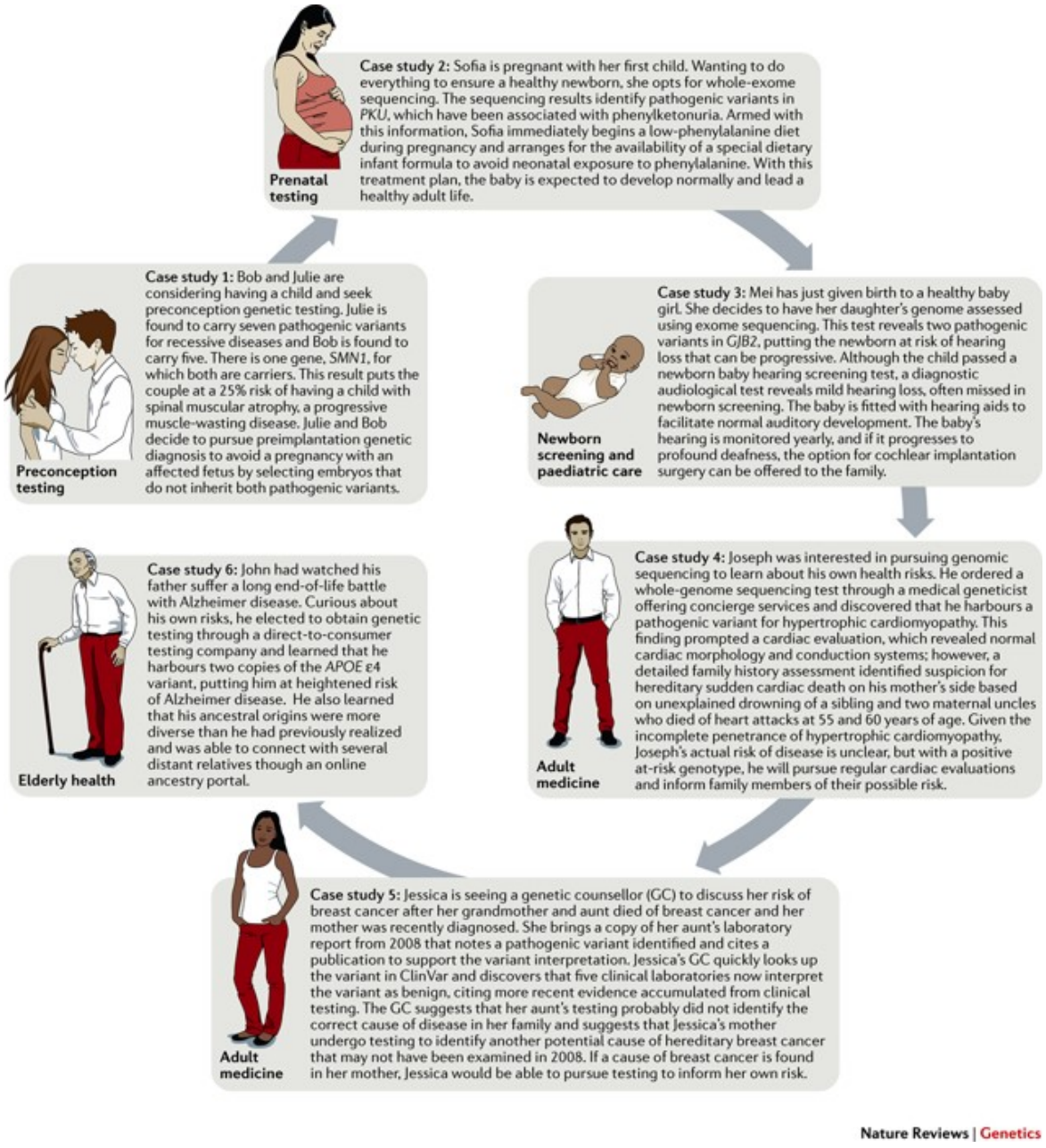


Image credit: Genome Research Limited

Clinical Genetic Testing



Rehm Nat Rev Gen 2017

Race, ethnicity, and identity

physical characteristics

categorization

biology

racism

sociology

psychology

oppression

discrimination

skin color

difference

social construction

language

www.radiolab.org

divisive

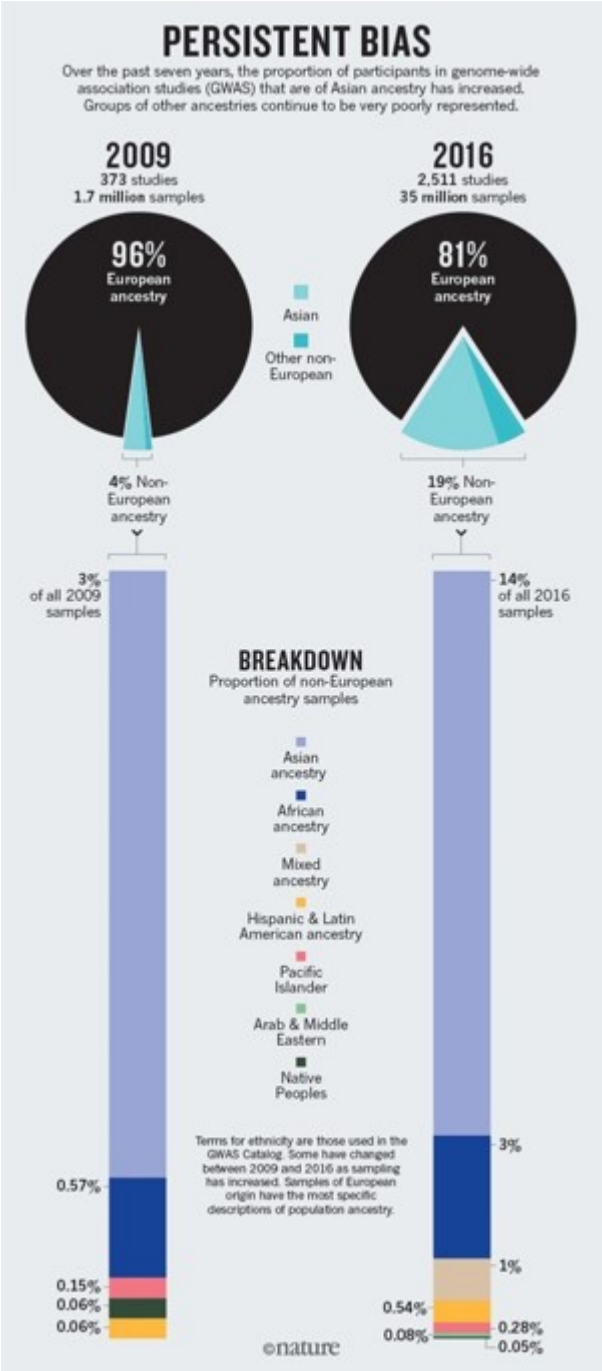
history

anthropology

culture

colonialism

inferiority/superiority



Genetic research and diversity

Genetic research and health disparities

New Online Views 6,785 | Citations 0 | Altmetric 0

Viewpoint

ONLINE FIRST FREE

March 27, 2017

More

Genomics, Health Disparities, and Missed Opportunities for the Nation's Research Agenda

Kathleen McGlone West, MS, MPH[1]; Erika Blacksher, PhD[2]; Wylie Burke, MD, PhD[2]

» Author Affiliations | Article Information

JAMA. Published online March 27, 2017. doi:10.1001/jama.2017.3096

The completion of the Human Genome Project occurred at a time of increasing public attention to health disparities. In 2004, Sankar and colleagues[1] suggested that this coincidental timing resulted in an inappropriate emphasis on the contribution of genomics to health disparities, conflating racial patterns of disease with genetic ancestry, and distracting attention from the large and compelling body of scientific evidence pointing to social determinants of health disparities.[2] For example, genomic research has emphasized discovery of genetic contributors to diabetes risk, but the recent increase in the prevalence of obesity and type 2 diabetes, which disproportionately affects minority populations, cannot be attributed to genetic changes and rather

Ownership of DNA

- Who owns your genetic information?
  - Court precedent that donors to do not have property rights in their biological samples
    - E.g., Moore v. Regents of the University of California (1990)
  - Many direct-to-consumer companies sell customer data to third parties
