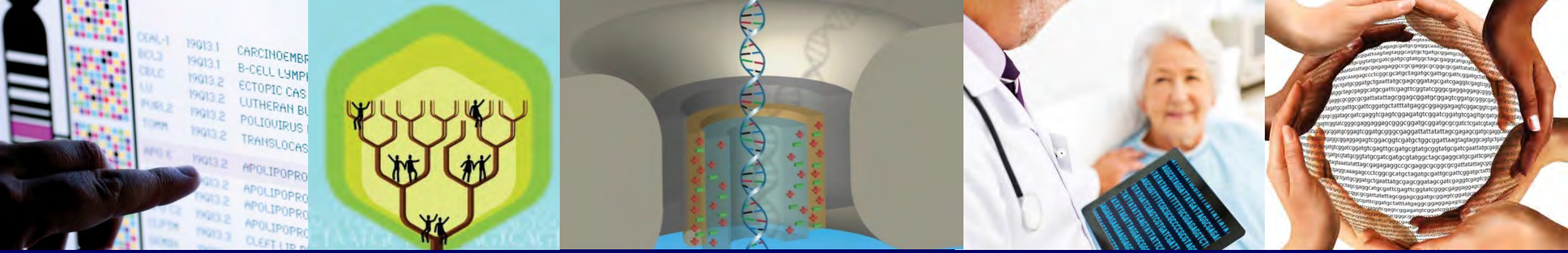


At *The Forefront of Genomics*: Making Genomic Medicine a Reality

Eric Green, M.D., Ph.D.

Director, NHGRI

National Human Genome Research Institute

The Forefront of Genomics®

Two Scientific Fields Launched Last Century Are Changing Medicine This Century

"Genetics"

London

ROYAL HORTICULTURAL SOCIETY.

FOUNDED A.D. 1804. INCORPORATED A.D. 1809.

VINCENT SQUARE, LONDON, S.W.

REPORT

OF THE

THIRD INTERNATIONAL CONFERENCE

1906

ON

GENETICS:

HYBRIDISATION

(THE CROSS-BREEDING OF GENERA OR SPECIES),

THE CROSS-BREEDING OF VARIETIES,

AND GENERAL PLANT-BREEDING.

EDITED BY THE

REV. W. WILKS, M.A.

SECRETARY.

LONDON:

Printed for the Royal Horticultural Society

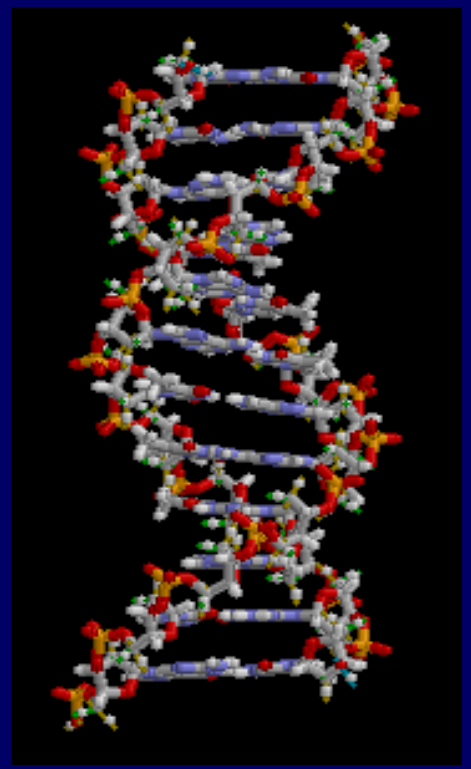
BY

SPOTTISWOODE & CO. LTD., NEW-STREET SQUARE, E.C.

Digitized by Google

1907

DNA's Double Helix



1953

"Genomics"

GENOMICS 1, 1–2 (1987)

EDITORIAL

A New Discipline, A New Name, A New Journal

In recent times there has been a rallying call for complete mapping/sequencing of the human genome. Technical advances in mapping beginning 20 years ago and in sequencing 10–15 years ago have made this feasible or at least conceivable. The two operations—mapping and sequencing—have the same objective, namely, analysis of the structure and organization of the human genome. Mapping determines the general location of genes on chromosomes and their positions relative to each other. The nucleotide sequence is the ultimate map. The two operations must go hand in hand. For example, the mapping of segments of DNA, e.g., overlapping cosmid clones, is seen as a desirable initial step for efficient sequencing of the human genome. Blind sequencing is not likely to be as efficient, and certainly not as interesting, as sequencing the expressed parts of the genome, whose chromosomal location is known. Mapping all expressed genes, cloned through their messenger RNAs, regardless of whether their function is known, sequencing these genes together with their introns, and sequencing out from these is seen by many as "the way to go." The ultimate map, the sequence, is seen as a rosetta stone from which the complexities of gene expression in development can be translated and the genetic mechanisms of disease interpreted.

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. Genomics involves workers competent in constructing and interpreting various types of genomic maps and interested in learning their biologic significance. Genetic mapping and nucleic acid sequencing should be viewed as parts of the same analytic process—a process intertwined with our efforts to understand development and disease.

In his essay entitled "What is Semantics?", Anatol Rapoport wrote:

> There are two suffixes in our language (and similar ones in other European languages) which suggest organized knowlege. One is the venerable, academic "ology," that reminds one of university curricula and scholarship. The other is the energetic and somewhat mysterious "ics," which has a connotative flavor of magic. Where "ology" suggests academic isolation (ichthyology, philology) "ics" suggests a method of attack on life's problems. It contains a faint throwback to the ancient dreams of the philosopher's stone and of "keys" to the riddles of the universe. Ancient words ending in "ics" are mathematics and metaphysics. Of more recent origin are economics, statistics, semantics, and cybernetics.

One might add *genetics*, and now, *genomics*.

While we are on words: *Genome* is an irregular hybrid of *gene* and *chromosome*. Both parents are Greek. In their *Glossary of Genetics and Cytogenetics*, Rieger, Michaelis, and Green (1976), stated that the hybrid term was first used in 1920 by Winkler, who also introduced the term *conversion* into genetics.

The necessity for communication, coordination, and education in this emerging field dictates the founding of a new journal dedicated to genomics in all of its ramifications. *Genomics* will not only report new data concerning genome maps and improved methods for mapping and sequencing—those will certainly be very important components of the journal—but also will publish analyses of the information, methods for those analyses, methods for storage, retrieval, searching, pattern recognition, comparisons, etc., as well as interpretation of structural findings in light of their biologic significance and biomedical applications.

Genomics will be a common meeting ground for molecular biologists and biochemists, human and somatic cell geneticists, cytogeneticists, population and evolutionary biologists, genetic epidemiologists, clinical geneticists, theoretical biologists, and computational scientists, all interested in the biology and genetics of the human and other complex genomes.

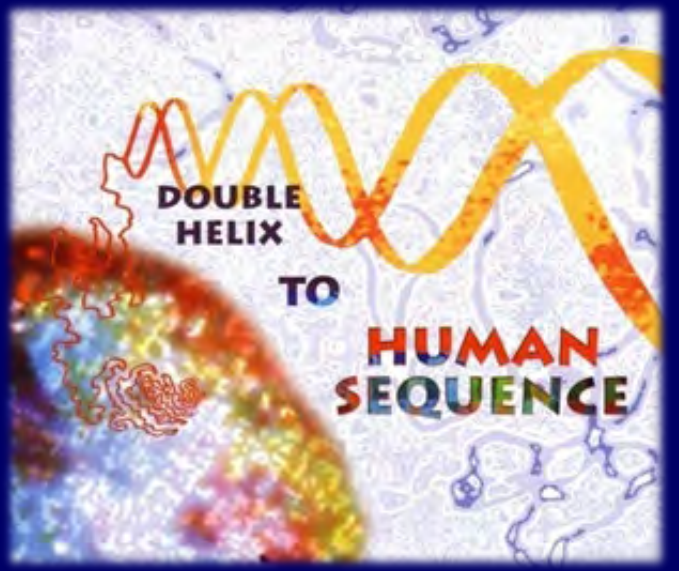
Topic areas for this interdisciplinary forum include:

- Chromosomal assignments of genes and DNA fragments by Mendelian and physical mapping approaches, including the description of new techniques
- Reports of nucleic acid sequences of cloned genes or other interesting portions of a genome
- Descripton of chromosomal and spatial distributions of gene families and genes that share nucleic acid or amino acid sequence domains

Copyright © 1987 by Academic Press, Inc.
All rights of reproduction in any form reserved.

1987

Human Genome Project: 1990-2003

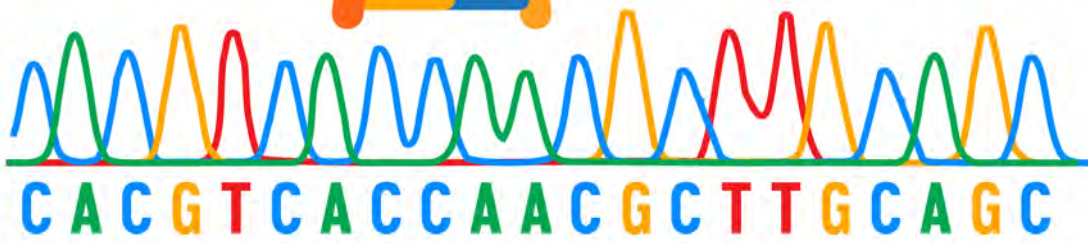


genome.gov/HGP

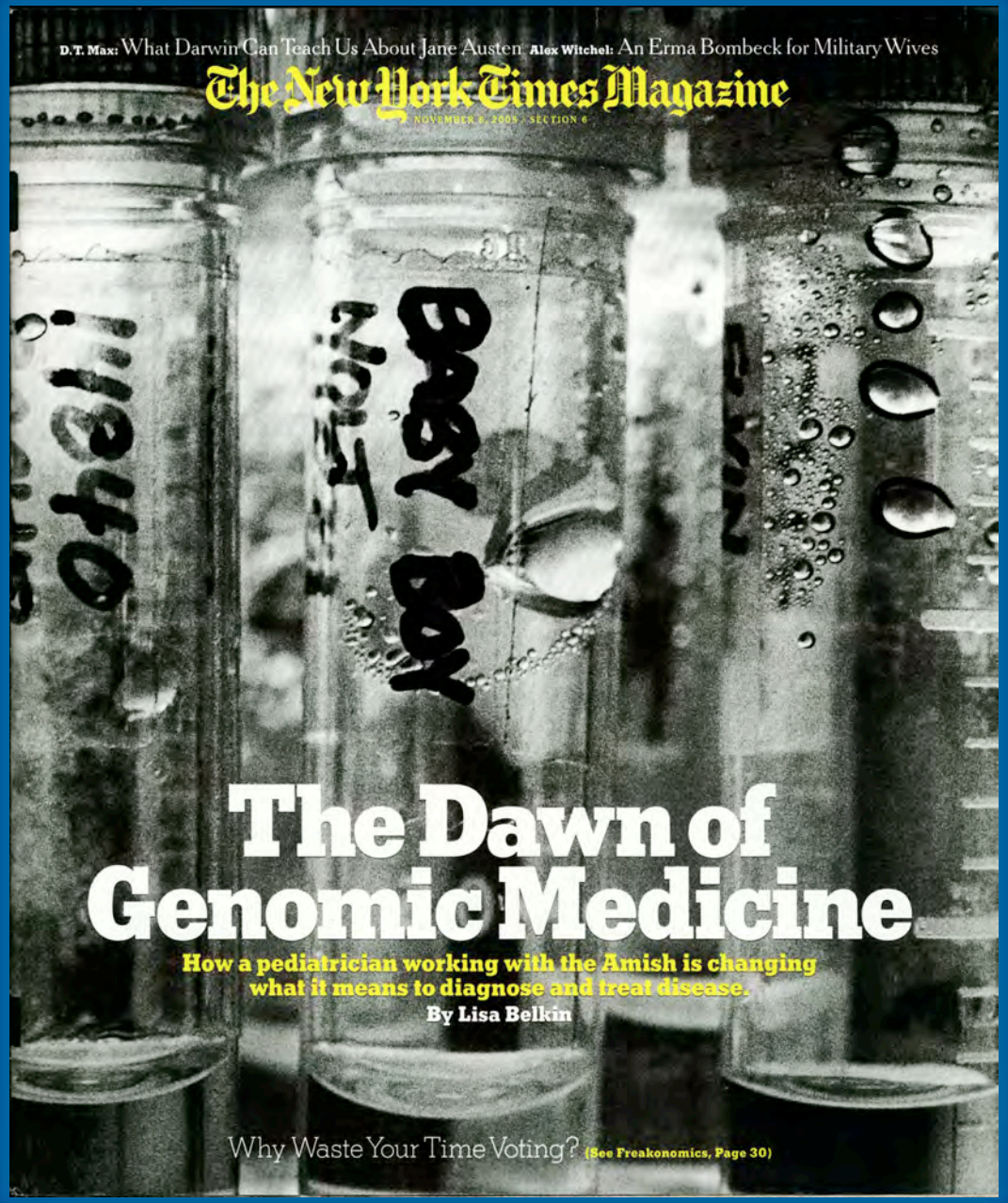
2003

2023

20th Anniversary of the Human Genome Project Completion



NIH National Human Genome Research Institute


D.T. Max: What Darwin Can Teach Us About Jane Austen Alex Witchel: An Erma Bombeck for Military Wives
The New York Times Magazine
NOVEMBER 6, 2005 · SECTION 6
The Dawn of Genomic Medicine
How a pediatrician working with the Amish is changing what it means to diagnose and treat disease.
By Lisa Belkin
Why Waste Your Time Voting? (See Freakonomics, Page 30)


Science
24 October 2003
Vol. 302 No. 5643
Pages 517–728 $10
Genomic Medicine
American Association for the Advancement of Science

Bringing Genomic Medicine Into Focus

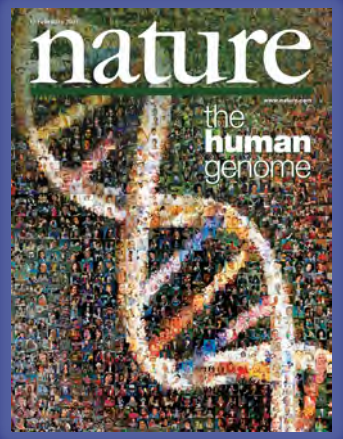
Genomic Medicine

An emerging medical discipline that involves using genomic information about an individual as part of their clinical care (e.g., for diagnostic or therapeutic decision-making) and the other implications of that clinical use

<u>Related (but not identical) terms</u>:

- **Personalized medicine**
- **Individualized medicine**
- **Precision medicine**

The Pivot to Genomic Medicine



Human Genome Project

→ → ? → →

✓ ✓ ✓ ✓

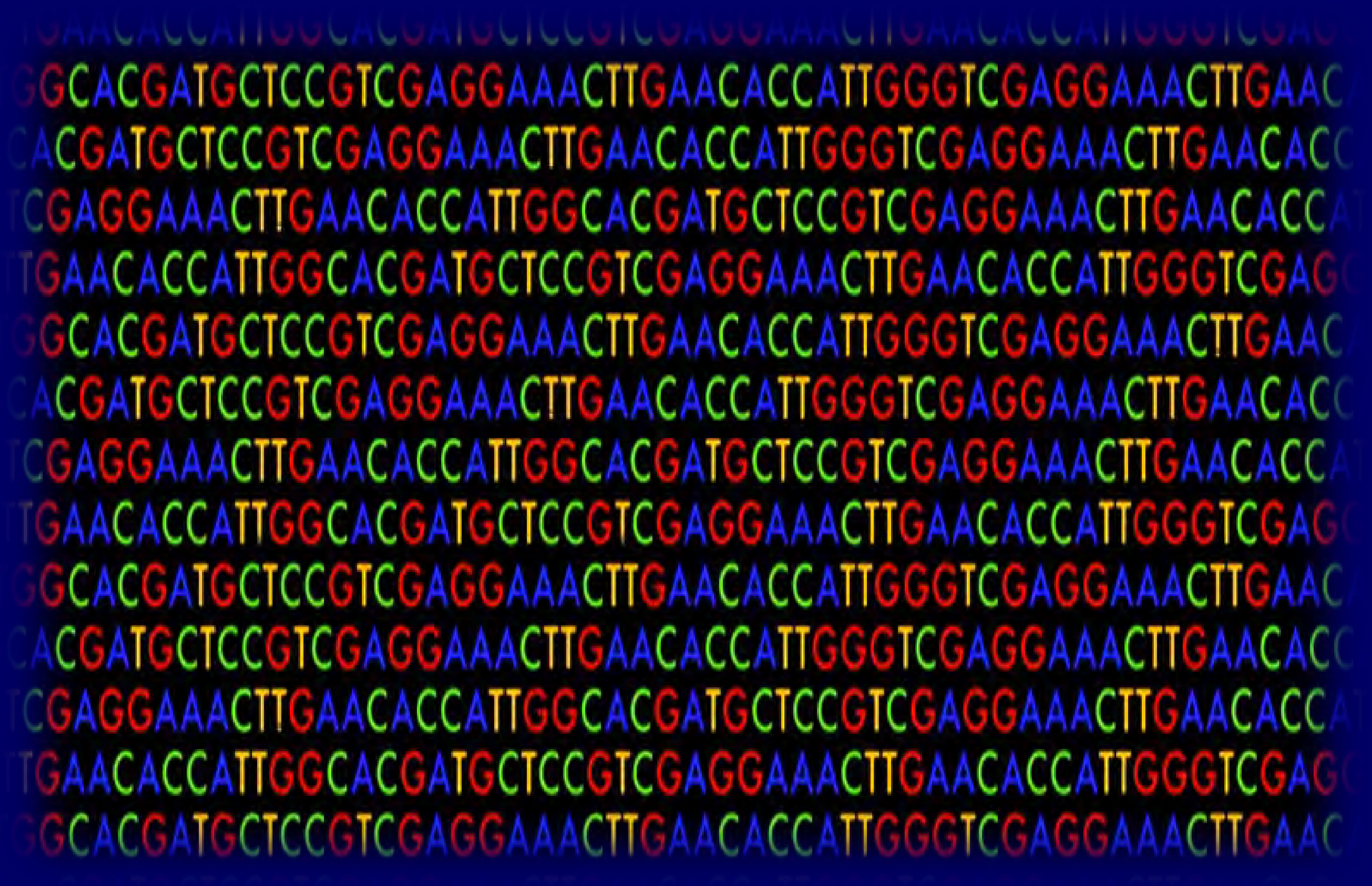
Realization of Genomic Medicine

En Route to Genomic Medicine

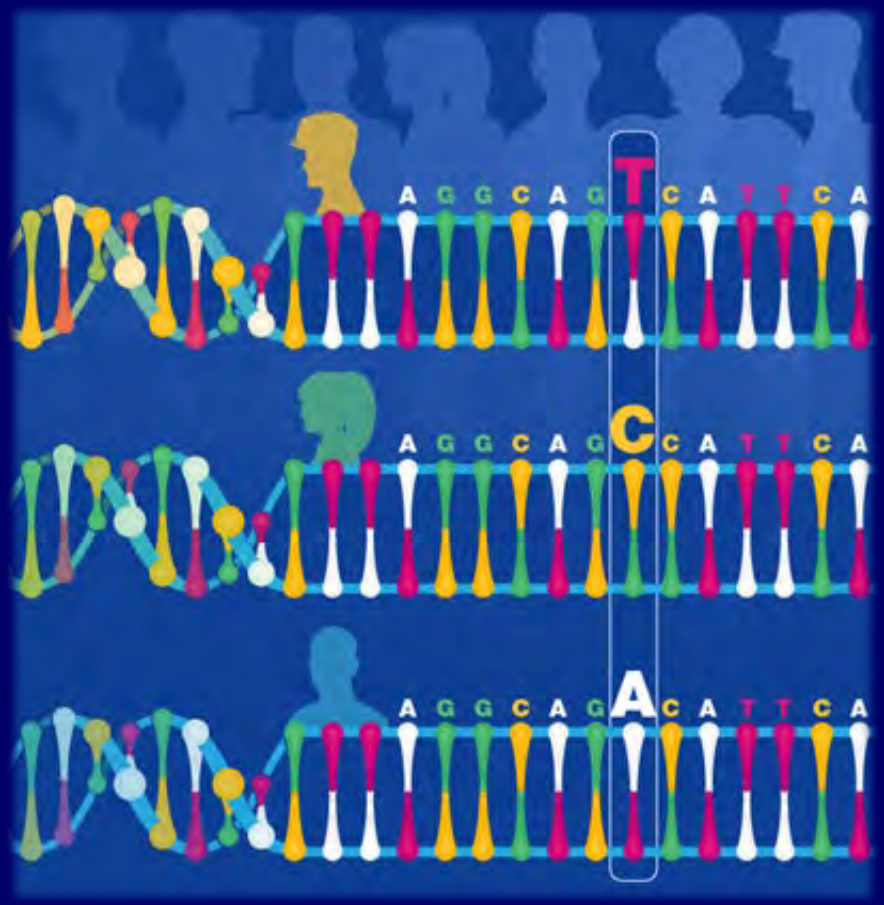
- Human Genome Sequenced for First Time by the Human Genome Project
- Cost of Sequencing a Human Genome Reduced >1 Million-Fold
- Millions of Human Genomes Sequenced
- Profound Advances in Understanding How the Human Genome Functions
- Significant Advances in Unraveling the Genomic Bases of Human Disease

