The Economic Impact and Functional Applications of Human Genetics and Genomics

Commissioned by the American Society of Human Genetics
Produced by TEConomy Partners, LLC.
Report Authors: Simon Tripp and Martin Grueber

May 2021

ASHG
American Society of Human Genetics

TECONOMY
PARTNERS LLC

Acknowledgements

ASHG and the project authors wish to thank the following organizations for their generous support of this study.

Invitae Corporation, San Francisco, CA
Regeneron Pharmaceuticals, Inc., Tarrytown, NY

The project authors express their sincere appreciation to the following individuals who provided their advice and input to this project.

ASHG Government and Public Advocacy Committee

Lynn B. Jorde, PhD
ASHG Government and Public Advocacy Committee (GPAC) Chair, President (2011)
Professor and Chair of Human Genetics
George and Dolores Eccles Institute of Human Genetics
University of Utah School of Medicine

Katrina Goddard, PhD
ASHG GPAC Incoming Chair, Board of Directors (2018-2020)
Distinguished Investigator, Associate Director, Science Programs
Kaiser Permanente Northwest

Melinda Aldrich, PhD, MPH
Associate Professor, Department of Medicine, Division of Genetic Medicine
Vanderbilt University Medical Center

Wendy Chung, MD, PhD
Professor of Pediatrics in Medicine and
Director, Clinical Cancer Genetics
Columbia University

Mira Irons, MD
Chief Health and Science Officer
American Medical Association

Peng Jin, PhD
Professor and Chair, Department of Human Genetics
Emory University

Allison McCague, PhD
Science Policy Analyst, Policy and Program Analysis Branch
National Human Genome Research Institute

Rebecca Meyer-Schuman, MS
Human Genetics Ph.D. Candidate, Antonellis Lab
University of Michigan

Charles Schwartz, PhD
Senior Research Scientist Emeritus
Greenwood Genetic Center

Wendy Uhlmann, MS, CGC
Clinical Professor of Internal Medicine and Clinical Professor of Human Genetics
Genetic Counselor and Clinic Coordinator, Medical Genetics Clinic
University of Michigan

David Wheeler, PhD
Director of Precision Genomics
Department of Computational Biology
St Jude Children's Research Hospital

Advisors

Lawrence Brody, PhD
Director, Division of Genomics and Society
National Human Genome Research Institute

Eric D. Green, MD, PhD
Director
National Human Genome Research Institute

Cristina Kapustij, MS
Chief, Policy and Program Analysis Branch
National Human Genome Research Institute

Cynthia Casson Morton, PhD
ASHG President (2014)
Professor of Obstetrics, Gynecology and Reproductive Biology and Professor of Pathology
Harvard Medical School
Chair, Obstetrics and Gynecology, and Director of Cytogenetics
Brigham and Women's Hospital

Heather C. Mefford, MD, PhD
ASHG Program Committee Chair (2018), Board of Directors (2021)
Associate Professor
Genetics, Epilepsy Program—Center for Integrative Brain Research
University of Washington and Seattle Children's

Kiran Musunuru, MD, PhD, MPH, ML
ASHG Program Committee Chair (2019), Board of Directors (2021)

Professor of Medicine
Director, Genetic and Epigenetic Origins of Disease Program, Cardiovascular Institute
Perelman School of Medicine at the University of Pennsylvania

David L. Nelson, PhD
ASHG President (2018)
Professor
Molecular and Human Genetics
Baylor College of Medicine

Telephone and Email Interview Participants

Todd R. Golub, MD
Director, Founding Core Institute Member, Broad Institute of MIT and Harvard, and
Founding Core Member and Director of the Gerstner Center for Cancer Diagnostics, Broad Institute

Leroy Hood, MD, PhD
Senior Vice President and Chief Science Officer
Providence St. Joseph Health, and
Chief Strategy Officer and Professor, Institute for Systems Biology

Kathryn A. Phillips, PhD
Professor of Health Economics and Health Services Research, and
Founding Director, UCSF Center for Translational and Policy Research on Personalized Medicine
University of California, San Francisco