    - Genetic data often more useful/valuable in the aggregate, e.g., as an R&D tool, than it is for the individual
- Ongoing case, Peerenboom v. Perlmutter, that may go against Moore precedent
  - See Genome Magazine Winter 2018 article, “Do You Belong to you?”

http://genomemag.com/do-you-belong-to-you/

Genetic Discrimination

- The Genetic Information Nondiscrimination Act (GINA) of 2008
  - Protects Americans from discrimination based on their genetic information in both health insurance (Title I) and employment (Title II)
- Does not apply to other insurance (e.g., life, long term disability)
- Legal issue: largely untested in court

https://www.genome.gov/10002077/genetic-discrimination/genetic-discrimination/

Ploidy

- Diploid
  - Two **sets** of chromosomes (2n)
    - One from father (paternal)
    - One from mother (maternal)
  - Human (non-germ) cells are diploid
    - Most Eukaryotes (species whose cells have nuclei) are diploid
- Haploid
  - One set of chromosomes (1n)
  - **Germ cells** (sperm and egg) are haploid
- Polyploidy
  - Multiple sets of chromosomes (>2n)
  - Common in plants
    - Fun fact: strawberries are octaploid (8n)

“CRISPR baby” controversy

- First gene-edited babies reported Nov 2018, from an academic group in Shenzhen, China
  - Twin girls born
- Edited *CCR5* gene in embryos prior to transplant to woman’s uterus
  - Attempt to create resistance to HIV, smallpox, and cholera
- Generated much controversy in scientific and medical communities
- Dec 2019: the lead scientist (He Jiankui) sentenced to 3 yrs in prison and ~$430K fine for “illegal medical practice by knowingly violating the country's regulations and ethical principles with their experiments”

Kennedy 2019, https://www.npr.org/2019/12/30/792340177/chinese-researcher-who-created-gene-edited-babies-sentenced-to-3-years-in-prison
Regalado 2018, https://www.technologyreview.com/s/612458/exclusive-chinese-scientists-are-creating-crispr-babies/

CRISPR: how does it work?