INNOVATION

Harnessing the human genome is the future of healthcare—and Johnson & Johnson is helping lead the way

The company’s partnership with the largest human genome sequencing project in the world will increase scientists’ understanding of genetic diseases and help create new interventions. Here, a look at the breakthroughs that have guided the understanding of the power of DNA.

WRITTEN BY BARBARA BRODY | NOVEMBER 30, 2023

Share

Listen to this article now 1.0x

00:00 11:14

Genetically speaking, humans are 99.5-99.9% identical to one another. That remaining 0.1-0.5% of unique genetic material? It helps make us who we are.

It's responsible for a number of traits you can see, like your hair color, as well as many that you can't, including disease risk. In the past few decades, scientists have made numerous advances in teasing out how such variability connects to human health, but they've merely scratched the surface.

Now the entire field of genetics (which focuses on single genes and their effects on phenotype) and genomics (the study of how all genes and their interactions influence an organism's biology) is about to get turbocharged.

On November 30, the UK Biobank Whole Genome Sequencing (WGS) Project unveiled data from the whole genome sequencing of its 500,000 volunteer participants. This monumental feat, in which Johnson & Johnson played a major role, means that approved researchers will have access to an unprecedented amount of data that can be used to speed novel discoveries. "Sequencing" technology to identify every component that makes up a person's genome enables scientists to uncover gene variants that may predispose an individual to a certain disease or make them more or less likely to respond to a specific treatment.

> “
>
> Scientists can apply to use this huge data set—the creation of which has been supported by Johnson & Johnson—to combine information about several gene

> variants and more accurately predict disease risk. ”
>
> Trevor Howe, Ph.D.

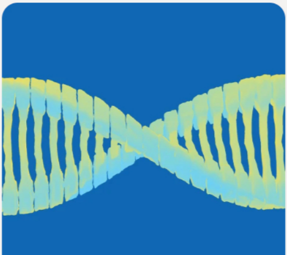
“The size of this database is phenomenal,” says Trevor Howe, Ph.D., Director, External Innovation, R&D Data Science & Digital Health and Fellow, Translational Genomics at Johnson & Johnson Innovative Medicine. “It contains not only the genomic sequences of half a million volunteers, but also links that data to extensive patient hospital health records, providing an end-to-end look at a patient’s overall health journey. If we can link genetics with health outcomes, we can better understand the drivers of disease which enable us to develop and deliver more targeted therapies to treat them.”

This study from UK Biobank, a comprehensive biomedical database that launched in 2012 (after scientists began recruiting participants from 2006 to 2010), would not have been possible without the foundations laid by earlier generations of genetic and genomic research. Here’s a look at some key milestones in the quest to harness DNA.

1953

The double helix is discovered

DNA was first identified in 1869 by Friedrich Miescher, a Swiss physician and biologist who called it “nuclein.” The next most notable DNA



discovery: Four UK-based researchers figure out that DNA is comprised of a double helix structure.

Why was this such a breakthrough? Because this unique shape suggested how DNA might be replicated when cells divide. It also marked the beginning of the biotechnology industry. Without knowledge of DNA's structure, scientific techniques including DNA sequencing, genetic engineering and functional genomics would not exist.

1977

A UK scientist develops the first DNA sequencing technique

The DNA double helix looks like a twisted ladder. Each "rung" is comprised of two nucleotides, the building blocks of DNA that form base pairs. DNA sequencing, also called genome sequencing, determines the exact order of these base pairs in a living organism. Humans have about 3 billion base pairs, which explains why the earliest attempts at genomic sequencing undertaken later in the 20th century were so costly and time-consuming.

In 1977, Fred Sanger, an English scientist, pioneers the Chain Termination Method of sequencing, aka the Sanger Sequencing. He and his team sequence the first full genome, not of a human but a virus. (All living things have a unique genome, including viruses.)

1989

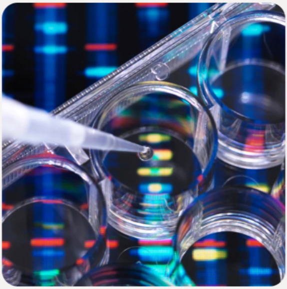
The first gene linked to a disease emerges

In the mid-1980s, Canadian scientist Lap-Chee Tsui and his team begin looking for the gene associated with cystic fibrosis, and in 1989 they pinpoint a gene called CFTR. It is the first time anyone has discovered a disease-causing gene.

Although the search for CFTR took years, scientists today could use modern gene sequencing to make a similar finding in just a few hours.