Philip Reilly, MD, JD
Venture Partner, Third Rock Ventures

David Veenstra, PharmD, PhD
Professor and Associate Director, Comparative Health Outcomes, Policy & Economics (CHOICE) Institute
School of Pharmacy, University of Washington

Carrie D. Wolinetz, PhD
Associate Director for Science Policy
National Institutes of Health

American Society of Human Genetics Staff

Mona Miller, MPP
Chief Executive Officer

Derek Scholes, PhD
Senior Director, Policy and Advocacy

Lyly G. Luhachack, PhD
Policy and Advocacy Specialist

Contents

Executive Summary

Modern life sciences and associated advancements in biopharmaceuticals, diagnostics, medical devices, and healthcare services have enabled unprecedented improvements in human health and longevity.

Perhaps nowhere has life science research advanced more in the modern age than through insights provided by genetics and genomics. This field is both fundamental in biological research—elucidating the basic code of life, DNA, upon which our form and function depend—and in enabling applied and translational discoveries across most diseases and health disorders.

This report examines and describes the positive impacts that are derived from modern human genetics and genomics science and its associated commercial and clinical applications on the nation's economy, society, and the health and well-being of individuals.

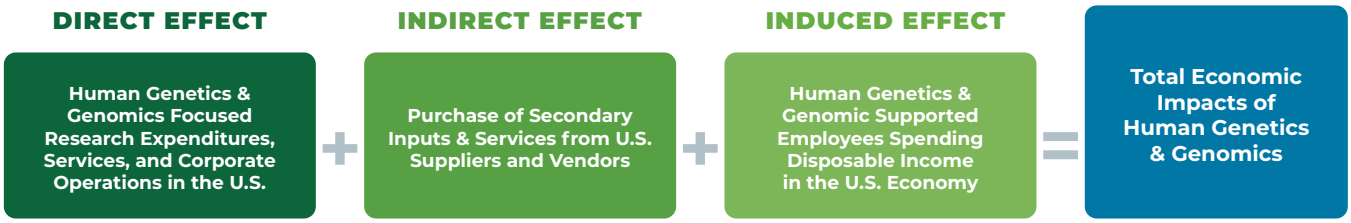
Twenty years after the completion of the Human Genome Project, there has been widespread expansion and application of human genetics and genomics technologies. Technologies for sequencing and for genome analysis have advanced quite spectacularly—to the extent that genome sequencing is now both fast and affordable. The technologies of genetics and genomics, and the research advancements they have enabled scientists to make, have now brought human genetics and genomics to a visible inflection point—a point in time where scientific discoveries are rapidly translating into clinical insights and significant human health and well-being advancements.

The Economic Impact of the Human Genetics and Genomics Sector

The U.S. economy has advanced on the back of scientific progress—progress that has enabled national leadership in diverse industries such as aerospace, energy, agriculture, transportation, advanced materials, information technology, and biotechnology. Continuing to strengthen the competitiveness of the U.S. economy requires ongoing expansion of the national capacity for innovation and the scientific and technological research and development (R&D) upon which innovation depends. Particularly important is leveraging science and innovation to give rise to new, fast-growing, advanced industries that spark economic growth and improved standards of living. Born out of federal investment in the Human Genome Project, the U.S. achieved early leadership in the genetics and genomics industry—leadership that has resulted in the growth of an important and dynamic economic sector.

Substantial U.S. economic activity, supporting a large volume of high-paying jobs across the nation, is generated from the performance of genetic and genomic research, the development and manufacturing of commercial genomic technologies, the broad range of diagnostics products and therapeutics on the

Figure ES-1: The Economic Impact of the Human Genetics and Genomics Sector in the United States



$3.3B
FEDERAL RESEARCH

Federal research funding, using a conservative definition of what constitutes human genetics and genomics research, reached $3.3 billion in 2019, with most of this coming from NIH.

152,000
INDUSTRY JOBS

89,464 core private sector industry jobs and an estimated 62,710 additional extended industry jobs (related employment share from major pharmaceutical and medical testing/diagnostics companies).

850,000
TOTAL SUPPORTED JOBS

With a direct employment estimate of nearly 166,000 academic and industry jobs, human genetics and genomics supports more than 850,000 total jobs. Each direct human genetics and genomics job supports 4.12 additional jobs in the U.S. economy.