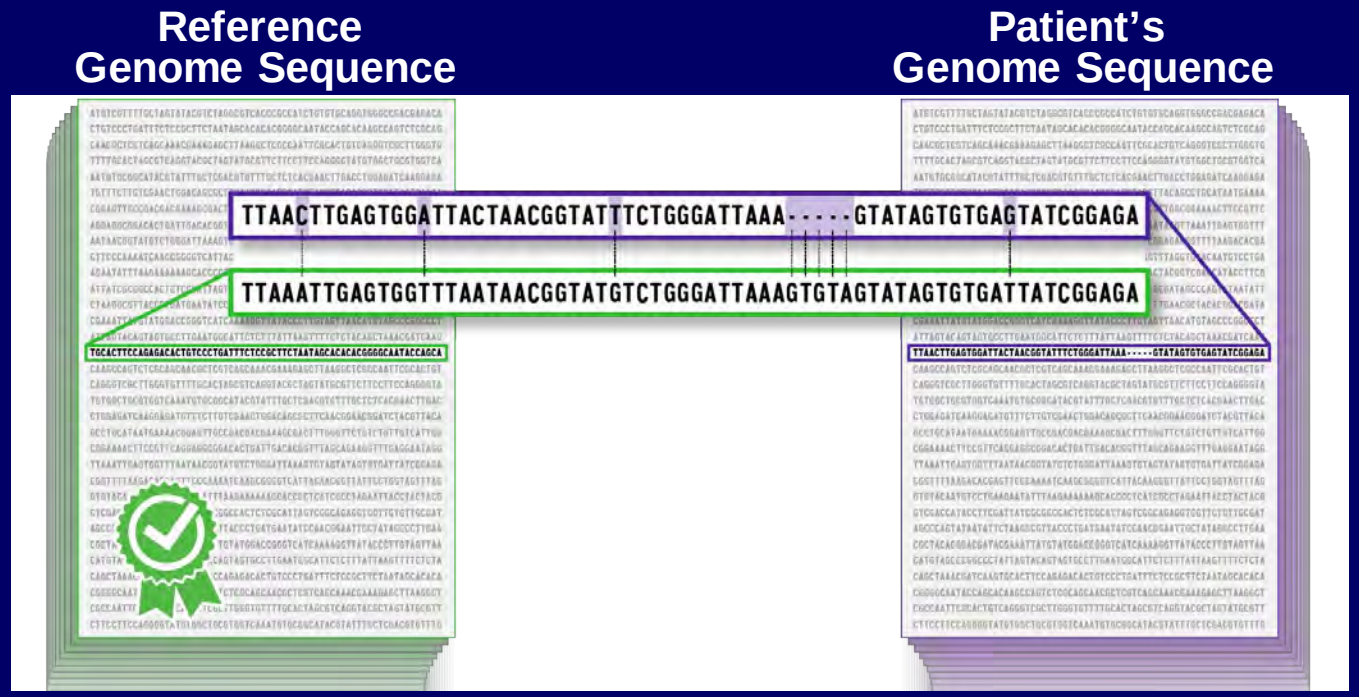
Making Genomic Medicine a Reality

Analyzing a Patient's Genome is Becoming More Routine



Genome Sequencing in Medicine





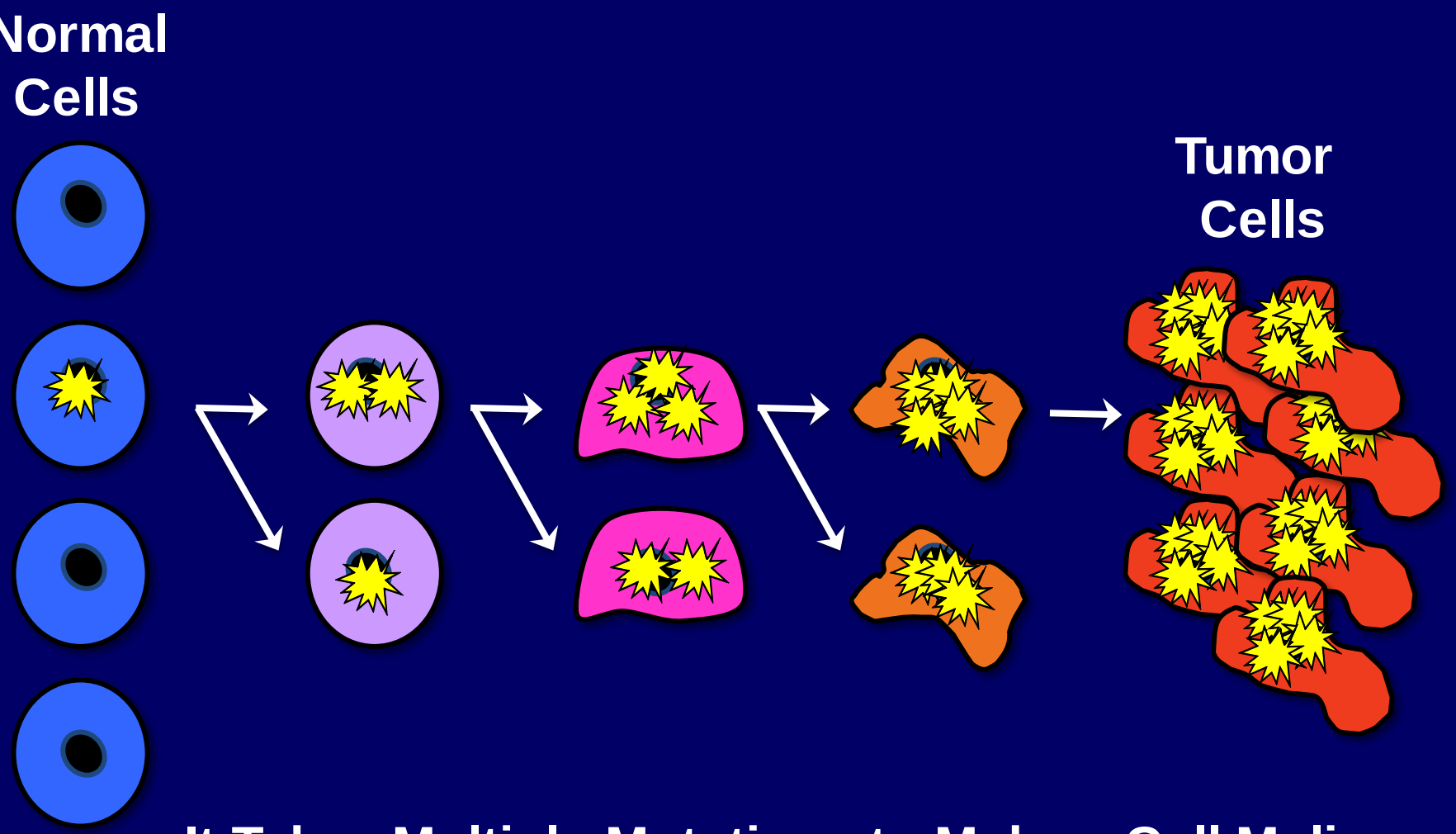
List of Genomic Variants

Implement Genomic Medicine

Genomic Medicine Implementation

Cancer Genomics

Cancer is a Disease of the Genome



It Takes Multiple Mutations to Make a Cell Malignant

Routine Cancer Diagnostic Tools

Cancer Histopathology

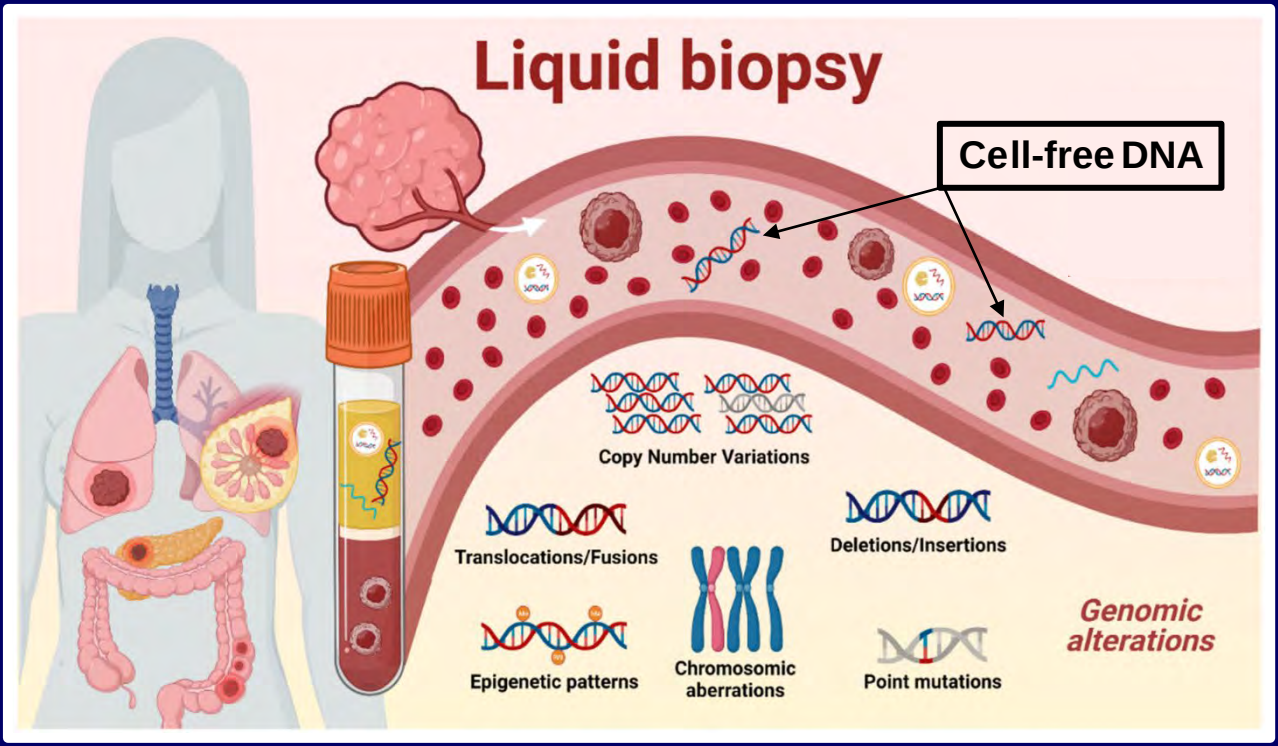
Morphology

Cancer Genome Sequencing

Genomic Signature

<u>Paradigm Change</u>: Genomic signature of a tumor often (perhaps almost always) provides more valuable clinical information than the tissue of origin.

Liquid Biopsy for Detecting Cancer



- Standard biopsies of human tissues are invasive and can be dangerous
- Tumor cells frequently die and release their DNA into the bloodstream
- Highly sensitive DNA-sequencing methods can detect and analyze that cell-free tumor DNA (accessed through simple blood draw)

Cancer Genomics Today

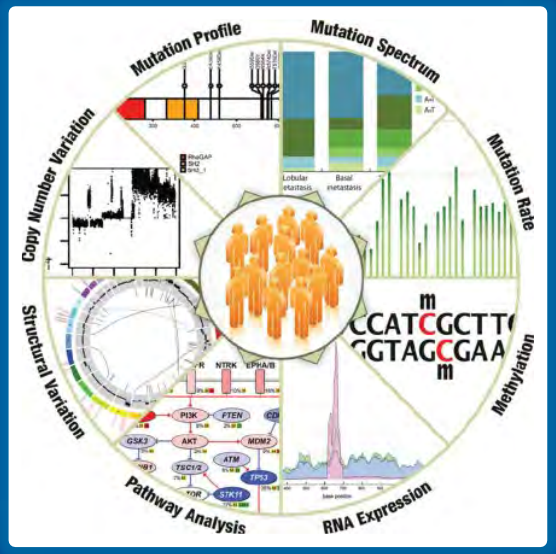
In retrospect

15 years after a giant leap for cancer genomics

Sheng F. Cai & Ross L. Levine

In 2008, the first comprehensive sequence of a cancer genome was reported, ushering in a new era of molecular diagnostic, prognostic and therapeutic advances informed by an essential framework to understand cancer's complexities.

Nature, 2023



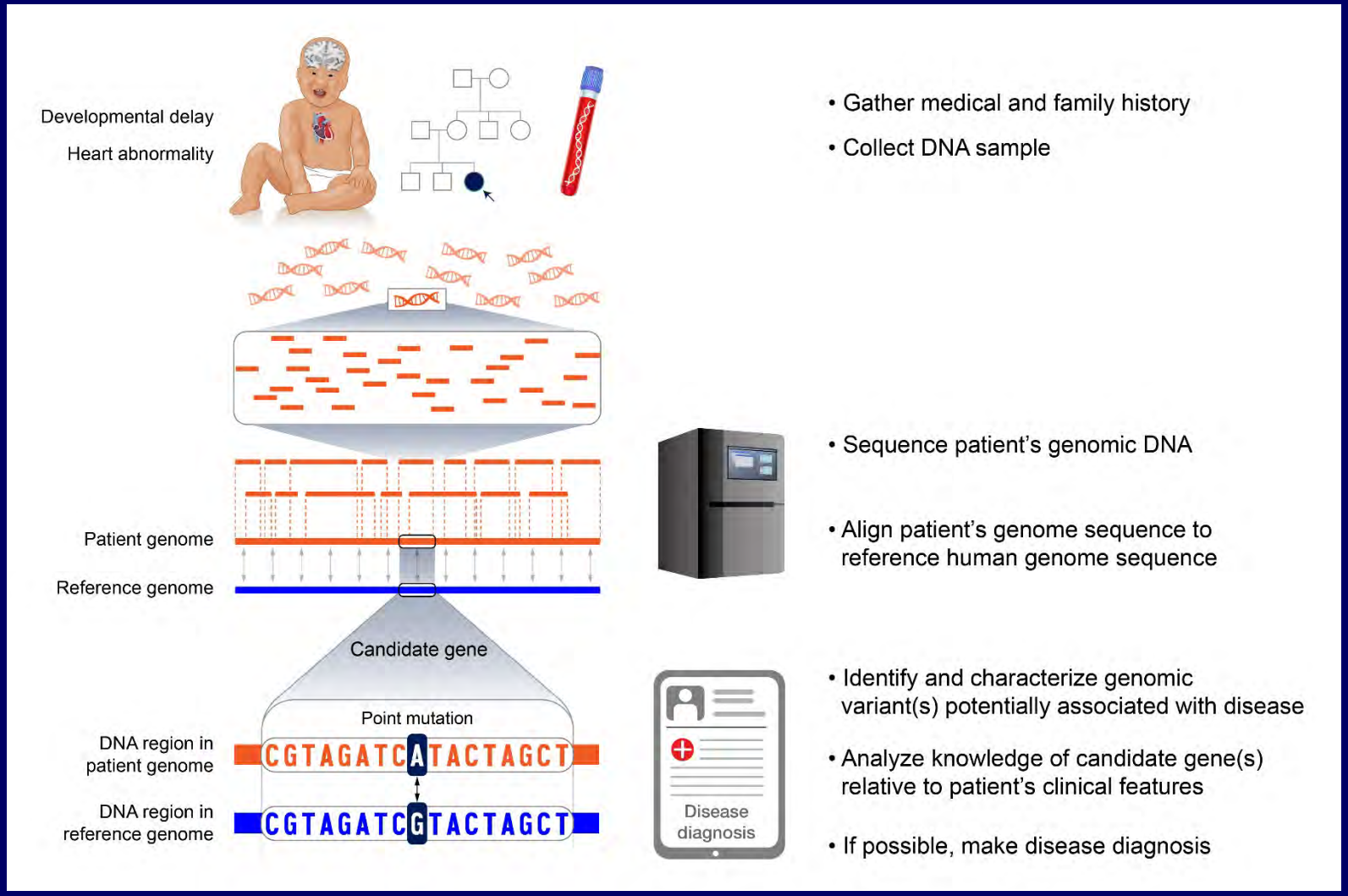
"Today, genomic sequencing as part of clinical care has transformed cancer diagnostics, clinical trials, and the use of new therapies to improve outcomes for people with cancer. Our unprecedented view of the cancer genome empowers clinicians, computational biologists, and bench scientists alike to define biologically relevant groups of people with cancer, direct genomic inquiry, and ultimately identify new therapies and biomarkers."

Genomic Medicine Implementation

Cancer Genomics

Rare Genetic Disease Diagnostics

Genome Sequencing is Becoming a Standard Diagnostic Tool in Medicine



Rare Disease Diagnostics

ORIGINAL ARTICLE

Genome Sequencing for Diagnosing Rare Diseases

M.H. Wojcik, G. Lemire, E. Berger, M.S. Zaki, M. Wissmann, W. Win, S.M. White, B. Weisburd, D. Wieczorek, L.B. Waddell, J.M. Verboon, G.E. VanNoy, A. Töpf, T.Y. Tan, S. Syrbe, V. Strehlow, V. Straub, S.L. Stenton, H. Snow, M. Singer-Berk, J. Silver, S. Shril, E.G. Seaby, R. Schneider, V.G. Sankaran, A. Sanchis-Juan, K.A. Russell, K. Reinson, G. Ravenscroft, M. Radtke, D. Popp, T. Polster, K. Platzer, E.A. Pierce, E.M. Place, S. Pajusalu, L. Pais, K. Õunap, I. Osei-Owusu, H. Opperman, V. Okur, K.T. Oja, M. O'Leary, E. O'Heir, C.F. Morel, A. Merkenschlager, R.G. Marchant, B.E. Mangilog, J.A. Madden, D. MacArthur, A. Lovgren, J.P. Lerner-Ellis, J. Lin, N. Laing, F. Hildebrandt, J. Hentschel, E. Groopman, J. Goodrich, J.G. Gleeson, R. Ghaoui, C.A. Genetti, J. Gburek-Augustat, H.T. Gazda, V.S. Ganesh, M. Ganapathi, L. Gallacher, J.M. Fu, E. Evangelista, E. England, S. Donkervoort, S. DiTroia, S.T. Cooper, W.K. Chung, J. Christodoulou, K.R. Chao, L.D. Cato, K.M. Bujakowska, S.J. Bryen, H. Brand, C.G. Bönnemann, A.H. Beggs, S.M. Baxter, T. Bartolomaeus, P.B. Agrawal, M. Talkowski, C. Austin-Tse, R. Abou Jamra, H.L. Rehm, and A. O'Donnell-Luria

NEJM, 2024

GENOME SEQUENCING: Yields a diagnosis for a rare genetic disease in ~30-50% of cases (and this % will increase over time!).

Undiagnosed Diseases

Rapid Genome Sequencing of Sick Newborns

CONRAD BORNMAN/GALLO/GETTY

The genomes of ill newborns can be sequenced in less than 24 hours to give clinicians a rapid diagnosis.

GENOMICS

Fast sequencing saves newborns

Rapid analysis of infant genomes is aiding diagnosis and treatment of inexplicably ill babies.