2003

The first human genome is successfully sequenced



The Human Genome Project was launched in 1990, and 13 years later results in the first-ever sequence of the human genome—or 92% of it, which was as far as scientists could get with the technology available at the time. There have been numerous refinements since then, and the most complete version was released in 2022 by scientists at the National Institutes of Health's National Human Genome Research Institute. Humans have 20,000 to 25,000 genes, and this latest reference genome is a map against which individual genomes can be compared.

Since the first human genome sequence, scientists around the world have been sequencing individual genomes and comparing them against the reference standards to identify disease-causing variants. By 2013, scientists had identified hundreds of variants linked to disease that could be treated if they are detected and acted upon with medical strategies.

The Human Genome Project helps jump-start the development of new diagnostic tools and treatments. The most growth has occurred in the oncology field, where scientists have developed targeted treatments that are effective in patients whose tumors express a particular genetic variant or biomarker.

2019

Janssen's first targeted therapy gets FDA approval

After 22 years of development, a Janssen drug developed to treat aggressive bladder cancer hits the market. This novel therapy is effective in patients whose tumor has a specific abnormal gene.

Around the same time, the company deepens its commitment to developing precision medicine strategies and bringing the innovations to patients. In 2021, a Janssen medicine became the first targeted therapy approved by the U.S. Food and Drug Administration (FDA) for advanced non-small cell lung cancer in patients with a specific genetic mutation. In 2023, a Janssen therapeutic was FDA-approved for use as a first-line targeted treatment for prostate cancer patients who carry a certain mutation and for whom chemotherapy is not clinically indicated.

The company is currently pursuing precision treatments for a variety of conditions, including those related to the cardiovascular, immune and neurological systems.

2019

Johnson & Johnson launches its partnership with UK Biobank

Johnson & Johnson joins forces with the UK's national funding agency investing in science and research, the Wellcome Trust (a global charitable science foundation) and three other pharmaceutical companies to support whole genome sequencing of all UK Biobank participants.

Set up almost 20 years ago, UK Biobank recruited half a million altruistic volunteers to create the world's most comprehensive source of health data. It is available to approved researchers worldwide via a protected database containing only de-identified data (e.g. name, address, date of birth and more stripped out), and drives scientific discoveries that improve human health. The rich database is regularly augmented with additional data from the half a million volunteers on lifestyle, whole body imaging scans, health information and proteins found in the blood.

This collaboration, facilitated through the London-based Johnson & Johnson Innovation team, produces the world's largest repository of WGS data.

"Nowhere in the world had anyone attempted human whole genome sequencing at this scale, but we could see the potential it had to generate novel, highly relevant human disease targets—and with targeted therapeutics, greatly improve the potential for clinical trial success," says Nerida Scott, Ph.D., Regional Head, Johnson & Johnson Innovation for Europe, the Middle East and Africa. "In addition, we believed it could give us the opportunity to screen all our internal programs and external partnering opportunities to confirm human target relevance and help increase the chance of successfully delivering products that transform human health."

Johnson & Johnson and the other partners were committed to whole genome sequencing rather than whole exome sequencing (WES), which covers only 2-3% of the genome. WGS is a major enhancement on WES.

"WGS captures the entire genome, enabling detection of coding and noncoding genetic variation, structural variants and variation in complex regions that are technically challenging to genotype," says Shuwei Li, Ph.D., Director, Population Analytics & Insights, R&D Data Science & Digital Health, Johnson & Johnson Innovative Medicine. "When you do whole exome sequencing, you're only looking at the coding region; it's like one thin slice of an apple."

For years scientists referred to noncoding DNA as "junk DNA." But it's become clear that some noncoding sequences regulate gene expression and protein formation. "We want to characterize the whole genomic landscape and integrate WGS into the human multi-omic framework built by the Population Analytics team to support target identification, prioritization, precision strategies and clinical trials," says Li.

2023

UK Biobank sequences 500,000 genomes

On November 30, after a period of exclusive access for Johnson & Johnson and the three industry WGS partners, UK Biobank announces that the sequenced genomes of all 500,000 participants are available to approved researchers across the globe.

Scientists can now apply to use this huge data set, which has over 10,000 variables for each participant, to combine information about several gene variants and more accurately predict disease risk, says Howe. That's key, because very few diseases directly stem from a single gene mutation. Most common illnesses—like heart disease, diabetes, depression and cancer—likely result from the interaction of several genes, as well as environmental factors.

This addition to UK Biobank's database also promises to streamline clinical development of new therapies by helping match patients to the best treatment for them, says Scott. Two people with rheumatoid arthritis, for instance, might show nearly identical symptoms, yet the biology of what's driving the underlying disease and how to best treat it might be drastically different.

The availability of the sequenced genomes could help expedite the development of novel drugs. "It will enhance precision strategies by making research more effective and yielding better, more precise treatments for patients around the world," she adds.