$265B
TOTAL ECONOMIC IMPACT

The direct economic activity generated by the human genetics and genomics industry exceeds $108 billion in 2019 and ultimately supports a total of more than $265 billion across the U.S. economy. Every $1.00 of direct human genetics and genomics activity generates an additional $1.45 in the U.S. economy.

$5.2B
DIRECT FEDERAL TAX REVENUES

The federal tax revenues of $5.2 billion generated by the direct operations of the human genetics and genomics domain alone surpasses the single year federal investment in human genetics and genomics of approximately $3.3 billion across all federal agencies.

4.75:1.00
FEDERAL RETURN ON INVESTMENT

In the simplest of terms, from a federal investment and revenue perspective, the overall economic impacts of U.S. human genetics and genomics generates a return on investment (ROI) of more than 4.75 to 1.00 ($3.3 billion in federal investment in human genetics and genomics – while the whole domain generates $15.5 billion in federal tax revenues).

Source: TEConomy Partners, LLC.

market that are derived from genomics knowledge and have pharmacogenomic associations, and the associated healthcare services that are delivered. The economic impact of the human genetics and genomics sector on the U.S. economy is assessed using the standard regional economics methodology of input-output analysis. The results demonstrate the growth of a powerful economic sector across the nation— a sector that has grown five-fold in its annual economic impact since 2010. Even more importantly, it is a sector that also generates robust functional impacts in terms of human health and well-being. Figure ES-1 summarizes some of the topline findings from the economic impact analysis.

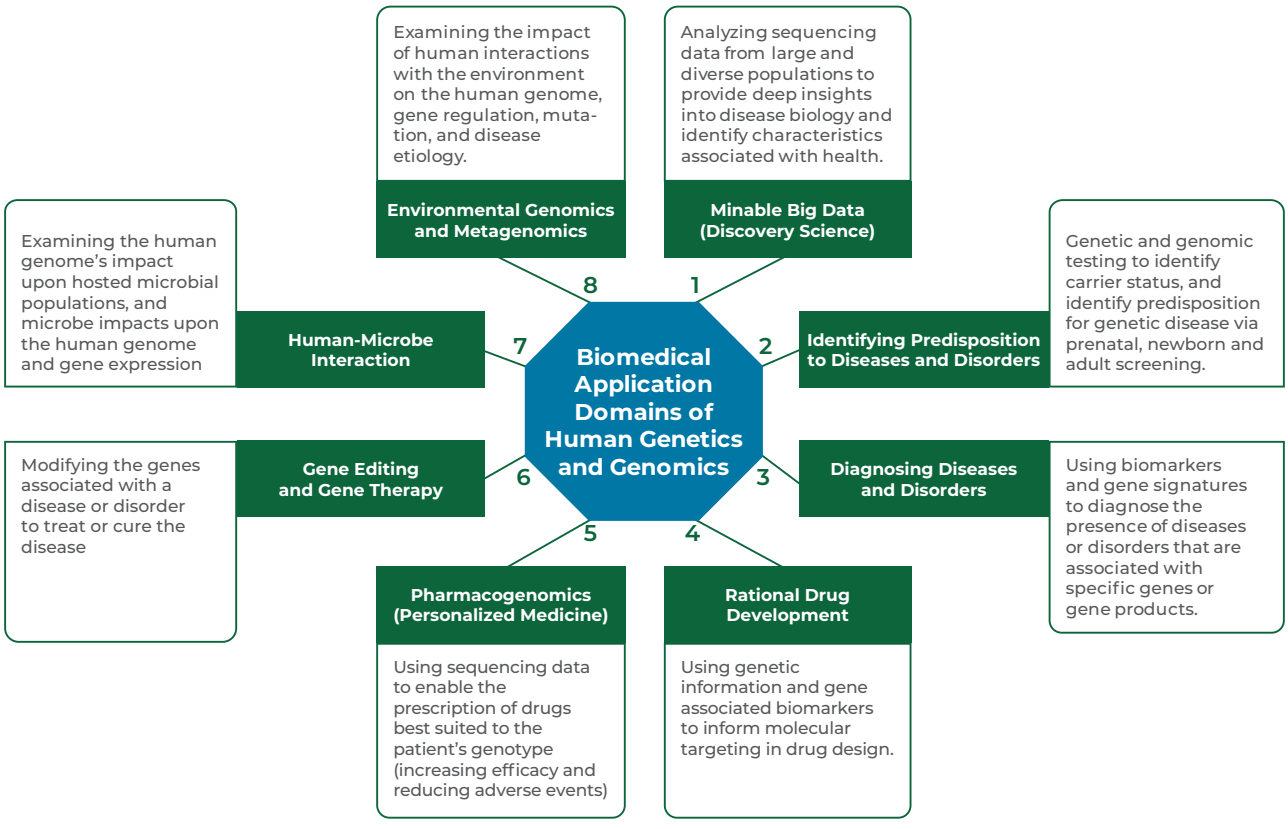
The sector also supports high wage jobs. Because it requires a well-educated and technically skilled workforce, direct jobs in the genetics and genomics sector pay more than $130,000 in annual total compensation (income and benefits) per worker, while the total jobs supported by human genetics and genomics economic activity (direct + induced) average greater than $81,000 in compensation per employee.