Nature, 2014

Rapid Genome Sequencing of Sick Newborns

npj | genomic medicine

Review article

Published in partnership with CEGMR, King Abdulaziz University

https://doi.org/10.1038/s41525-024-00404-0

Rapid genomic sequencing for genetic disease diagnosis and therapy in intensive care units: a review

Check for updates

Stephen F. Kingsmore, Russell Nofsinger & Kasia Ellsworth

NPJ Genome Med, 2024

"In 44 studies of children in ICUs with diseases of unknown etiology, 37% received a genetic diagnosis, 26% had consequent changes in management, and net healthcare costs were reduced by $14,265 per child tested…

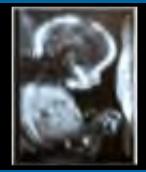
In five years, there is the potential for infant and childhood mortality in the US and UK to have been reduced by several percent through use of [rapid genome sequencing] as a first-tier, standard of care test for children in ICUs with diseases of uncertain etiology."

Genomic Medicine Implementation

Cancer Genomics

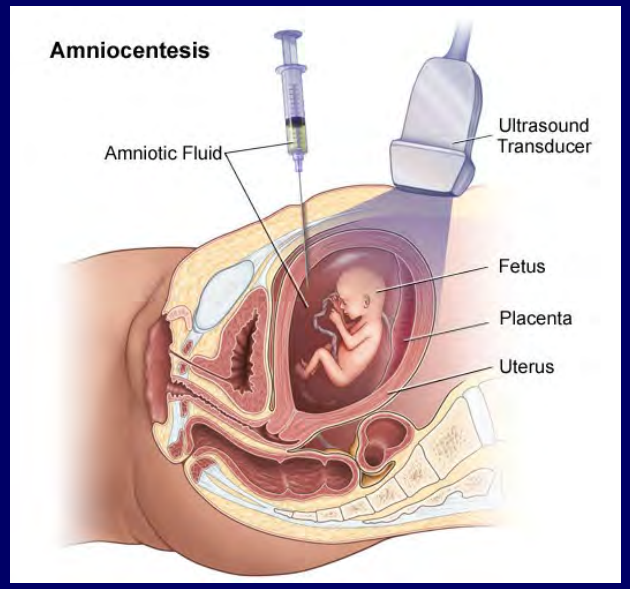
Rare Genetic Disease Diagnostics

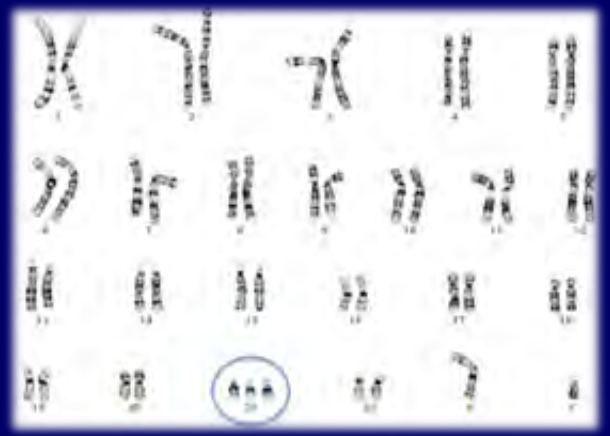
Noninvasive Prenatal Genomic Testing



Noninvasive Prenatal Testing (NIPT)

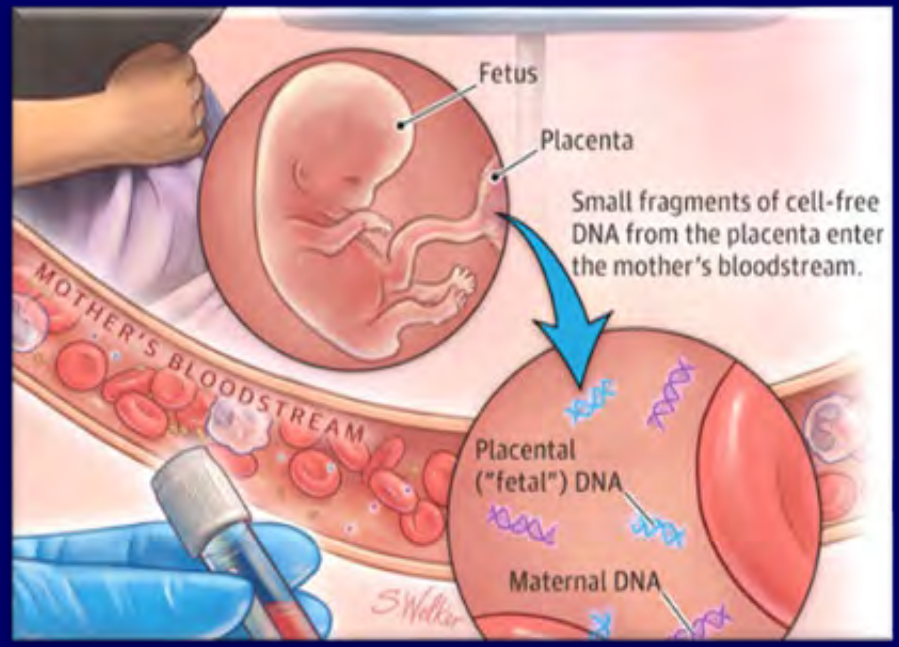
Then (before Genomics)





"Aneuploidy"

Now (with Genomics)



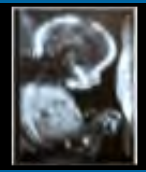
- Both mother and fetus release cell-free DNA from dying cells into the blood
- As an alternative to an invasive procedure, sequencing of cell-free DNA in maternal plasma now used to <u>screen</u> for aneuploidy
- #1 genomic medicine test worldwide

Genomic Medicine Implementation

- Cancer Genomics

- Rare Genetic Disease Diagnostics

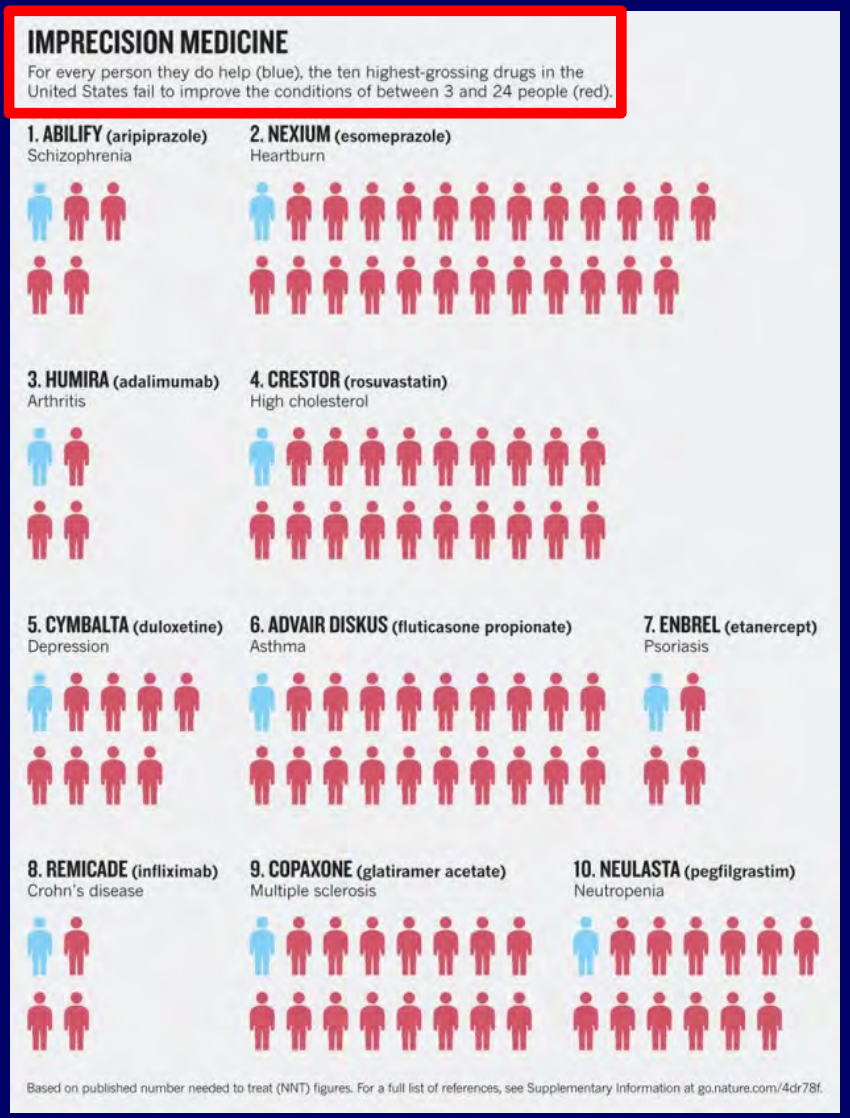
- Noninvasive Prenatal Genomic Testing



- Pharmacogenomic Testing

People Respond Differently to Medications

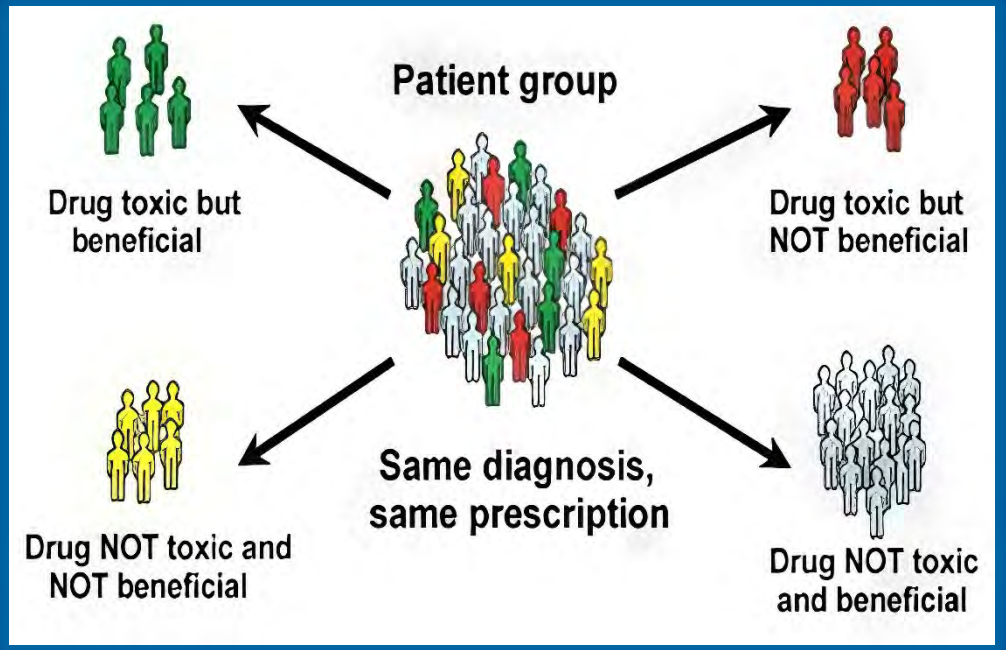
Prescribing Medications is Imprecise



Nature (2015)

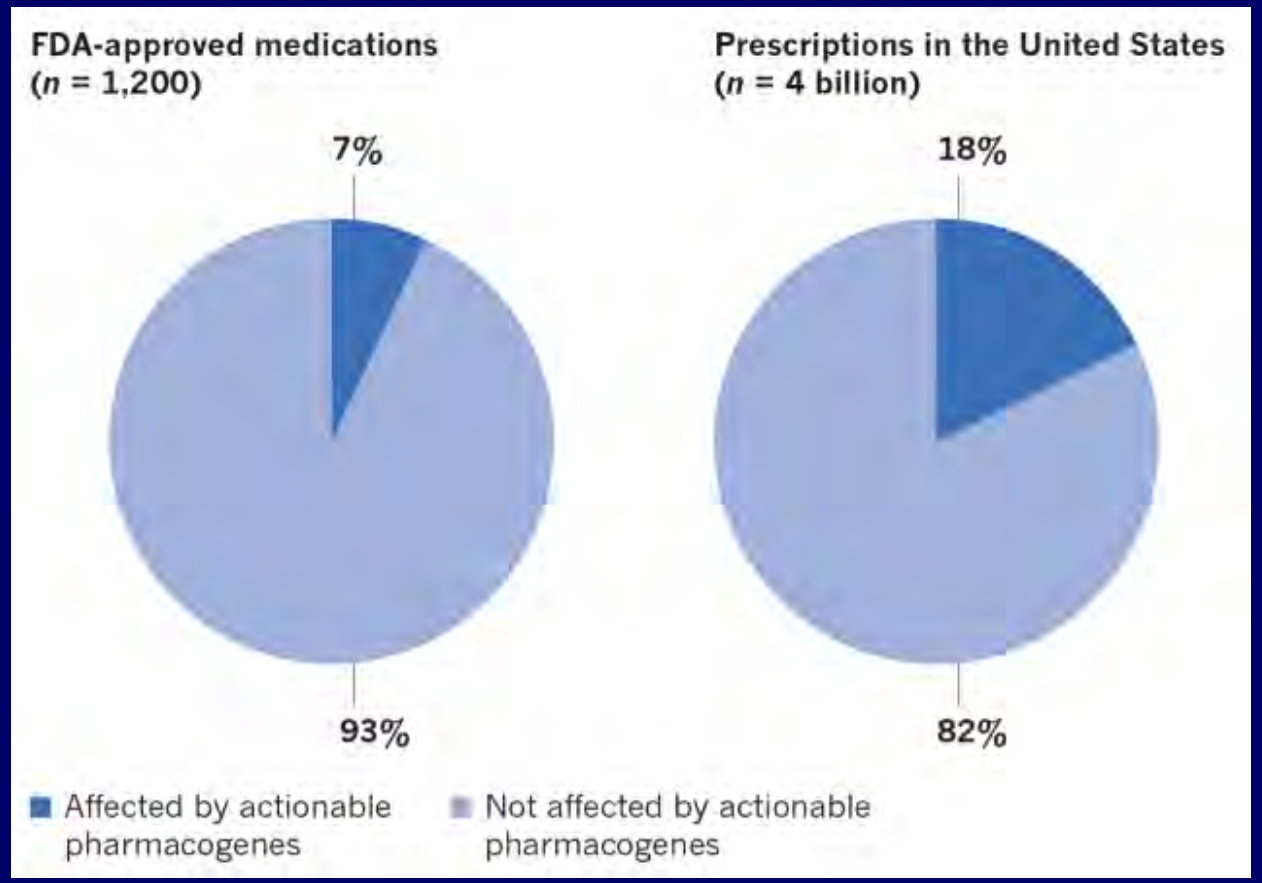
Nature (2016)

Pharmacogenomics: Basic Rationale



- 'One size does <u>not</u> fit all'
- Stratify patients based on detected genomic variants
- Use genomics to 'get the right drug to the right patient at the right dose'

Pharmacogenomics: Getting Real



Drug:	Used to Treat:
Abacavir	HIV
Allopurinol	Gout
Azathioprine	Autoimmune disease
Carbamazepine	Seizures
Clopidogrel	Blood clots
Methotrexate	Autoimmune disease
Phenytoin	Seizures

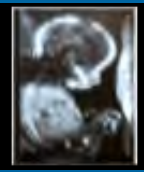
Bottom Line: Pharmacogenomic testing now appropriate for a small subset of prescription medication, but that subset expected to grow in the future. Meanwhile, efforts to increase clinical usage are ongoing.

Genomic Medicine Implementation

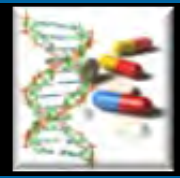
Cancer Genomics

Rare Genetic Disease Diagnostics

Noninvasive Prenatal Genomic Testing



Pharmacogenomic Testing



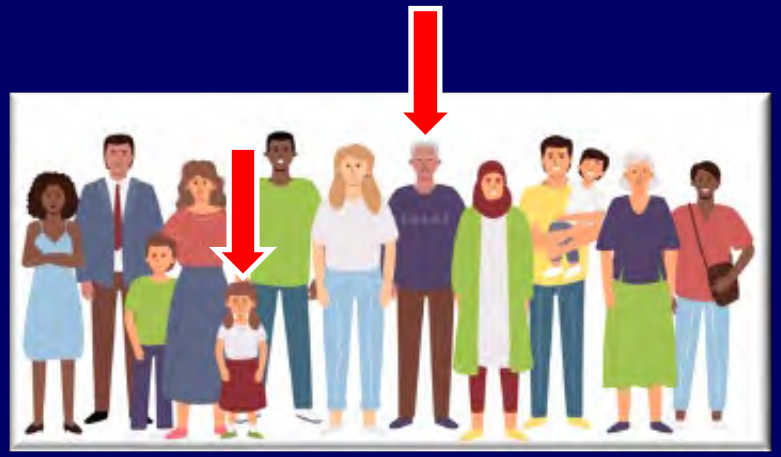
Genomics-based Prevention

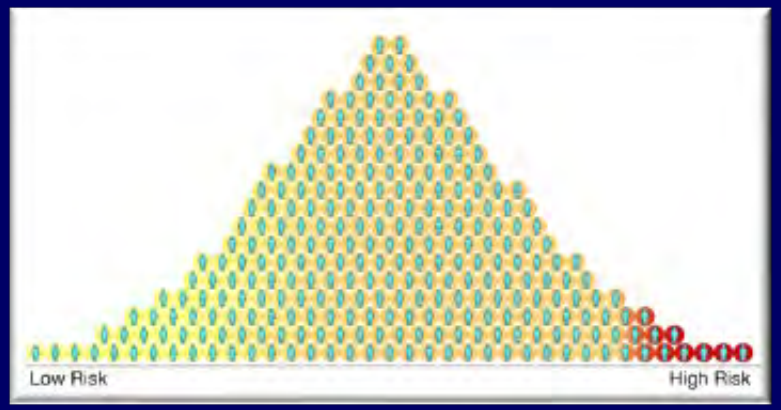
Genomics-based Prevention

PREVENTIVE
GENOMICS CLINIC

Rare Diseases

Common Diseases





Polygenetic Risk Scores

En Route to Genomic Medicine

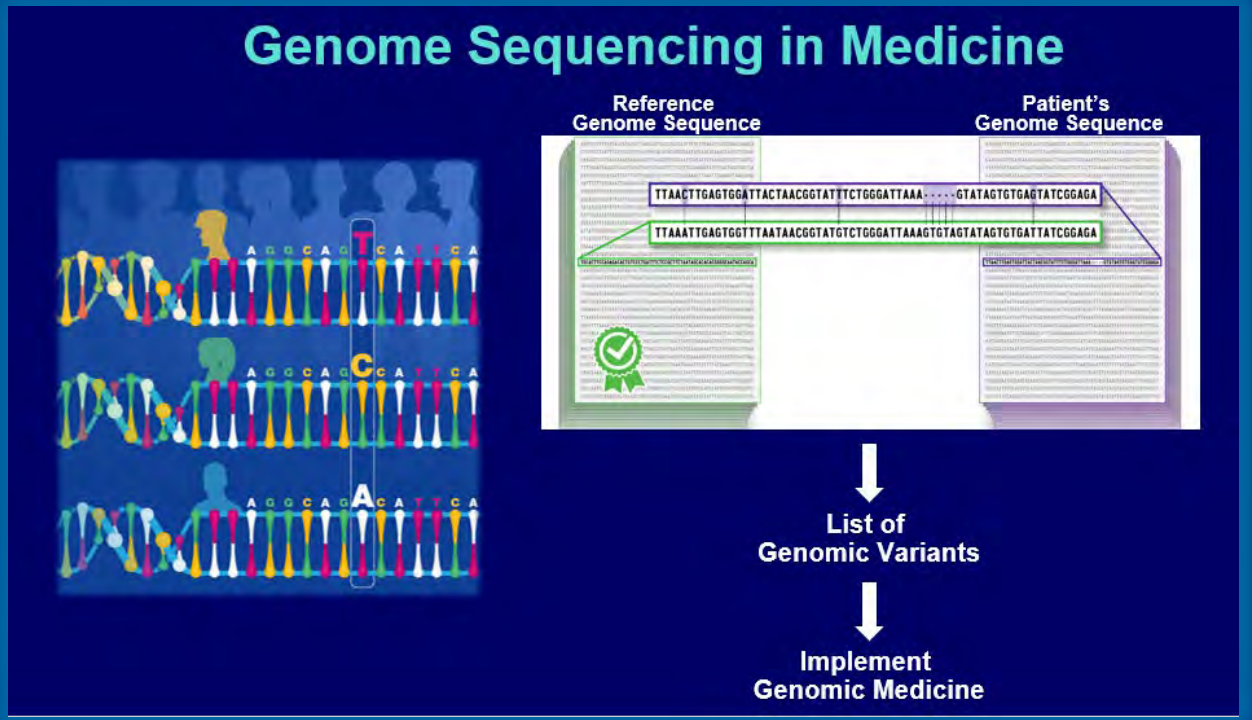
- Human Genome Sequenced for First Time by the Human Genome Project
- Cost of Sequencing a Human Genome Reduced >1 Million-Fold
- Millions of Human Genomes Sequenced
- Profound Advances in Understanding How the Human Genome Functions
- Significant Advances in Unraveling the Genomic Bases of Human Disease
- Vivid Examples of Genomic Medicine Now Emerging