The Functional Impacts of Human Genetics and Genomics

The speed and affordability of gene sequencing and advanced genomic data analytics have helped produce deep biomedical insights and innovations, which are being combined with advancements in biopharmaceuticals, diagnostics, and other medical technologies that leverage genomic information. An evident tipping point has been achieved where the utility of genomics and wide-spread use of sequencing is clearly advantageous for significantly enhancing human health outcomes. As this report highlights, the functional application of human genetics and genomics to clinical healthcare is now a daily reality in some medical fields (e.g., cancer diagnosis and treatment) and is increasingly front-and-center in neurological, psychiatric, gastrointestinal, immunologic, rheumatologic, dermatologic, pain management, and other application areas of clinical medicine. It is also fundamental to advancements being made in the diagnosis and treatment of a wide range of rare diseases and disorders—helping to end the diagnostic odysseys of millions of patients afflicted with rare diseases that have been difficult to diagnose and sparsely served in terms of available treatments.

In reviewing the functional applications of human genetics and genomics, the authors find that the positive impacts being generated are highly diverse—generated within eight major domains

Figure ES-2: Functional Biomedical Impact Domains (Applications) of Human Genetics and Genomics



Source: TEConomy Partners, LLC.

of activities impacting human health. These are summarized and briefly described in Figure ES-2.

The eight domains identified in Figure ES-2 are already having profound impacts in advancing clinical health sciences and health outcomes. Each of these areas is profiled briefly below and detailed further in the full body of the report.

1. Minable Big Data (Discovery Science)

Advancements in high-speed gene sequencing technologies have facilitated the assembly of exabytes[1] of genomic information that can be analyzed (assisted by highly advanced and automated analytical systems) for unique insights into genome structure and function and the association of gene variants with human diseases and health disorders. It is anticipated that by 2025 more than 60 million patients will have had their genome sequenced in a healthcare context.[2] Access to extremely large volumes of sequenced individuals provides a rich platform for important scientific discovery and for advancing the identification and classification of genomic variant pathogenicity (variants associated with causation of disease). Both science and technological capabilities are now at the point where the analysis of genomic and phenomic big data provides a powerful pathway forward for biomedical discovery and clinical applications to improve human health.

2. Identifying Predisposition to Diseases and Disorders

One of the primary research and clinical applications of human genetics and genomics is identification of the potential predisposition for individuals to develop specific diseases or health disorders. Modern genetic screening for such predispositions divides into three key categories: 1) carrier screening, which tests a prospective parent for the presence of gene variants that have been shown to be associated with risk of passing down a hereditary disorder (thereby helping to inform family planning and associated decisions); 2) pre-natal and post-natal testing, which focuses on testing for genetic predisposition to disease in the fetus or in newborns; and, 3) child and adult testing.

Information provided by predisposition screening enables patients and their physicians to make informed healthcare decisions, plan follow-up health monitoring strategies, and identify strategies for care using evidence-based clinical best practices.

3. Diagnosing Disease, Rare Diseases, and Disorders

Whole genome and whole exome sequencing are increasingly being used in clinical practice to facilitate the diagnosis of diseases or health disorders. In addition to the many common chronic diseases (such as heart disease, diabetes, cancer, etc.), approximately 7,000 rare diseases have been recognized[3] and have historically been a significant challenge to diagnose. Rare diseases, by their inherent nature of being rare, present diagnostic challenges because so few physicians have encountered them. Often, these diseases may present symptoms seen in other, more common diseases, resulting in an understandable misdiagnosis and inappropriate treatment strategies being adopted. Patients, and their families, may embark on long "diagnostic odysseys", seeing dozens of practitioners, undergoing multiple tests and procedures, enduring fruitless attempts at treatment over many years without ever getting a definitive, accurate diagnosis. Genetic and genomic testing provides a pathway to solving this dilemma in multiple diseases and disorders impacting many thousands of patients.

Collectively, rare diseases have a significant population impact, with approximately 1 in 10 individuals having a rare disease (estimated at between 25-30 million patients in the U.S. and 350 million worldwide).[4] Modern genetic and genomic diagnostic tools, informed by

1 An exabyte = 10006 bytes (1,000,000,000,000,000,000 bytes).
2 Birney, Ewan. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine,* vol. 6, no. 2, March-April 2019.
3 National Center for Advancing Translational Sciences, Genetic and Rare Diseases Information Center. "FAQs About Rare Diseases." https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases. Accessed 12 May 2021.
4 Ibid.

> The ability to tailor a drug regimen to a specific genetic code that is truly personalized to that specific DNA double helix has been a dream of researchers, physicians, and patients alike. Advances in precision medicine, specifically around the genome…are making this dream a reality."
>
> **Kristen Ciriello Pothier. *Personalizing Precision Medicine*. A Global Voyage from Vision to Reality. John Wiley & Sons, Inc., 2017.**

scientific advancements in identifying gene variants associated with specific diseases, are providing clear diagnostic benefits. By deploying genetic and genomic testing, up to and including whole genome sequencing, diagnostic odysseys may be ended for many patients—not only providing a pathway to appropriate treatment but also reducing significant waste in the healthcare system and the associated costs of incorrect diagnosis. Even if no treatment is available, peace of mind can result through simply having an "answer" and being able to end the costly hunt for diagnosis. It has been noted that "this is clearly the most powerful diagnostic tool ever developed for the millions of children with rare diseases."[5]

4. Rational Drug Development

Rational drug development uses genetic and genomic information to advance the development of new biopharmaceuticals to treat diseases. Biomarkers (genes or gene products) are providing molecular targets for purposefully designed drugs that are engineered to bind to targets. The application of genetics and genomics to drug development has resulted in multiple clinical successes, with specific examples highlighted in this report. Biopharmaceutical companies are now able to use genetic and genomic information to target the trials of their pharmaceutical and biologic molecules to patients who have been preselected through the presence of biomarkers (often genetic). This has the potential to advance more drugs successfully to market since they are more likely to demonstrate efficacy in their trials by virtue of being rationally targeted.

It is also notable that progress in genetics and genomics has enabled pharmaceutical research to increasingly address rare diseases—helping to rebalance biopharmaceutical research in terms of work on chronic diseases versus rare diseases.

5. Precision Medicine and Targeted Therapeutics (Pharmacogenetics)

Having an ability to sequence a patient's whole genome rapidly and cost-effectively has opened the door to a new paradigm in healthcare termed "precision medicine" whereby an individual's genetic profile is used to guide decisions made in regard to the prevention, diagnosis, and treatment of disease. The discipline of "pharmacogenetics" (also "pharmacogenomics") has developed as a field of research and, increasingly, clinical practice, that addresses the genetically determined variation in how individuals respond to specific drugs in terms of differences in dose requirement, efficacy, and the risk of adverse drug reactions (ADRs). It is increasingly being employed to help physicians select the "right drug and the right dose" for a patient based on their genome (assuming there is statistically significant clinical information linking a drug to specific gene variants in terms of efficacy and side effects). Currently, pharmacogenetics is improving health outcomes along three primary paths:

5 Kingsmore, Stephen. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine,* vol. 6, no. 2, March-April 2019.

- Selection of the therapeutic (among multiple choices) that is likely to prove most efficacious based on the patient's genome and a drug's proven efficacy for their specific genotype.
- Ruling-out a therapeutic (among multiple choices) based on the patient's genome and a drug's potential for unacceptable adverse side effects given their specific genotype.
- Development of an optimized drug dosage for a patient based on their genotype's influence on the rate at which they will metabolize the drug.