Research
It's A Journey


Challenges
ahead

Challenges of Analyzing a Patient's Genome Sequence

An Intentional Oversimplification



Analyzing a Person's Genome Sequence

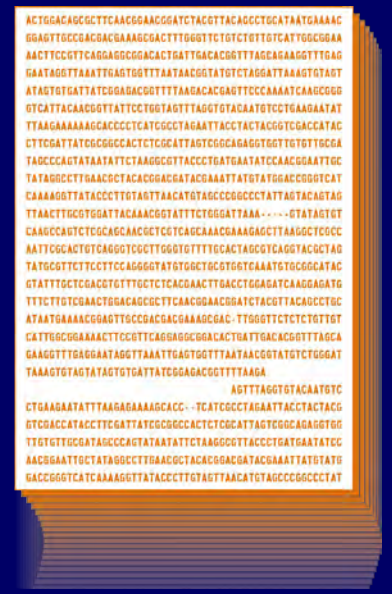
1. **Detecting all genomic variants in the generated genome sequence**
   - **Required: Reference Genome Sequence (or Pangenome Reference)**
   - **Examples: GRCh38.p13 (Build 38) or T2T-CHM13**

Reference vs. Routine Genome Sequences

Reference genome sequence

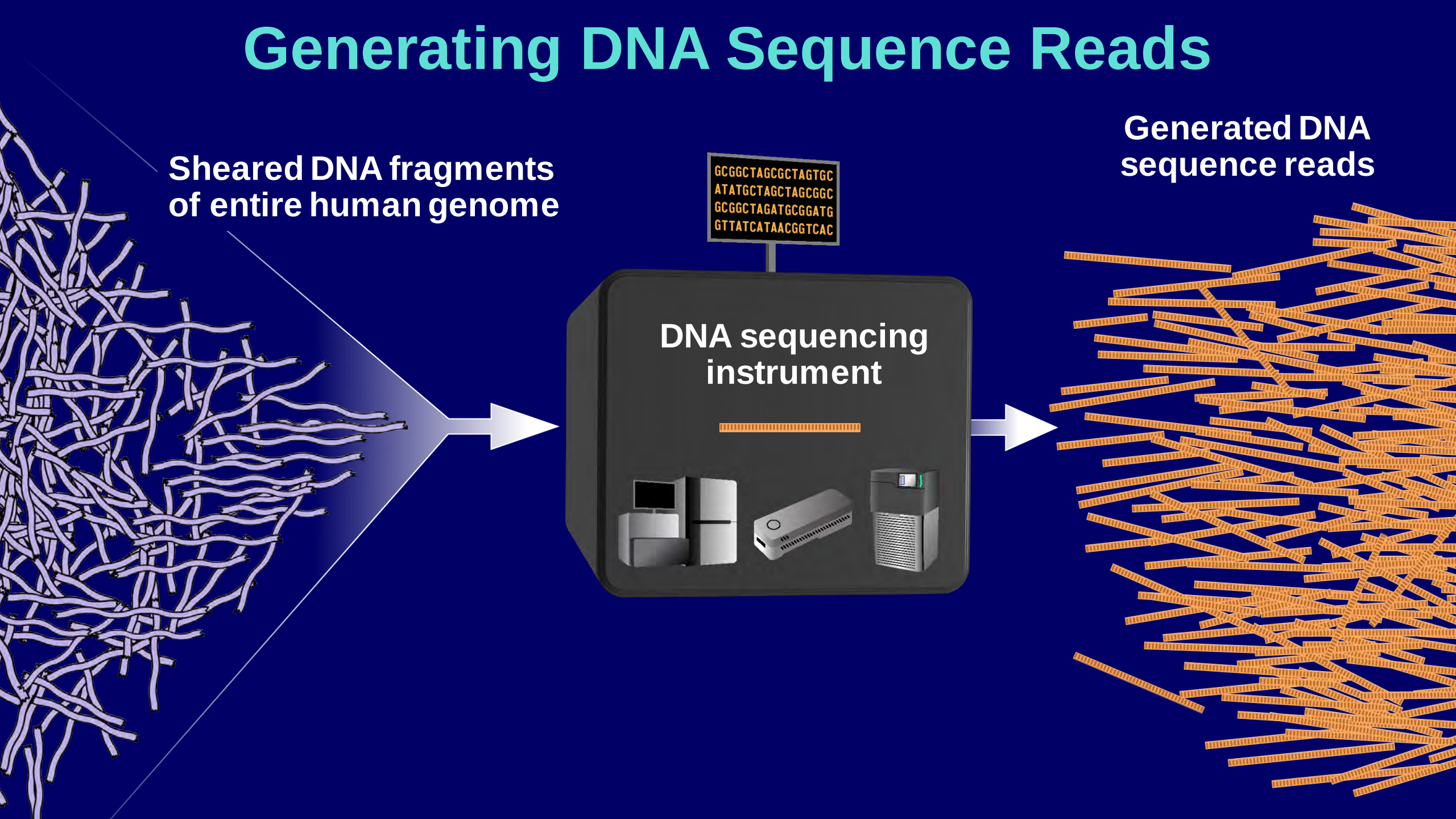
- Purpose: ‘High-quality’ representation
- Meticulously generated
- Multiple DNA-sequencing technologies
- No (or little) missing sequences
- Cost: ~$10,000

Routine genome sequence

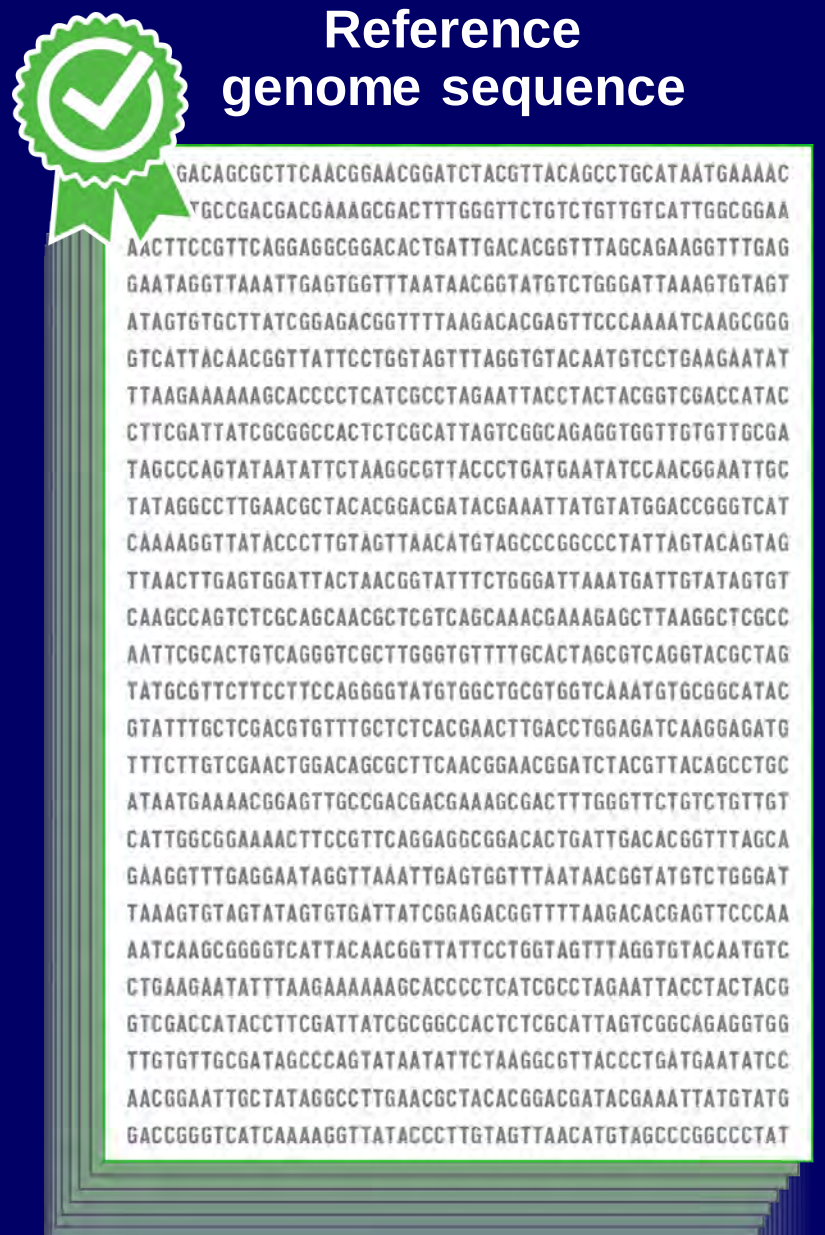


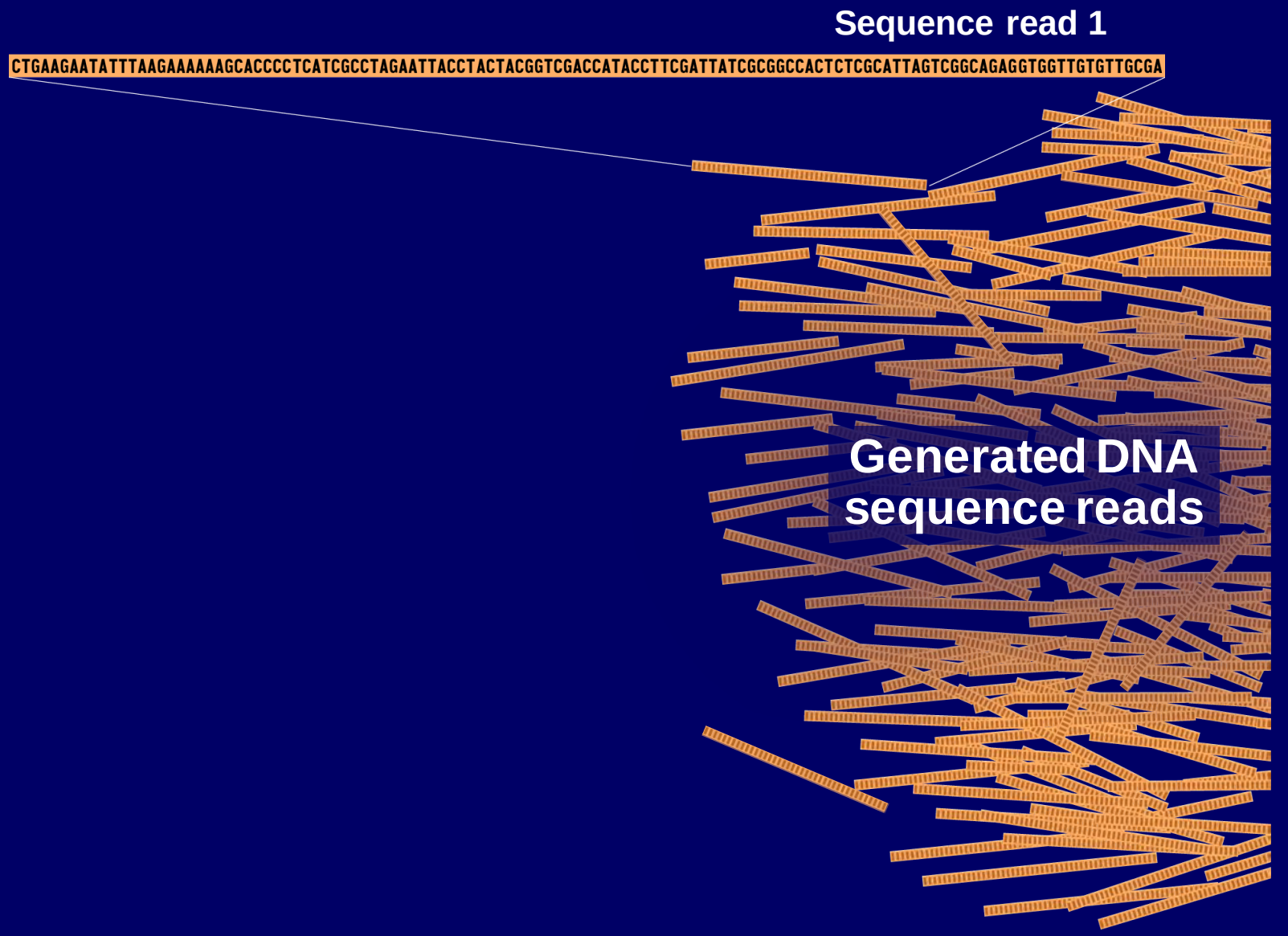
- Purpose: Identify genomic variants
- High-throughput generated
- Single DNA-sequencing technology
- Always has ‘missing’ sequences
- Cost: <$1,000

Generating DNA Sequence Reads



Reference-Assisted Genome Sequence Generation





Reference-Assisted Genome Sequence Generation

Routine genome sequence

Sequence read 1

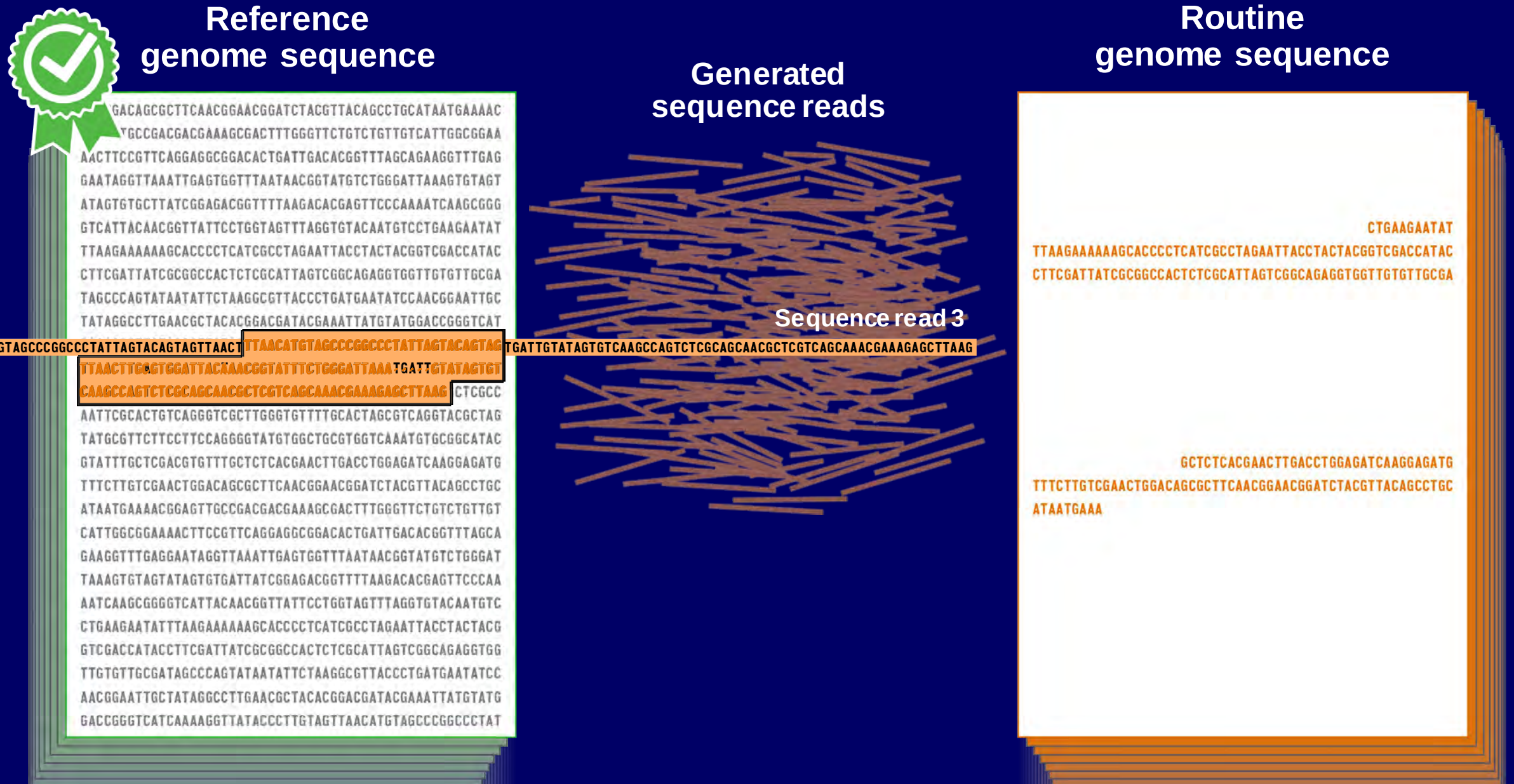
CTGAAGAATATTTAAGAAAAAAGCACCCCTCATCGCCTAGAATTACCTACTACGGTCGACCATACCTTCGATTATCGCGGCCACTCTCGCATTAGTCGGCAGAGGTGGTTGTGTTGCGA

Sequence read 1

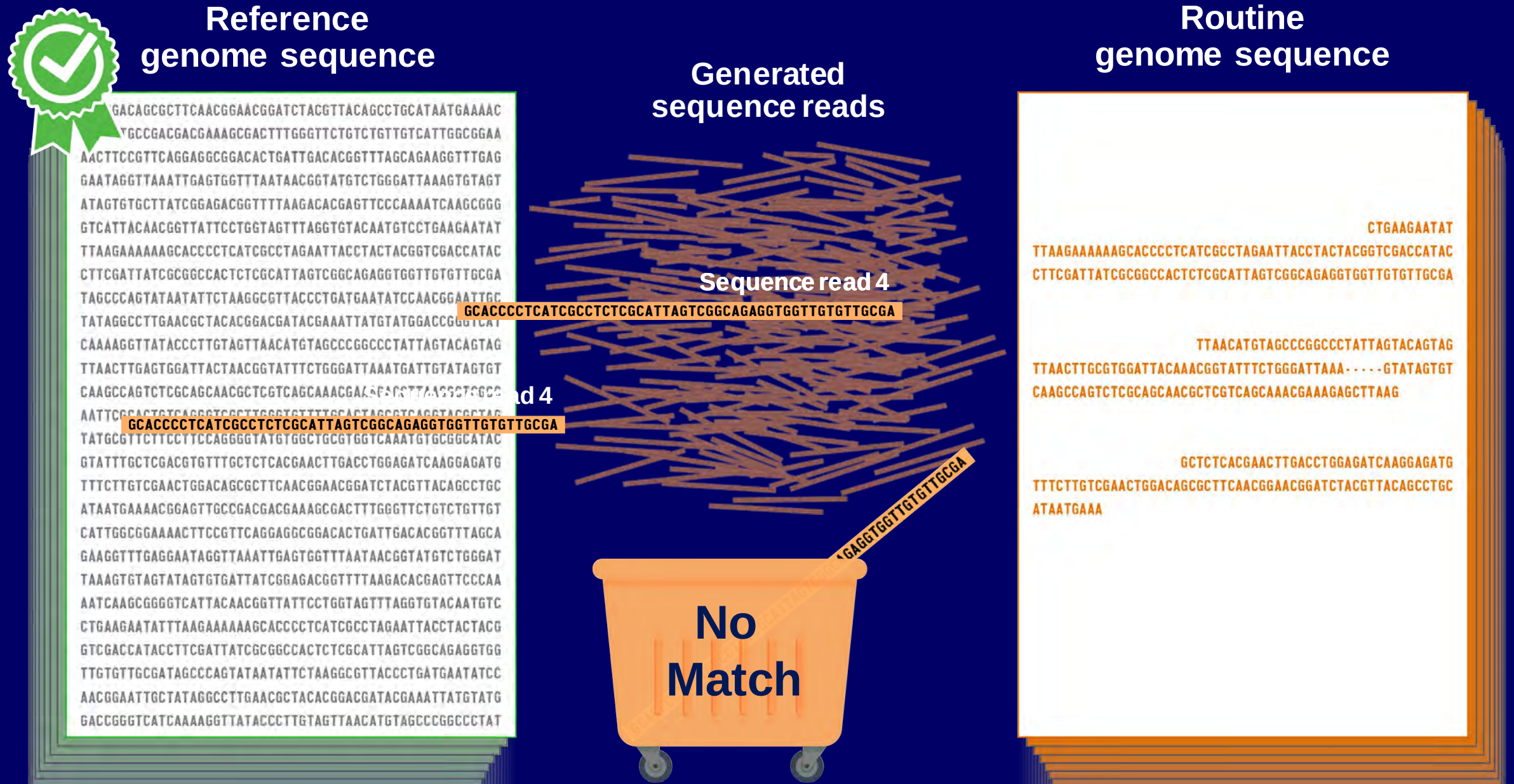
CTGAAGAATATTTAAGAAAAAAGCACCCCTCATCGCCTAGAATTACCTACTACGGTCGACCATACCTTCGATTATCGCGGCCACTCTCGCATTAGTCGGCAGAGGTGGTTGTGTTGCGA