Cancer is perhaps the most well-recognized cluster of disease for which genetic tests may impact drug selection and dosing; however, analysis of U.S. Food and Drug Administration (FDA) data shows that pharmacogenetic associations are also in place for multiple chronic diseases and conditions, covering applications in major categories such as cardiovascular disease, gastroenterological diseases and disorders, infectious diseases, neurological diseases and disorders, psychiatric conditions, and rheumatologic diseases. Pharmacogenetic associations now span a range from relatively rare diseases, such as Tourette's syndrome and Tardive dyskinesia, to common conditions, such as hypercholesterolemia and depression. There are more than 100 drugs for which the associations are now listed by the FDA.

6. Gene Editing and Gene Therapy

As noted above, genetic and genomic advancements are elucidating gene variant associations with the predisposition for disease, providing enhanced diagnosis of diseases, and providing increasingly effective pathways for therapeutics and disease treatment. Another developing approach is to use the expanding knowledge of gene variants associated with disease to provide targets for potential modification of a patient's genes themselves—modification that has the goal of treating, and potentially curing, the target disease through what is termed gene editing or gene therapy. Ultimately, gene editing and gene therapy represent new pathways to the treatment and curing of diseases, but these approaches are still in the early stages of clinical application.[6] Part of the caution in clinical application arises from a need for further study of the potential for off-target gene edits (mutagenesis) to occur in non-targeted genes and for unintended mosaicism to occur. Despite these challenges, there are several important gene therapies that have successfully advanced through clinical trials, helping to treat a series of previously untreatable rare diseases. It is a promising field for ongoing advancement.

7. Human-Microbe Interactions

Each of us is host to communities of trillions of microbes. Microbes serve important functions for humans, for example aiding our digestion and the breakdown of micronutrients, defending us from pathogenetic microbes, and priming our immune system. Recent research has shown that we have a symbiotic two-way genetic interaction with microbes, with microbes impacting our genes and gene expression, and human genotype impacting the make-up of the microbial communities we host.

While microbes play an important positive role in our health, many microbes are pathogenic, being the causative agents for human infectious diseases. Research is finding that individual genomes can be associated with resistance or susceptibility to certain infectious diseases, and the recent COVID-19 pandemic, coinciding with the current significant volumes of patients for which genome sequences are available, has enabled significant clinical study of genome effects on viral susceptibility and resistance.

6 It should be noted that the discussion of gene editing and gene therapy pertains to modifying non-hereditable (somatic) genes—changes to an individual's genes that will only affect the individual being treated but not the genes of future generations. There is ongoing discussion and public debate about the potential use of gene editing to make heritable genetic changes (changes to the germline). Such genome edits would result in changes to an individual's DNA being passed to their progeny and subsequent generations. At the present time, the general consensus of leading organizations in medical genetics, genetics research, and genetic counseling is that genome editing which culminates in human pregnancy should not be undertaken, and that further research is required into the scientific, clinical, and ethical implications of germline editing.

Genomics in the COVID-19 Pandemic

Genomics rapidly assumed crucial roles in COVID-19 research and clinical care in areas such as: (1) the deployment of DNA and RNA sequencing technologies for diagnostics, tracking of viral isolates, and environmental monitoring; (2) the use of synthetic nucleic acid technologies for studying SAR- CoV-2 virulence and facilitating vaccine development; (3) examination of how human genomic variation influences infectivity, disease severity, vaccine efficacy, and treatment response; (4) the adherence to principles and values related to open science, data sharing, and consortia based collaborations; and (5) the provision of genomic data science tools to study COVID-19 pathophysiology. The growing adoption of genomic approaches and technologies into myriad aspects of the global response to the COVID-19 pandemic serves as another important and highly visible example of the integral and vital nature of genomics in modern research and medicine.

Eric D. Green, et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." Nature, vol. 586, no. 29, October 2020.

8. Metagenomics and Environmental Genomics

There exists a vast network of interactions between individual genomes and other biological and environmental systems. Each of us walks a slightly different path through life, experiencing different influences upon our physiology in terms of the food we eat, the amount of sun we expose ourselves to, the environments we experience in our jobs, the pathogens that we by chance encounter, etc. Any and all of these and more may be subtly changing (mutating) letters in our genome or periodically influencing gene regulation or expression. Metagenomics is the field of genomics that investigates these interactions and their effects.

Obviously, the human genome is highly complex. Add to that all the genomes in the environment with which one may come into contact, and the enormity of the subject comes into focus. Large-scale sequencing programs are, however, providing a rich resource of data for scientists to mine in metagenomic studies.

Conclusion

The fields of human genetics and genomics are having profound positive impacts not only in terms of biomedical discovery, but also in terms of the clinical practice of medicine—working to improve the lives for millions of patients and demonstrating great promise for future highly positive contributions to human health and well-being worldwide.

As the eight functional domains for human health application of genetics and genomics illustrate, this field of science (and the expanding industry associated with it) generates a profound impact on biomedical research and the practice of clinical healthcare. In addition to applications in human medicine and wellness, there are also several non-medical human applications of genetics and genomics, including forensic science, anthropology and genealogy, evolutionary biology, and paternity testing. These are also highlighted in the report.

It is readily evident that, as fundamental genomic knowledge has expanded, the enhanced understanding of genetic mechanisms generated, in concert with access to rich whole exome and genome datasets (and associated reference compendia of human gene variants), has opened the door to a new era of discovery and progress in medicine. The impacts of these advancements are now increasingly reverberating across medicine, a fact highlighted by Eric Green, the Director of the National Human Genome Research Institute (NHGRI), and colleagues who note that:"

> *With 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, opportunities beyond their initial expectations, with many more anticipated in the next decade.*[7]

While generating these positive functional impacts is the raison d'etre for pursuing the advancement of human genetics and genomics, it has also had the very positive spillover effect of building a powerful science- and technology-based economic sector for the U.S.—a sector that supports over 850,000 jobs across the nation and generates a $265 billion economic impact in terms of U.S. economic output. The continued innovation in human genetics and genomics is expanding the stock of knowledge upon which our continued advancement depends and shows great promise to continue to do so long into the future. The field represents a particularly strong example of how investing in fundamental and applied science generates robust economic, social, and individual benefits for humankind.

7 Green, Eric D., et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 Oct. 2020.

I. Introduction

A. Science as a Driver of Economic and Social Advancement

Scientific research is of high importance not only because it reveals fundamental truths but also because it increases the stock of human knowledge upon which economic, societal, and technological progress depends. The U.S. economy, in particular, has advanced on the back of scientific progress that has enabled national leadership in diverse industries such as aerospace, energy, agriculture, transportation, materials, digital technology, communications, and healthcare.

In the past two decades, scientific advancements have seen new industries advanced, typically with the U.S. being an early innovator in commercial applications of both science and associated technologies. Developments in physics, chemistry, biology, and mathematical and computational sciences have paved the way for new industries in nanotechnology and advanced materials, renewable energy, AI-powered autonomous systems, biotechnology, and genetic engineering. There has also been an observable trend of "convergence," whereby multiple science and technological disciplines combine to advance new opportunities. Advancements in computational and digital analytics converging with large-scale data from other sciences have helped accelerate this trend.

What has become clear is that continuing to strengthen the competitiveness of the U.S. economy requires ongoing expansion of national innovation capacity and the scientific and technological research and development (R&D) upon which that innovation

From Fundamental to Applied Research

Public/Private and Market/ Non-Market Returns

Scientific research may produce both private and social returns. Research that leads to improved knowledge, national security, public health, enhanced food security, etc. provides social (public) returns. Research generating a patented technology or improving the productivity of a production process provides private returns to those inventing and using the research-based tool or knowledge for commerce. Often, research leads to both forms of return at the same time. For example, the development of a vaccine for an infectious disease provides private monetary returns to the vaccine developer and social returns through enhanced public health. Both types of returns motivate investment in research. Research finds that the rate of social return on R&D exceeds the rate of private return (although both are strong). Despite the large U.S. national investment in research, analysis shows that optimal investment in research would be more than four times actual investment.