Reference-Assisted Genome Sequence Generation

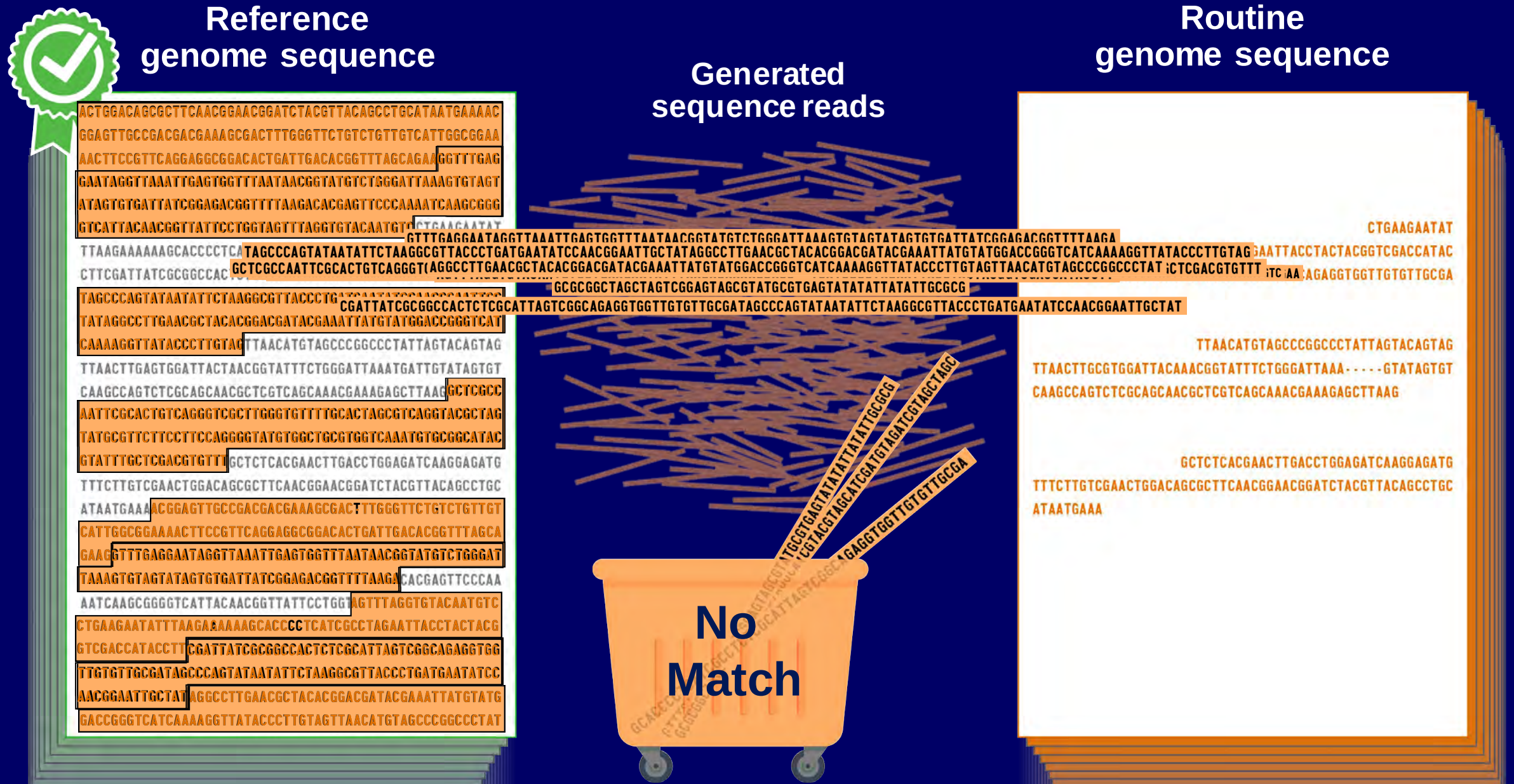
Reference-Assisted Genome Sequence Generation



Reference-Assisted Genome Sequence Generation

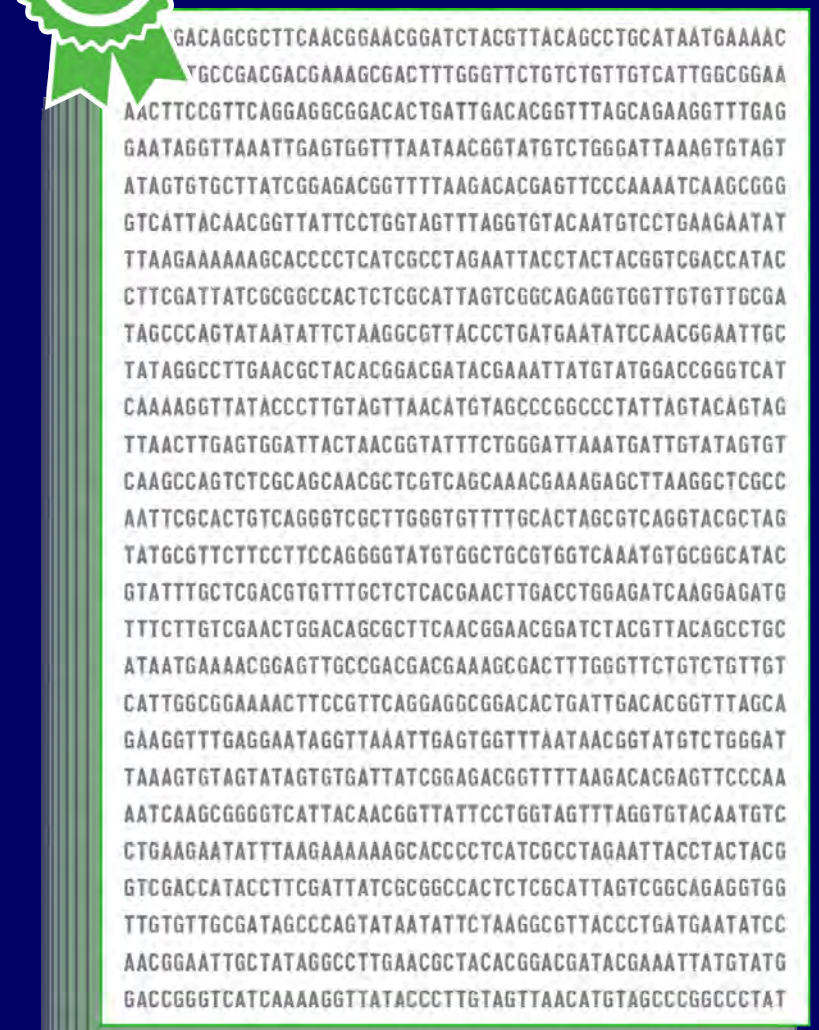


Reference-Assisted Genome Sequence Generation



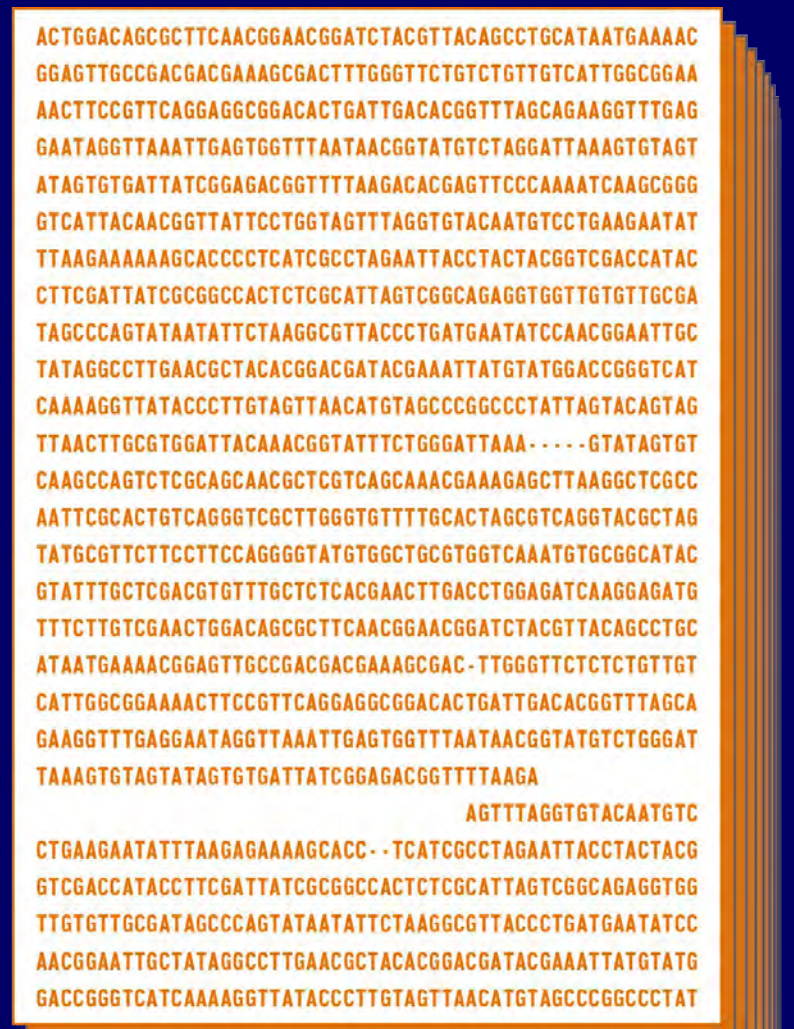
Reference-Assisted Genome Sequence Generation

Reference genome sequence



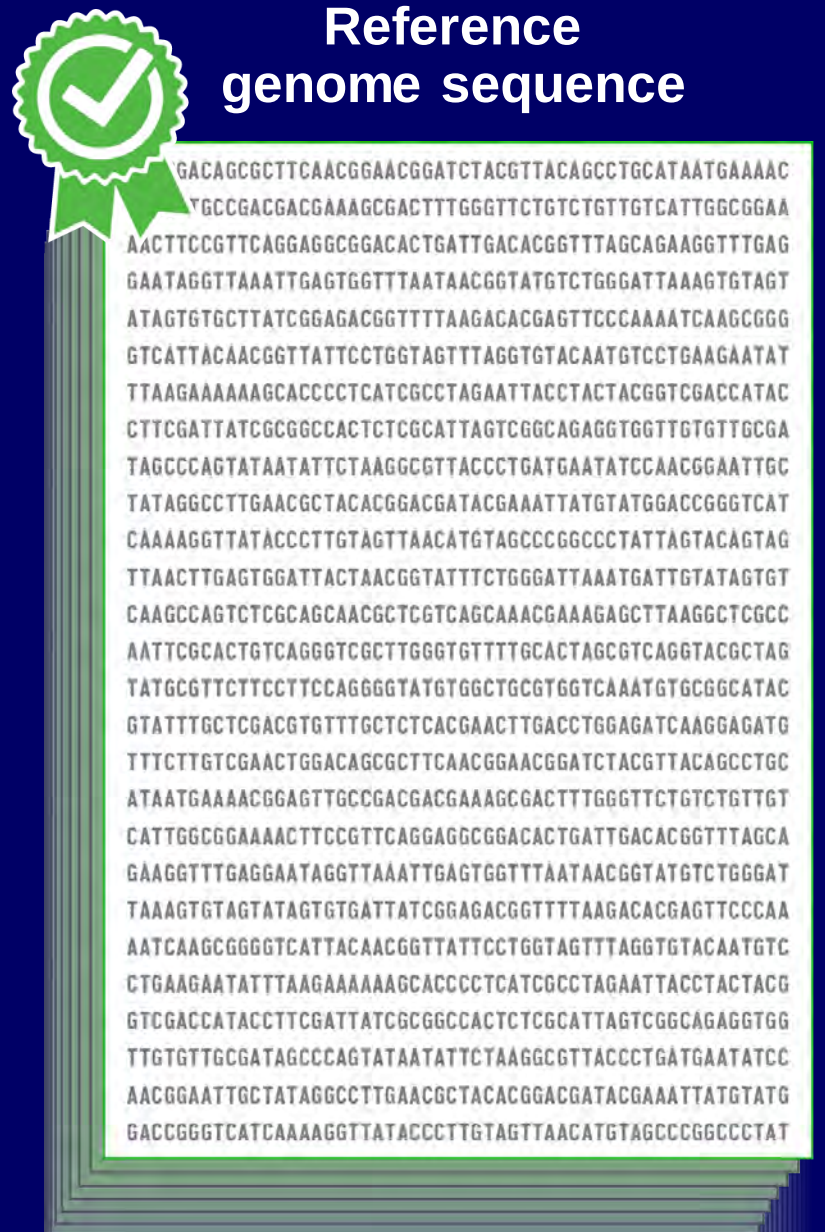
Each line of a routine genome sequence is read and localized many times to ensure its accuracy – often more than 30 times!

Routine genome sequence



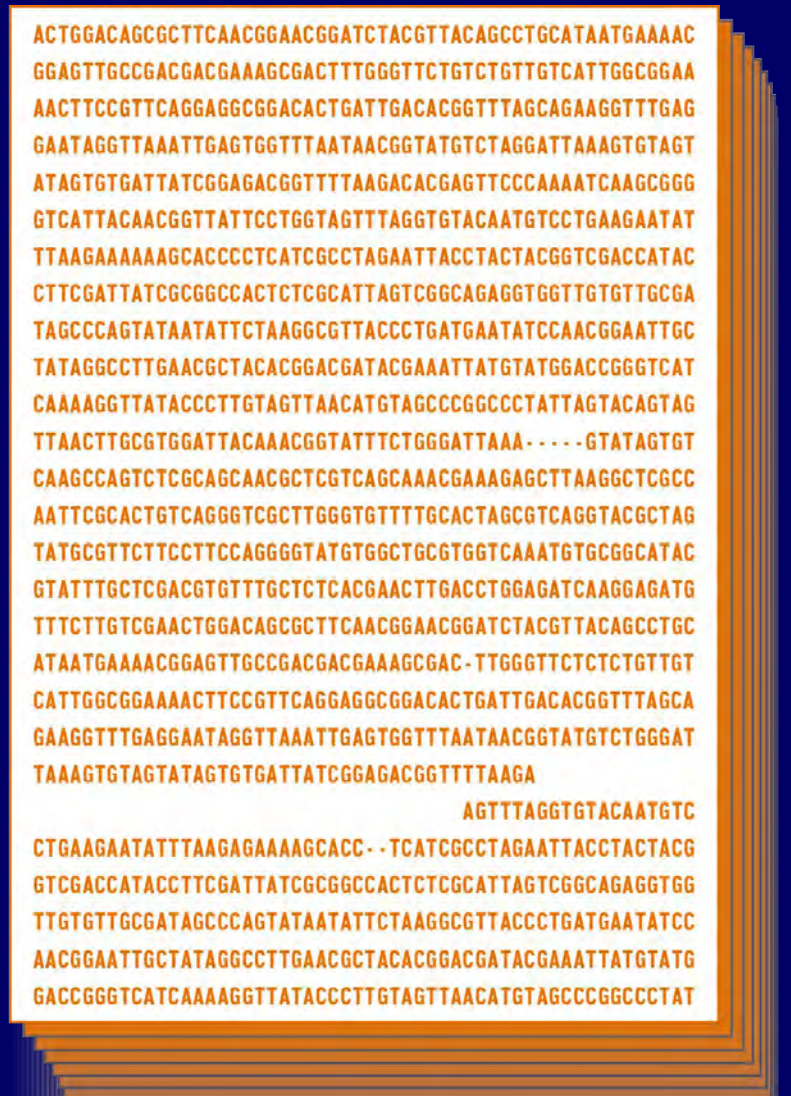
Comparing the Two Genome Sequences

Reference genome sequence



Compare sequences to find differences

Routine genome sequence



Detecting Human Genomic Variants

Reference genome sequence

GACAGCGCTTCAACGGAACGGATCTACGTTACAGCCTGCATAATGAAAAC
TGCCGACGACGAAAGCGACTTTGGGTTCTGTCTGTTGTCATTGGCGGAA
AACTTCCGTTCAGGAGGCGGACACTGATTGACACGGTTTAGCAGAAGGTTTGAG
GAATAGGTTAAATTGAGTGGTTTAATAACGGTATGTCTGGGATTAAAGTGTAGT
ATAGTGTGCTTATCGGAGACGGTTTTAAGACACGAGTTCCCAAAATCAAGCGGG
GTCATTACAACGGTTATTCCTGGTAGTTTAGGTGTACAATGTCCTGAAGAATAT
TTAAGAAAAAAGCACCCCTCATCGCCTAGAATTACCTACTACGGTCGACCATAC
CTTCGATTATCGCGGCCACTCTCGCATTAGTCGGCAGAGGTGGTTGTGTTGCGA
TAGCCCAGTATAATATTCTAAGGCGTTACCCTGATGAATATCCAACGGAATTGC
TATAGGCCTTGAACGCTACACGGACGATACGAAATTATGTATGGACCGGGTCAT
CAAAAGGTTATACCCTTGTAGTTAACATGTAGCCCGGCCCTATTAGTACAGTAG
TTAACTTGAGTGGATTACTAACGGTATTTCTGGGATTAAATGATTGTATAGTGT
CAAGCCAGTCTCGCAGCAACGCTCGTCAGCAAACGAAAGAGCTTAAGGCTCGCC
AATTCGCACTGTCAGGGTCGCTTGGGTGTTTTGCACTAGCGTCAGGTACGCTAG
TATGCGTTCTTCCTTCCAGGGGTATGTGGCTGCGTGGTCAAATGTGCGGCATAC
GTATTTGCTCGACGTGTTTGCTCTCACGAACTTGACCTGGAGATCAAGGAGATG
TTTCTTGTCGAACTGGACAGCGCTTCAACGGAACGGATCTACGTTACAGCCTGC
ATAATGAAAACGGAGTTGCCGACGACGAAAGCGACTTTGGGTTCTGTCTGTTGT
CATTGGCGGAAAACTTCCGTTCAGGAGGCGGACACTGATTGACACGGTTTAGCA
GAAGGTTTGAGGAATAGGTTAAATTGAGTGGTTTAATAACGGTATGTCTGGGAT
TAAAGTGTAGTATAGTGTGATTATCGGAGACGGTTTTAAGACACGAGTTCCCAA
AATCAAGCGGGGTCATTACAACGGTTATTCCTGGTAGTTTAGGTGTACAATGTC
CTGAAGAATATTTAAGAAAAAAGCACCCTCATCGCCTAGAATTACCTACTACG
GTCGACCATACCTTCGATTATCGCGGCCACTCTCGCATTAGTCGGCAGAGGTGG
TTGTGTTGCGATAGCCCAGTATAATATTCTAAGGCGTTACCCTGATGAATATCC
AACGGAATTGCTATAGGCCTTGAACGCTACACGGACGATACGAAATTATGTATG
GACCGGGTCATCAAAAGGTTATACCCTTGTAGTTAACATGTAGCCCGGCCCTAT

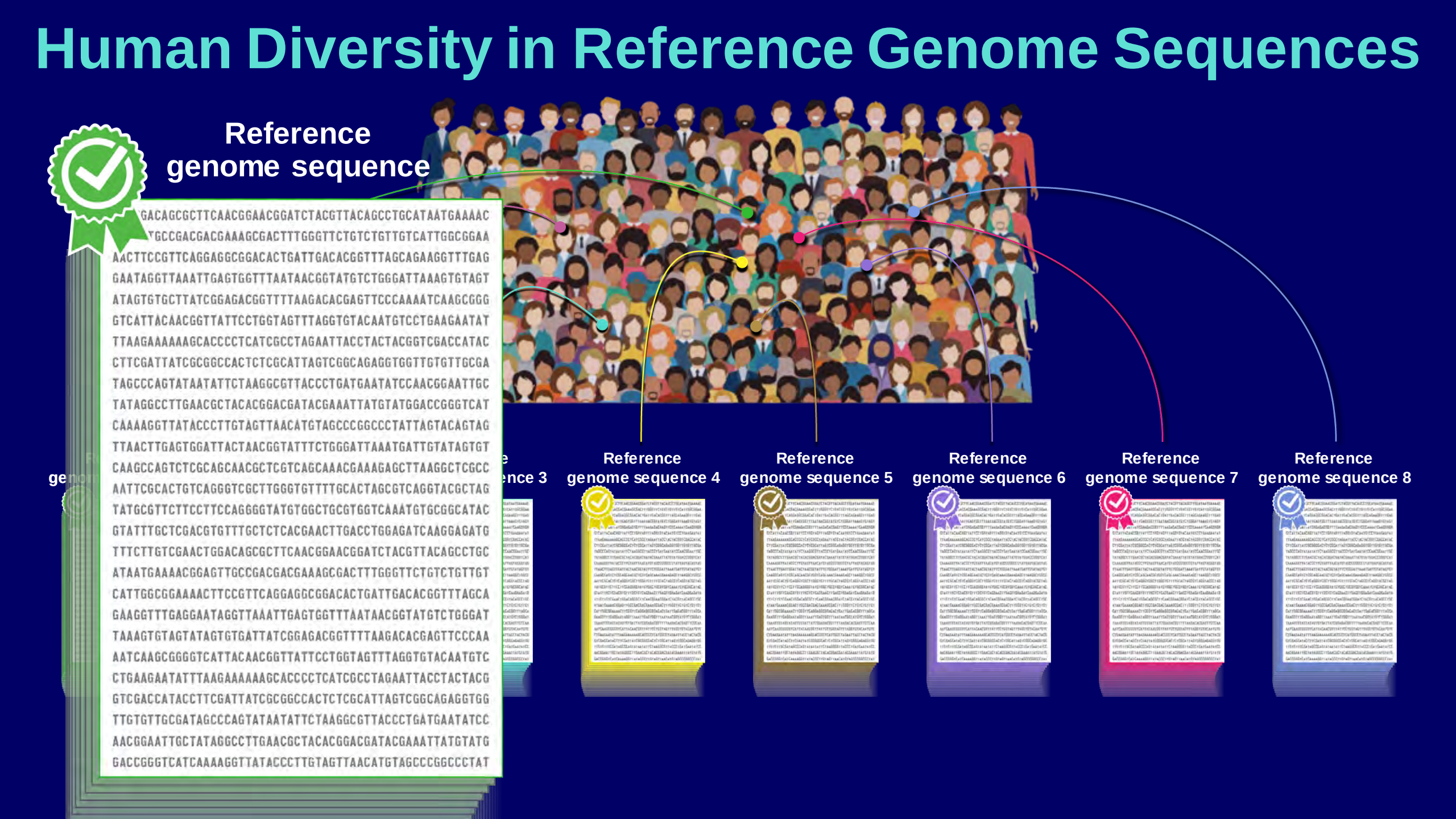
Genomic variants

Yellow vs. *Red*

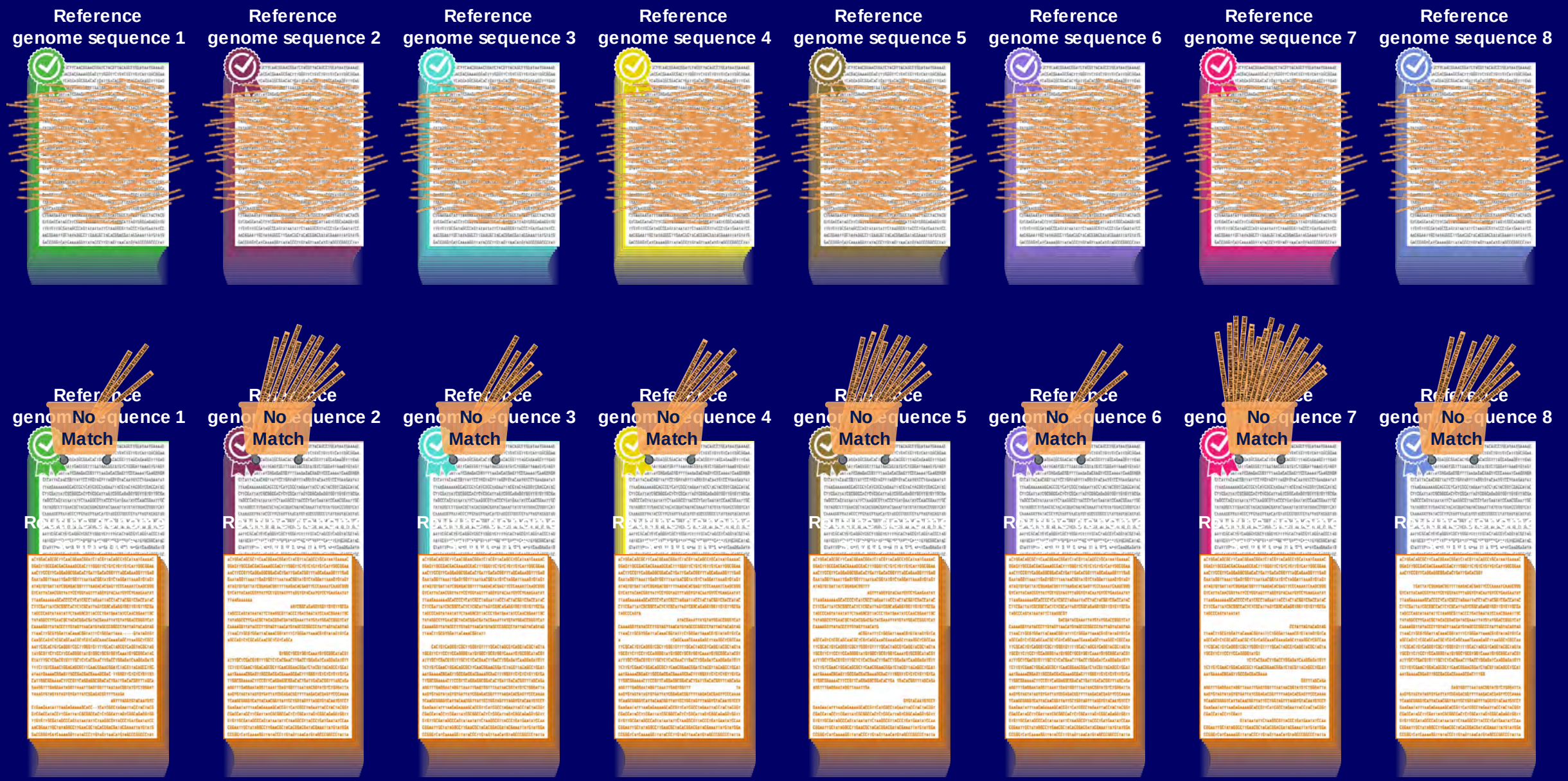
Routine genome sequence

ACTGGACAGCGCTTCAACGGAACGGATCTACGTTACAGCCTGCATAATGAAAAC
GGAGTTGCCGACGACGAAAGCGACTTTGGGTTCTGTCTGTTGTCATTGGCGGAA
AACTTCCGTTCAGGAGGCGGACACTGATTGACACGGTTTAGCAGAAGGTTTGAG
GAATAGGTTAAATTGAGTGGTTTAATAACGGTATGTCTAGGATTAAAGTGTAGT
ATAGTGTGATTATCGGAGACGGTTTTAAGACACGAGTTCCCAAAATCAAGCGGG
GTCATTACAACGGTTATTCCTGGTAGTTTAGGTGTACAATGTCCTGAAGAATAT
CCATCGCCTAGAATTACCTACT
TAGCCCAGTATAATATTCTAAGGCGTTACCCTGATGAATATCCAACGGAATTGC
TATAGGCCTTGAACGCTACACGGACGATACGAAATTATGTATGGACCGGGTCAT
CAAAAGGTTATACCCTTGTAGTTAACATGTAGCCCGGCCCTATTAGTACAGTAG
TTAACTTGCGTGGATTACAAACGGTATTTCTGGGATTAAA----GTATAGTGT
CAAGCCAGTCTCGCAGCAACGCTCGTCAGCAAACGAAAGAGCTTAAGGCTCGCC
AATTCGCACTGTCAGGGTCGCTTGGGTGTTTTGCACTAGCGTCAGGTACGCTAG
TATGCGTTCTTCCTTCCAGGGGTATGTGGCTGCGTGGTCAAAGGGTATGTGGCT
GCGTGGTCAAATGTGCGGCATACGTATTTGCTCGACGTGTTTGCTCTCACGAAC
TTGACCTGGAGATCAAGGAGATGTTTCTTGTCGAACTGGACAGCGCTTCAACGG
AACGGATCTACGTTACAGCCTGCATAATGAAAACGGAGTTGCCGACGACGAAAG
CGAC-TTGGGTTCTCTGTTGTCATTGGCGGAAAACTTCCGTTCAGGAGGCGG
ACACTGATTGACACGGTTTAGCAGAAGGTTTGAGGAATAGGTTAAATTGAGTGG
TTTAATAACGGTATGTCTGGGATTAAAGTGTAGTATAGTGTGATTATCGGAGAC
GGTTTTAAGA
AGTTTAGGTGTACAATGTCCTGAAGAATATTTAAGAGAAAAGCACC--TC
ATCGCCTAGAATTACCTACTACGGTCGACCATACCTTCGATTATCGCGGCCACT
CTCGCATTAGTCGGCAGCAGCAGCAGCAGAGGTGGTTGTGTTGCGATAGCCCAG
TATAATATTCTAAGGCGTTACCCTGATGAATATCCAACGGAATTGCTATAGGCC
TTGAACGCTACACGGACGATACGAAATTATGTATGGACCGGGTCATCAAAAGGT

Human Diversity in Reference Genome Sequences



No Single Reference Genome Sequence is Ideal



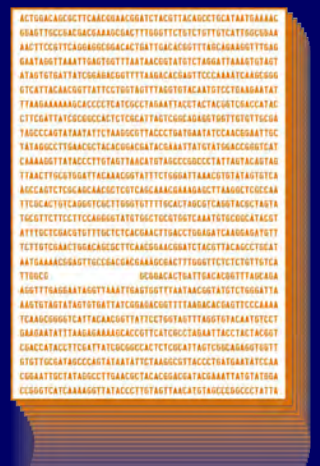
Human Pangenome Reference

- Composite of multiple human reference genome sequences
- Captures the breadth of human genomic variation much better than any one human reference genome sequence
- Enables more accurate and complete detection of genomic variants across diverse human populations

Using a Human Pangenome Reference

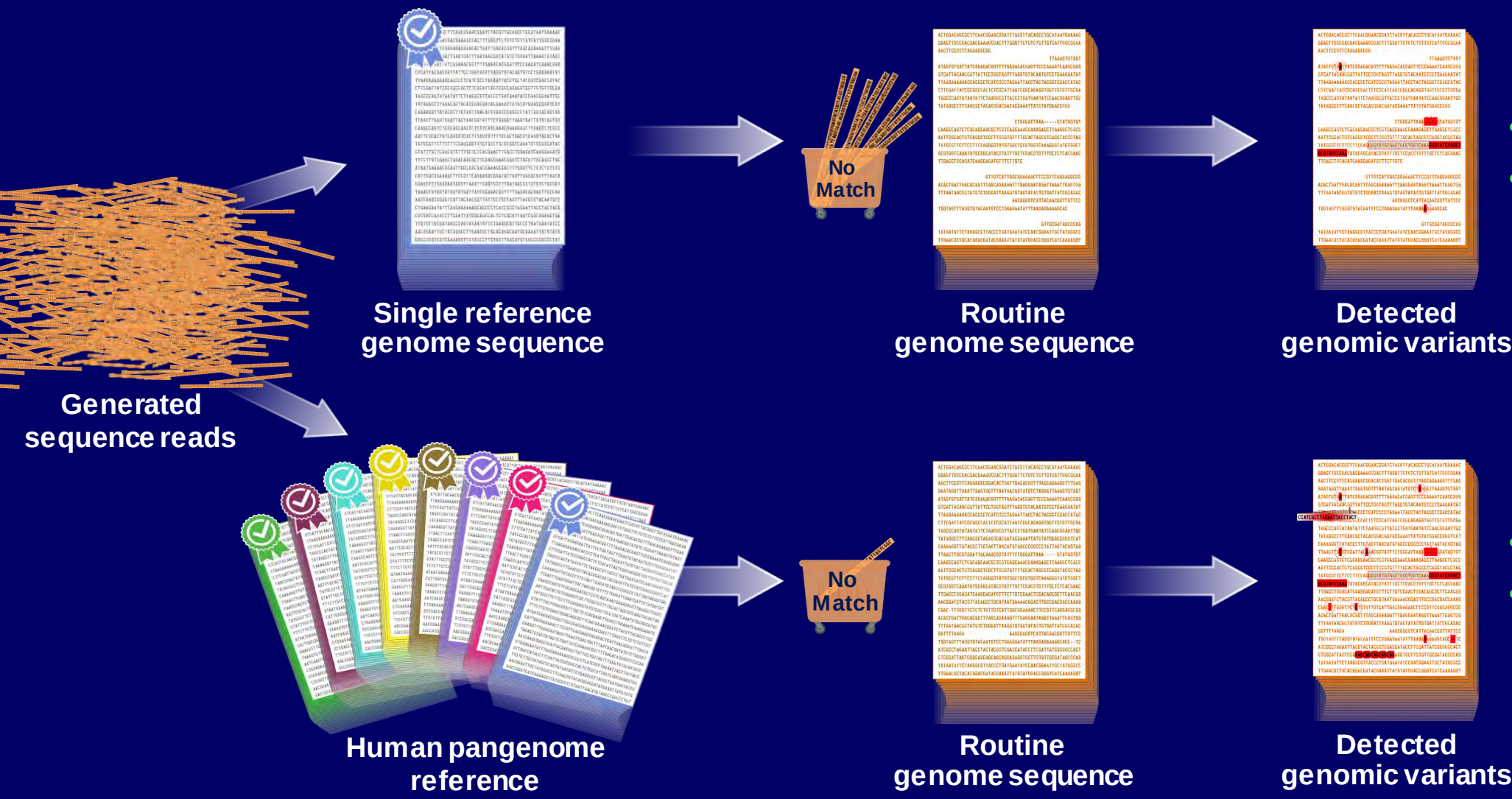
Human pangenome reference

Routine genome sequence

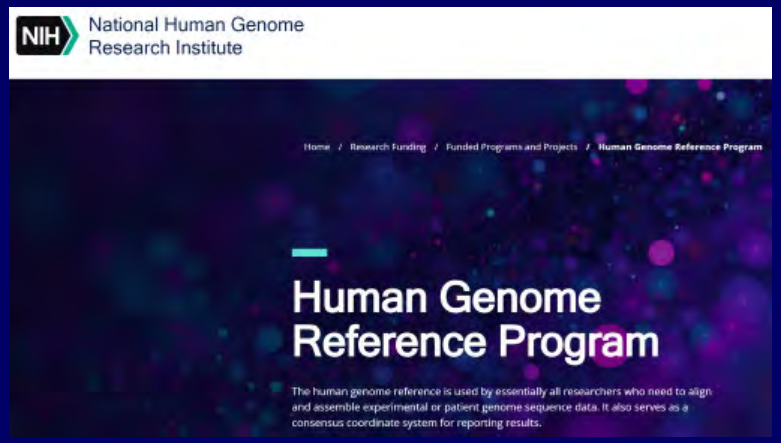


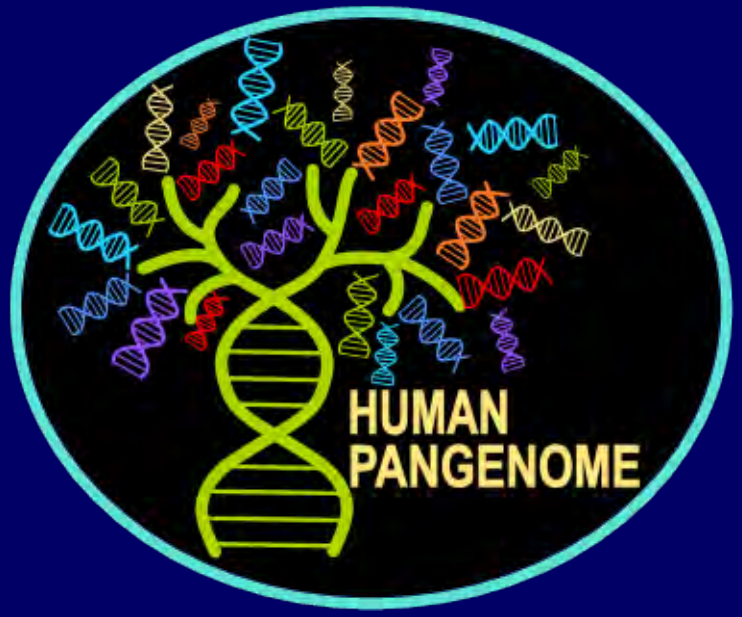
More complete routine genome sequence

Improved Genomic Variant Detection Using a Human Pangenome Reference



NHGRI's Human Genome Reference Program





Article

A draft human pangenome reference

https://doi.org/10.1038/s41586-023-05896-x
Received: 9 July 2022
Accepted: 28 February 2023
Published online: 10 May 2023
Open access
Check for updates

Wen-Wei Liao, Mobin Asri, Jana Ebler, Daniel Doerr, Marina Haukness, Glenn Hickey, Shuangjia Lu, Julian K. Lucas, Jean Monlong, Haley J. Abel, Silvia Buonaiuto, Xian H. Chang, Haoyu Cheng, Justin Chu, Vincenza Colonna, Jordan M. Eizenga, Xiaowen Feng, Christian Fischer, Robert S. Fulton, Shilpa Garg, Cristian Groza, Andrea Guarracino, William T. Harvey, Simon Heumos, Kerstin Howe, Miten Jain, Tsung-Yu Lu, Charles Markello, Fergal J. Martin, Matthew W. Mitchell, Katherine M. Munson, Moses Njagi Mwaniki, Adam M. Novak, Hugh E. Olsen, Trevor Pesout, David Porubsky, Pjotr Prins, Jonas A. Sibbesen, Jouni Sirén, Chad Tomlinson, Flavia Villani, Mitchell R. Vollger, Lucinda L. Antonacci-Fulton, Gunjan Baid, Carl A. Baker, Anastasiya Belyaeva, Konstantinos Billis, Andrew Carroll, Pi-Chuan Chang, Sarah Cody, Daniel E. Cook, Robert M. Cook-Deegan, Omar E. Cornejo, Mark Diekhans, Peter Ebert, Susan Fairley, Olivier Fedrigo, Adam L. Felsenfeld, Giulio Formenti, Adam Frankish, Yan Gao, Nanibaa' A. Garrison, Carlos Garcia Giron, Richard E. Green, Leanne Haggerty, Kendra Hoekzema, Thibaut Hourlier, Hanlee P. Ji, Eimear E. Kenny, Barbara A. Koenig, Alexey Kolesnikov, Jan O. Korbel, Jennifer Kordosky, Sergey Koren, HoJoon Lee, Alexandra P. Lewis, Hugo Magalhães, Santiago Marco-Sola, Pierre Marijon, Ann McCartney, Jennifer McDaniel, Jacquelyn Mountcastle, Maria Nattestad, Sergey Nurk, Nathan D. Olson, Alice B. Popejoy, Daniela Puiu, Mikko Rautiainen, Allison A. Regier, Arang Rhie, Samuel Sacco, Ashley D. Sanders, Valerie A. Schneider, Baergen I. Schultz, Kishwar Shafin, Michael W. Smith, Heidi J. Sofia, Ahmad N. Abou Tayoun, Françoise Thibaud-Nissen, Francesca Floriana Tricomi, Justin Wagner, Brian Walenz, Jonathan M. D. Wood, Aleksey V. Zimin, Guillaume Bourque, Mark J. P. Chaisson, Paul Flicek, Adam M. Phillippy, Justin M. Zook, Evan E. Eichler, David Haussler, Ting Wang, Erich D. Jarvis, Karen H. Miga, Erik Garrison, Tobias Marschall, Ira M. Hall, Heng Li & Benedict Paten

Analyzing a Person's Genome Sequence

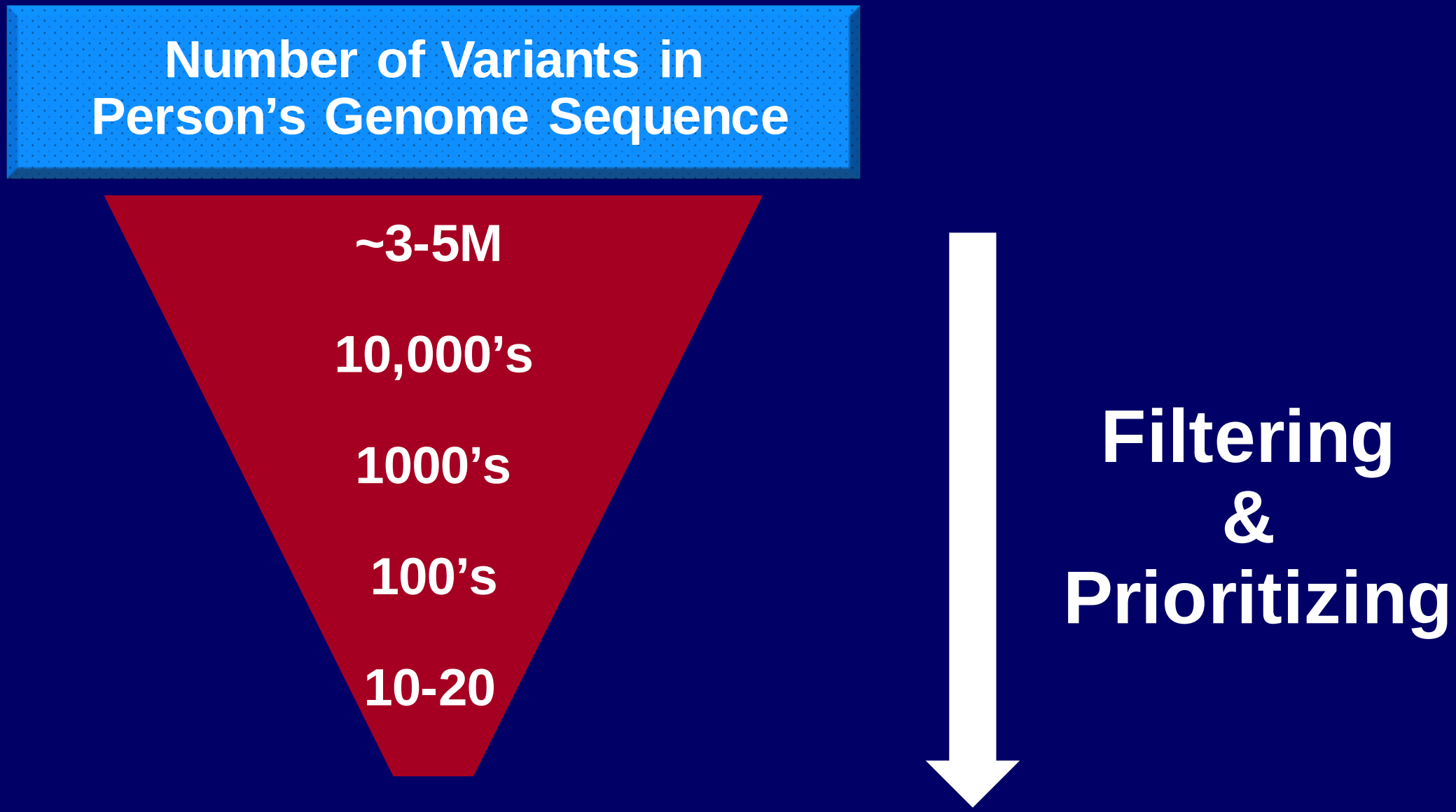
1. **Detecting all genomic variants in the generated genome sequence**
   - Required: Reference Genome Sequence (or Pangenome Reference)
   - Examples: GRCh38.p13 (Build 38) or T2T-CHM13

2. **Filtering & Prioritizing detected genomic variants to identify those most likely to be clinically relevant (e.g., pathogenic variants in the case of rare genetic diseases)**
   - Required: Reference Population Databases (Aggregated Genomic Variants)
   - Example: gnomAD

Frequencies of Genomic Variants: Rare Disease Diagnostics as Prototype

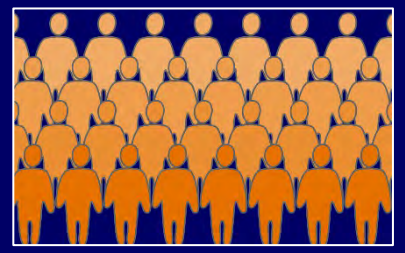
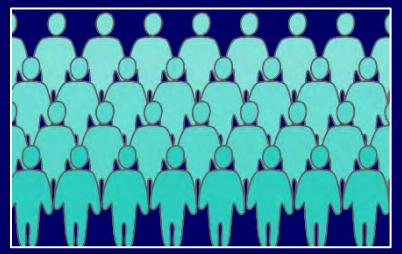
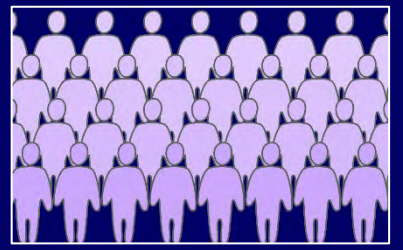
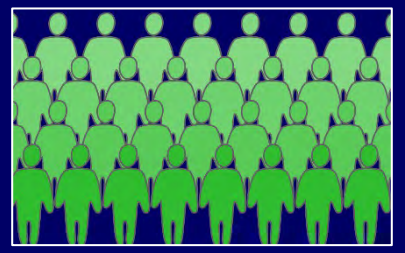
- Vast majority of pathogenic genomic variants are rare
- But being rare does not mean a genomic variant is pathogenic
- However, being common means a genomic variant is unlikely to be pathogenic
- Therefore, genomic variants are initially FILTERED into groups that are common (removed) and rare (prioritized)
- Following filtering, prioritized variants are further assessed for possible pathogenicity

Analyzing a Person's Genomic Variants



Frequencies of Genomic Variants Vary Among Ancestral Populations

Different Ancestral Populations

Variant 1	30%	35%	20%	18%
Variant 2	0.001%	0%	0.001%	0.0001%
Variant 3	0.001%	30%	25%	20%
Variant 4	0.01%	0.0001%	0.001%	30%

Frequency of Each Variant in Each Ancestral Population

Overcoming Inequities in Genomic Diagnoses

Decreasing Incorrect Genomic Variant Classifications

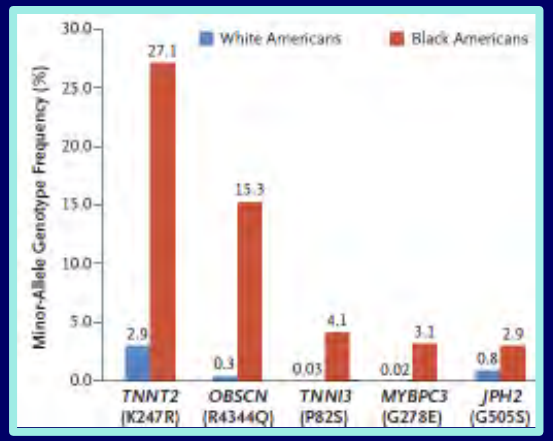
The NEW ENGLAND JOURNAL of MEDICINE

SPECIAL ARTICLE

Genetic Misdiagnoses and the Potential for Health Disparities

Arjun K. Manrai, Ph.D., Birgit H. Funke, Ph.D., Heidi L. Rehm, Ph.D., Morten S. Olesen, Ph.D., Bradley A. Maron, M.D., Peter Szolovits, Ph.D., David M. Margulies, M.D., Joseph Loscalzo, M.D., Ph.D., and Isaac S. Kohane, M.D., Ph.D.

Manrai et al., N Engl J Med (2016)



Misclassification of 5 genomic variants for cardiomyopathy

Increasing Diagnoses via Equitable Study Recruitment and Clinical Testing

Petrovski and Goldstein Genome Biology (2016) 17:157
DOI 10.1186/s13059-016-1016-y

Genome Biology

OPEN LETTER Open Access

Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine

Slavé Petrovski[1,2*] and David B. Goldstein[1*]

Petrovski & Goldstein Genome Biol (2016)

communications biology

ARTICLE

OPEN

The frequency of pathogenic variation in the *All of Us* cohort reveals ancestry-driven disparities

Eric Venner[1✉], Karynne Patterson[2], Divya Kalra[1], Marsha M. Wheeler[2], Yi-Ju Chen[1], Sara E. Kalla[1], Bo Yuan[1], Jason H. Karnes[3,4], Kimberly Walker[1], Joshua D. Smith[2], Sean McGee[2], Aparna Radhakrishnan[2], Andrew Haddad[5], Philip E. Empey[6], Qiaoyan Wang[1], Lee Lichtenstein[7], Diana Toledo[7], Gail Jarvik[8,9], Anjene Musick[10] & Richard A. Gibbs[1] on behalf of the All of Us Research Program Investigators[+]

Venner et al., Commun Biol (2024)

Analyzing a Person's Genome Sequence

1. **Detecting** all genomic variants in the generated genome sequence
   - Required: Reference Genome Sequence (or Pangenome Reference)
   - Examples: GRCh38.p13 (Build 38) or T2T-CHM13

2. **Filtering** & **Prioritizing** detected genomic variants to identify those most likely to be clinically relevant (e.g., pathogenic variants in the case of rare genetic diseases)
   - Required: Reference Population Databases (Aggregated Genomic Variants)
   - Example: gnomAD

3. **Establishing** the clinical relevance of prioritized genomic variants
   - Required: Knowledgebase with Information about Pathogenicity of Genomic Variants
   - Example: ClinGen

Clinical Genome Resource (ClinGen)

Mission: Build and support an authoritative central resource that defines the clinical relevance of genes and variants for use in precision medicine and research

Global Network of Contributors:
>2,700 Experts from 69 Countries

2,682 Gene-Disease Validity Curations

7,130 Variant Pathogenicity Curations

248 Clinical Actionability Reports

Required for Accurate and Equitable Analyses of a Person's Genome Sequences

1. Appropriately matched human genome reference sequence – or a human pangenome reference

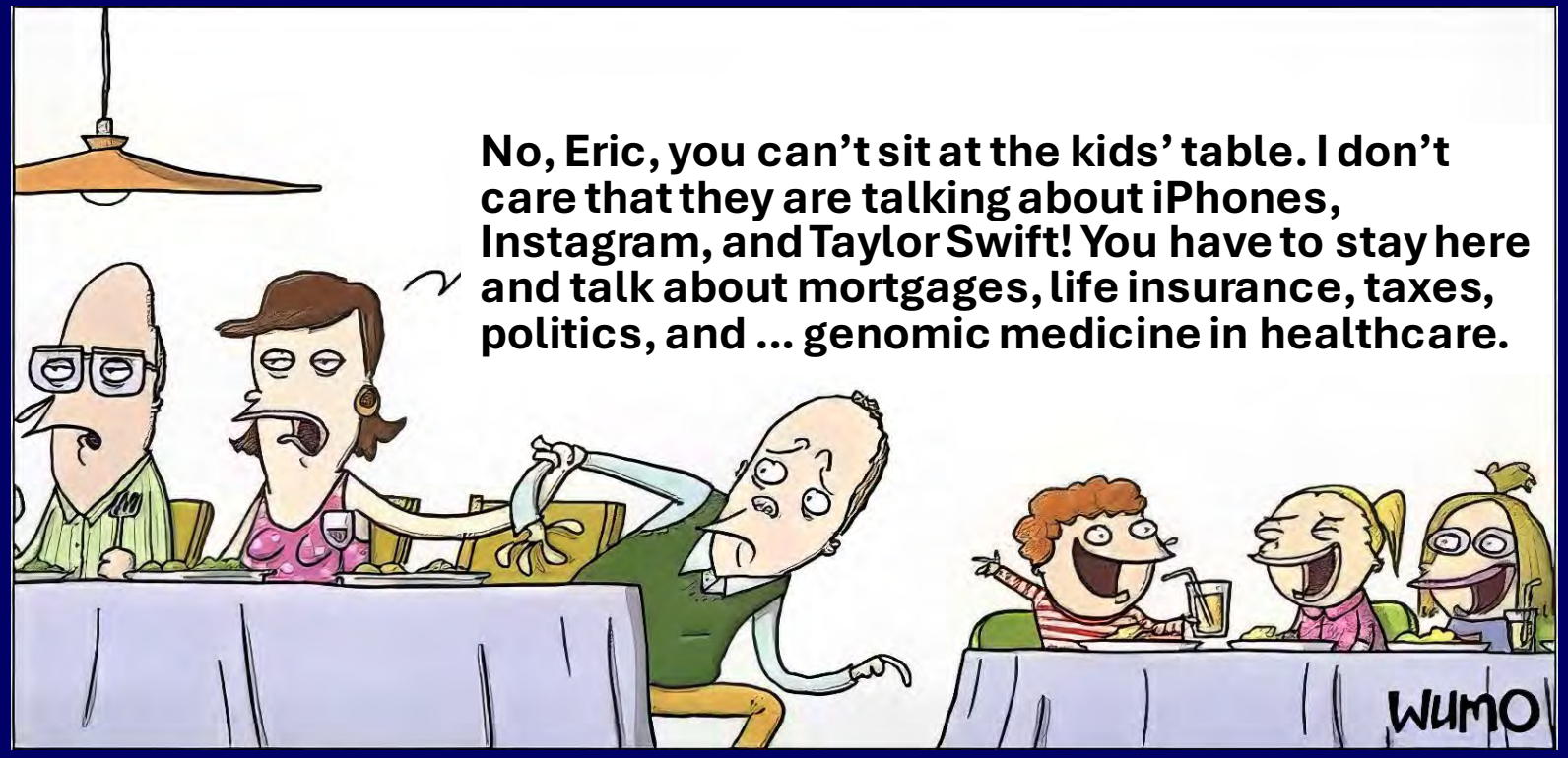
2. Reference population database (with aggregated genomic variant information) for appropriately matched ancestral population(s)

gnomAD
genome aggregation database

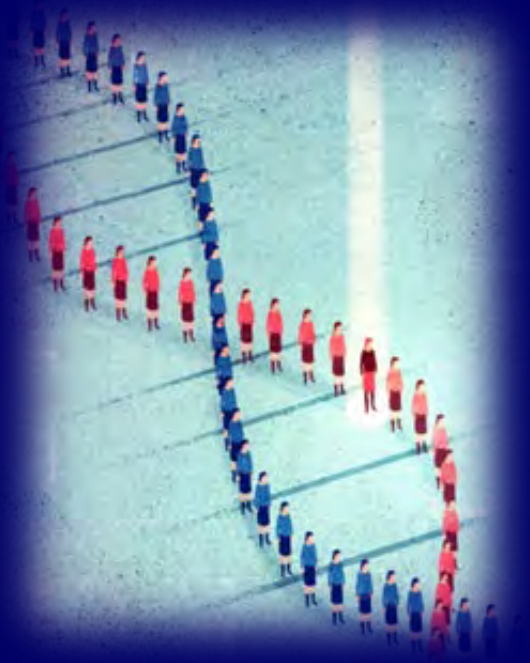
3. Robust knowledgebase of curated information about the likely pathogenicity of genomic variants (developed by expert panels)

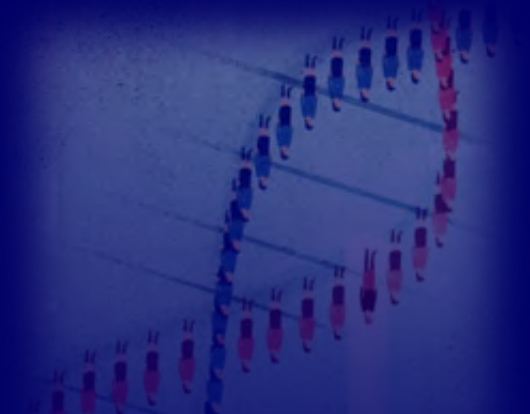
ClinGen
Clinical Genome Resource

Genomics Arrives at the ‘Adult Table’



Genomics and Society





Societal Challenges with Genomic Medicine

The **Forefront** of **Genomics**®

2020 NHGRI Strategic Vision

Perspective

Strategic vision for improving human health at The Forefront of Genomics

https://doi.org/10.1038/s41586-020-2817-4
Received: 30 June 2020
Accepted: 4 September 2020
Published online: 28 October 2020
Check for updates

Eric D. Green[1]✉, Chris Gunter[1], Leslie G. Biesecker[1], Valentina Di Francesco[1], Carla L. Easter[1], Elise A. Feingold[1], Adam L. Felsenfeld[1], David J. Kaufman[1], Elaine A. Ostrander[1], William J. Pavan[1], Adam M. Phillippy[1], Anastasia L. Wise[1], Jyoti Gupta Dayal[1], Britny J. Kish[1], Allison Mandich[1], Christopher R. Wellington[1], Kris A. Wetterstrand[1], Sarah A. Bates[1], Darryl Leja[1], Susan Vasquez[1], William A. Gahl[1], Bettie J. Graham[1], Daniel L. Kastner[1], Paul Liu[1], Laura Lyman Rodriguez[1], Benjamin D. Solomon[1], Vence L. Bonham[1], Lawrence C. Brody[1], Carolyn M. Hutter[1] & Teri A. Manolio[1]

Starting with the launch of the Human Genome Project three decades ago, and continuing after its completion in 2003, genomics has progressively come to have a central and catalytic role in basic and translational research. In addition, studies increasingly demonstrate how genomic information can be effectively used in clinical care. In the future, the anticipated advances in technology development, biological insights, and clinical applications (among others) will lead to more widespread integration of genomics into almost all areas of biomedical research, the adoption of genomics into mainstream medical and public-health practices, and an increasing relevance of genomics for everyday life. On behalf of the research community, the National Human Genome Research Institute recently completed a multi-year process of strategic engagement to identify future research priorities and opportunities in human genomics, with an emphasis on health applications. Here we describe the highest-priority elements envisioned for the cutting-edge of human genomics going forward—that is, at 'The Forefront of Genomics'.

Beginning in October 1990, a pioneering group of international researchers began an audacious journey to generate the first map and sequence of the human genome, marking the start of a 13-year odyssey called the Human Genome Project[1–3]. The successful and early completion of the Project in 2003, which included parallel studies of a set of model-organism genomes, catalysed enormous progress in genomics research. Leading the signature advances has been a greater than one million-fold reduction in the cost of DNA sequencing[4]. This decrease has allowed the generation of innumerable genome sequences, including hundreds of thousands of human genome sequences (both in research and clinical settings), and the continuous development of assays to identify and characterize functional genomic elements[4,5]. These new tools, together with increasingly sophisticated statistical and computational methods, have enabled researchers to create rich catalogues of human genomic variants[2,3], to gain an ever-deepening understanding of the functional complexities of the human genome[5], and to determine the genomic bases of thousands of human diseases[5,6]. In turn, the past decade has brought the initial realization of genomic medicine[7], as research successes have been converted into powerful tools for use in healthcare, including somatic genome analysis for cancer (enabling development of targeted therapeutic agents)[7], non-invasive prenatal genetic screening[7], and genomics-based tests for a growing set of paediatric conditions and rare disorders[8], among others.

In essence, with growing insights about the structure and function of the human genome and ever-improving laboratory and computational technologies, genomics has become increasingly woven into the fabric of biomedical research, medical practice, and society. The scope, scale, and pace of genomic advances so far were nearly unimaginable when the Human Genome Project began; even today, such advances are yielding scientific and clinical opportunities beyond our initial expectations, with many more anticipated in the next decade.

Embracing its leadership role in genomics, the National Human Genome Research Institute (NHGRI) has developed strategic visions for the field at key inflection points, in particular at the end of the Human Genome Project in 2003[1] and then again at the beginning of the last decade in 2011[9]. These visions outlined the most compelling opportunities for human genomics research. In each case informed by a multi-year engagement process, NHGRI endeavoured to start the new decade with an updated strategic vision for human genomics research. Through a planning process that involved more than 50 events (such as dedicated workshops, conference sessions, and webinars) over the past two years (see http://genome.gov/genomics2020), the institute collected input from a large number of stakeholders, with the resulting input catalogued and synthesized using the framework depicted in Fig. 1.

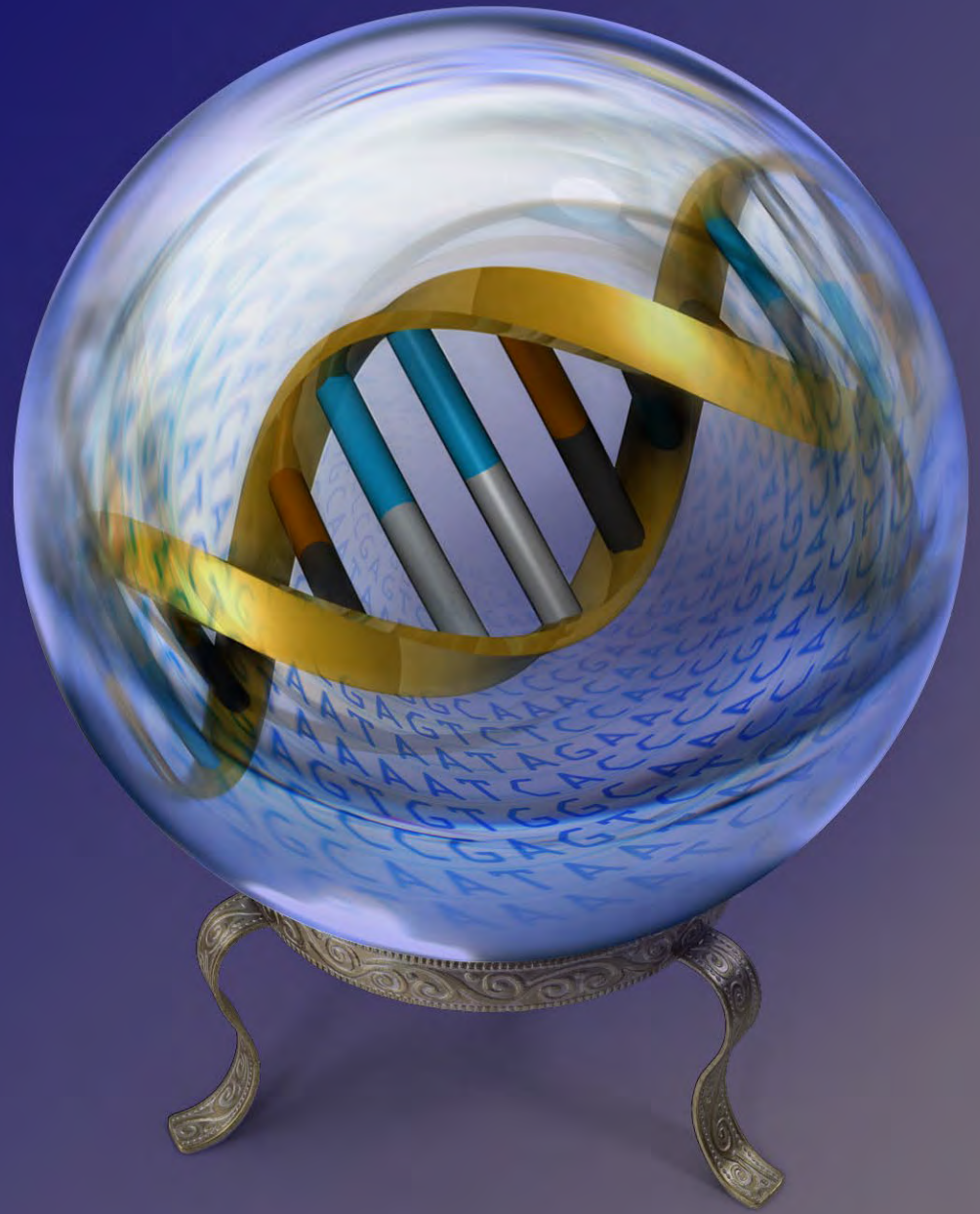
Unlike the past, this round of strategic planning was greatly influenced by the now widely disseminated nature of genomics across biomedicine. A representative glimpse into this historic phenomenon is illustrated in Fig. 2. During the Human Genome Project, NHGRI was the primary funder of human genomics research at the US National Institutes of Health (NIH), but the past two decades have brought a greater than tenfold increase in the relative fraction of funding coming from other parts of the NIH.

National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA. ✉e-mail: egreen@mail.nih.gov

Nature (2020)

genome.gov/2020sv

Bold Predictions for Human Genomics by 2030



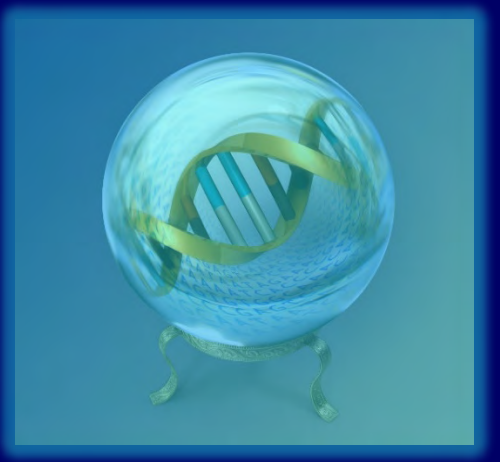
Bold Predictions for Human Genomics by 2030

Box 5

Bold predictions for human genomics by 2030

Some of the most impressive genomics achievements, when viewed in retrospect, could hardly have been imagined ten years earlier. Here are ten bold predictions for human genomics that might come true by 2030. Although most are unlikely to be fully attained, achieving one or more of these would require individuals to strive for something that currently seems out of reach. These predictions were crafted to be both inspirational and aspirational in nature, provoking discussions about what might be possible at The Forefront of Genomics in the coming decade.

1. Generating and analysing a complete human genome sequence will be routine for any research laboratory, becoming as straightforward as carrying out a DNA purification.
2. The biological function(s) of every human gene will be known; for non-coding elements in the human genome, such knowledge will be the rule rather than the exception.
3. The general features of the epigenetic landscape and transcriptional output will be routinely incorporated into predictive models of the effect of genotype on phenotype.
4. Research in human genomics will have moved beyond population descriptors based on historic social constructs such as race.
5. Studies that involve analyses of genome sequences and associated phenotypic information for millions of human participants will be regularly featured at school science fairs.
6. The regular use of genomic information will have transitioned from boutique to mainstream in all clinical settings, making genomic testing as routine as complete blood counts.
7. The clinical relevance of all encountered genomic variants will be readily predictable, rendering the diagnostic designation 'variant of uncertain significance (VUS)' obsolete.
8. An individual's complete genome sequence along with informative annotations will, if desired, be securely and readily accessible on their smartphone.
9. Individuals from ancestrally diverse backgrounds will benefit equitably from advances in human genomics.
10. Breakthrough discoveries will lead to curative therapies involving genomic modifications for dozens of genetic diseases.



Embracing that role, NHGRI formulated the strategic vision described here, which provides an optimistic outlook that the successes in human genomics over the past three decades will be amplified in the coming decade. Many of the details about what is needed to fulfil the promise of genomics have now come into focus. Major unsolved problems remain—among them determining the role for the vast majority of functional elements in the human genome (especially those outside of protein-coding regions), understanding the full spectrum of genomic variation (especially that implicated in human disease), developing data-science capabilities (especially those that keep pace with data generation), and improving healthcare through the implementation of genomic medicine (especially in the areas of prevention, diagnosis, and therapeutic development). The new decade also brings research questions related to the societal implications of genomics, including those related to social inequities, pointing to the continued importance of investigating the ethical, legal, and social issues related to genomics. But now more than ever, solutions to these problems seem to be within striking distance. Towards that end (and with the characteristic spirit of genomics audacity), we offer ten bold predictions of what might be realized in human genomics by 2030 (Box 5).

Box 5

Bold predictions for human genomics by 2030

Some of the most impressive genomics achievements, when viewed in retrospect, could hardly have been imagined ten years earlier. Here are ten bold predictions for human genomics that might come true by 2030. Although most are unlikely to be fully attained, achieving one or more of these would require individuals to strive for something that currently seems out of reach. These predictions were crafted to be both inspirational and aspirational in nature, provoking discussions about what might be possible at The Forefront of Genomics in the coming decade.

1. Generating and analysing a complete human genome sequence will be routine for any research laboratory, becoming as straightforward as carrying out a DNA purification.
2. The biological function(s) of every human gene will be known; for non-coding elements in the human genome, such knowledge will be the rule rather than the exception.
3. The general features of the epigenetic landscape and transcriptional output will be routinely incorporated into predictive models of the effect of genotype on phenotype.
4. Research in human genomics will have moved beyond population descriptors based on historic social constructs such as race.
5. Studies that involve analyses of genome sequences and associated phenotypic information for millions of human participants will be regularly featured at school science fairs.
6. The regular use of genomic information will have transitioned from boutique to mainstream in all clinical settings, making genomic testing as routine as complete blood counts.
7. The clinical relevance of all encountered genomic variants will be readily predictable, rendering the diagnostic designation 'variant of uncertain significance (VUS)' obsolete.
8. An individual's complete genome sequence along with informative annotations will, if desired, be securely and readily accessible on their smartphone.
9. Individuals from ancestrally diverse backgrounds will benefit equitably from advances in human genomics.
10. Breakthrough discoveries will lead to curative therapies involving genomic modifications for dozens of genetic diseases.

Embracing that role, NHGRI formulated the strategic vision described here, which provides an optimistic outlook that the successes in human genomics over the past three decades will be amplified in the coming decade. Many of the details about what is needed to fulfil the promise of genomics have now come into focus. Major unsolved problems remain—among them determining the role for the vast majority of functional elements in the human genome (especially those outside of protein-coding regions), understanding the full spectrum of genomic variation (especially that implicated in human disease), developing data-science capabilities (especially those that keep pace with data generation), and improving healthcare through the implementation of genomic medicine (especially in the areas of prevention, diagnosis, and therapeutic development). The new decade also brings research questions related to the societal implications of genomics, including those related to social inequities, pointing to the continued importance of investigating the ethical, legal, and social issues related to genomics. But now more than ever, solutions to these problems seem to be within striking distance. Towards that end (and with the characteristic spirit of genomics audacity), we offer ten bold predictions of what might be realized in human genomics by 2030 (Box 5).

Box 5

Bold predictions for human genomics by 2030

Some of the most impressive genomics achievements, when viewed in retrospect, could hardly have been imagined ten years earlier. Here are ten bold predictions for human genomics that might come true by 2030. Although most are unlikely to be fully attained, achieving one or more of these would require individuals to strive for something that currently seems out of reach. These predictions were crafted to be both inspirational and aspirational in nature, provoking discussions about what might be possible at The Forefront of Genomics in the coming decade.

1. Generating and analysing a complete human genome sequence will be routine for any research laboratory, becoming as straightforward as carrying out a DNA purification.
2. The biological function(s) of every human gene will be known; for non-coding elements in the human genome, such knowledge will be the rule rather than the exception.
3. The general features of the epigenetic landscape and transcriptional output will be routinely incorporated into predictive models of the effect of genotype on phenotype.
4. Research in human genomics will have moved beyond population descriptors based on historic social constructs such as race.
5. Studies that involve analyses of genome sequences and associated phenotypic information for millions of human participants will be regularly featured at school science fairs.
6. The regular use of genomic information will have transitioned from boutique to mainstream in all clinical settings, making genomic testing as routine as complete blood counts.
7. The clinical relevance of all encountered genomic variants will be readily predictable, rendering the diagnostic designation 'variant of uncertain significance (VUS)' obsolete.
8. An individual's complete genome sequence along with informative annotations will, if desired, be securely and readily accessible on their smartphone.
9. Individuals from ancestrally diverse backgrounds will benefit equitably from advances in human genomics.
10. Breakthrough discoveries will lead to curative therapies involving genomic modifications for dozens of genetic diseases.

Embracing that role, NHGRI formulated the strategic vision described here, which provides an optimistic outlook that the successes in human genomics over the past three decades will be amplified in the coming decade. Many of the details about what is needed to fulfil the promise of genomics have now come into focus. Major unsolved problems remain—among them determining the role for the vast majority of functional elements in the human genome (especially those outside of protein-coding regions), understanding the full spectrum of genomic variation (especially that implicated in human disease), developing data-science capabilities (especially those that keep pace with data generation), and improving healthcare through the implementation of genomic medicine (especially in the areas of prevention, diagnosis, and therapeutic development). The new decade also brings research questions related to the societal implications of genomics, including those related to social inequities, pointing to the continued importance of investigating the ethical, legal, and social issues related to genomics. But now more than ever, solutions to these problems seem to be within striking distance. Towards that end (and with the characteristic spirit of genomics audacity), we offer ten bold predictions of what might be realized in human genomics by 2030 (Box 5).

the use of genomics in medicine from diagnosing and treating disease to maintaining health.

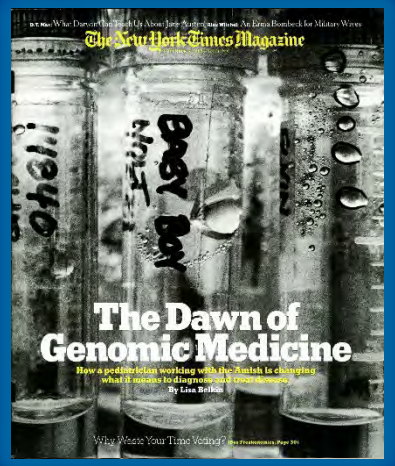
Sharp barriers between research and clinical care obstruct the virtuous cycle of moving scientific discoveries rapidly into clinical care and bringing clinical observations back to the research setting[42] (Fig. 3). Learning healthcare systems—in which real-time data on outcomes of healthcare delivery are accessed and used to enhance clinical practice—can lead to continuous care improvement, but only if the barriers between research and clinical care are reduced[142]. For example, offering genome sequencing to all members of a healthcare system, performed in conjunction with research and participant engagement and provided in real time[84], could help to assess the clinical utility of genomic information and may allow providers to improve disease diagnosis and management. System-wide implementation of such an experiment requires not only extensive patient and provider education, sophisticated informatics capabilities, and genomics-based clinical decision support, but also the development and evaluation of data security and privacy protections to ensure patient confidentiality[136]. Patients should be engaged in the design of such systems and informed at entry to them (and periodically thereafter), so as to be fully aware of the nature of the ongoing research with their clinical data and the goals and potential risks of their participation[137]. Extending such studies across many healthcare systems should reveal common challenges and solutions[138,139], thereby enhancing the learning healthcare model for genomic medicine more broadly (Fig. 3).

Concluding thoughts

The dawn of genomics featured the launch of the Human Genome Project in October 1990[1]. Three decades later, the field has seen stunning technological advances and high-profile programmatic successes, which in turn have led to the widespread infusion of genomic methods and approaches across the life sciences and, increasingly, into medicine and society.

NHGRI has for the third time[133] since the Human Genome Project undergone an extensive horizon-scanning process to capture, synthesize, and articulate the most compelling strategic opportunities for the next phase of genomics—with particular attention to elements that are most relevant to human health. The now near-ubiquitous nature of genomics (including in the complex healthcare ecosystem) presented practical challenges for attaining a holistic assessment of the field. Another reality was that the NHGRI investment in genomics has now been multiplied many-fold by the seeding of human genomics throughout the broader research community. These changes reflect a continued maturation of both the field (in general) and NHGRI (more specifically), nicely aligning with the institute's evolving leader-

A ~21-year Pivot:
Bringing Genomic Medicine Into Focus



Let's Stay Connected!

genome.gov/stayconnected

Subscribe:

genome.gov/email

Follow:

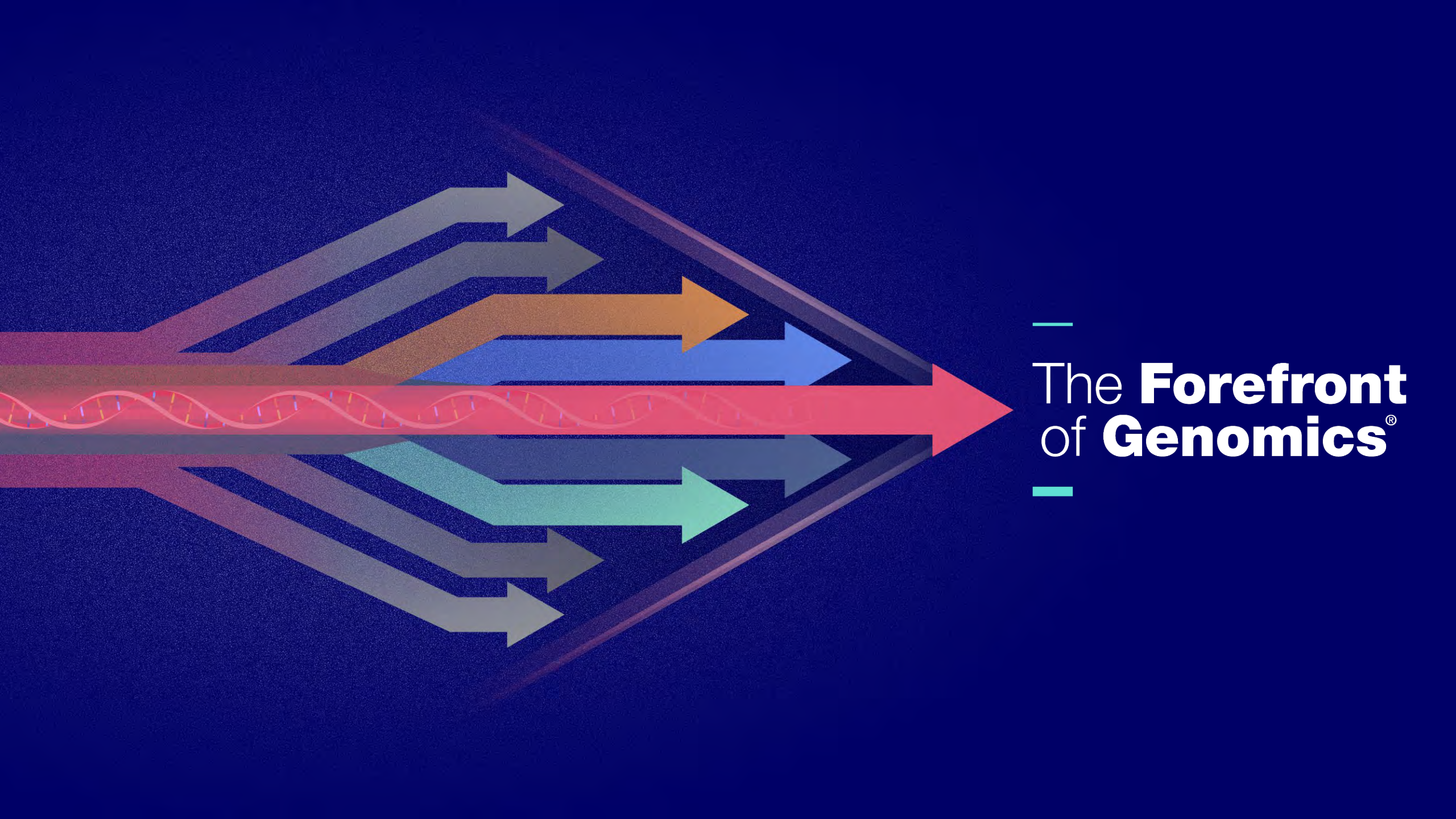
@NHGRI_Director

Follow:

linkedin.com/in/eric-green-md-phd

Follow:

@NHGRI_Director


The Forefront
of Genomics®

How to Represent a Human Pangenome Reference

Human Pangenome Reference Graph

Visualizing a Human Pangenome Reference