The robust levels of social return on research underpin the core rationale for public investment in research. Indeed, without public support for research, a wide-ranging and important suite of research topics would go unaddressed. Because so much important research is non-market (focused on phenomena or subject matter without an immediate line-of-sight to a market application) or focused on the generation of social returns (benefiting society overall, but perhaps not able to realize private returns to investment), the public sector plays a critically important funding function in the R&D ecosystem. Further, private sector investment in basic science is relatively scarce because of the speculative nature of early fundamental research, the long time-horizons involved in the performance of basic inquiry, the risk of experiment failures, and, most importantly, the lack of immediate line-of-sight to market.

What is critically important to understand is that the U.S. economy, and the innovations and technologies upon which it depends, is built upon a bedrock of fundamental scientific advancements and an overlying strata of applied and translational discoveries that leverage fundamental knowledge.

depends. This requires funding for research and the institutions that perform research, together with funding for the education of the scientists, technologists, engineers, mathematicians, and other skilled intellectual talent that innovates and produces the products and services that result from innovation.

B. Funding and Supporting American Science

The federal government has been and continues to be an essential component of the U.S. research ecosystem, funding research performed at universities, independent research institutions, and other organizations and conducting research within national institutes and laboratories. Federal funding for research has been especially important in supporting fundamental science, which in turn represents the platform upon which applied research may advance. Federal funding also plays an important role in translating research into early-stage commercialization, through translational research funding and dedicated early-stage venture support through the federal Small Business Innovation Research (SBIR) and Small Business Technology Transfer (STTR) programs.

Support for scientific research in the U.S. comes from both public and private sources. As noted in the prior sidebar textbox, public resources focus primarily on supporting basic through applied research while private (commercial) funding focuses primarily on powering applied and translational research that advances innovations into commercial application. In total, the U.S. R&D enterprise spent $596.6 billion in 2019, representing 2.84% of GDP.[8] However, the extent to which the economy is driven by and built upon the innovative output of R&D makes the impact of R&D on overall GDP many times larger and a dominant factor in U.S. economic success. The economic future of the U.S. hinges upon research and development.

C. Human Life Sciences

Scientific research produces many benefits and returns for society, but perhaps none are as important as the preservation and extension of human life itself. Modern life sciences and associated advancements in biopharmaceuticals, diagnostics, medical devices, and healthcare services have enabled unprecedented improvements in human health and longevity. Today, the average life expectancy for a newborn female and male in the U.S. is 81.4 years and 76.3 years, respectively.[9] In 1950, those same metrics were 71.1 and 65.6 years.[10] Average lifespans have expanded by more than a decade in less than two generations through advancements in health, hygiene, and safety.

> “Today, a powerful argument can be made for substantially increased investment in research in the United States. Equally great is the need to train the next generation of scientists and citizens for what will be a very different world.”
>
> **National Academy of Sciences. 2021. “The Endless Frontier. The Next 75 Years in Science.”**

8 “2020 Global R&D Funding Forecast”, *R&D World*, February 2020, WTWH Media, LLC.
9 Kochanek, Kenneth D., et al. “Mortality in the United States, 2019.” NCHS Data Brief, No. 395, Dec. 2020.
10 Ibid.

Investing in People and Infrastructure

Science relies on an educated base of scientists who make discoveries and is also advanced through the development of new tools and technologies that power experiments and enable new insights into biological processes. For example, advancements in ultra-high resolution imaging technology, functional imaging of real-time processes, analytical chemistry instruments, computational systems, and gene sequencing equipment have facilitated leaps forward in fundamental and applied scientific insights.

The impact and importance of human life science advancement have come sharply into focus during the COVID-19 pandemic, where R&D-based innovation of diagnostic tests, advanced therapeutics, and the rapid development of vaccines has proven crucial in forging a path for a return to normal life and commerce. As noted in a recent report: "the COVID-19 crisis has vividly illustrated the critical importance of life science research and innovation systems and the ecosystems that support the advancement of innovations through commercial deployment to address health needs."[11]

Scientific advancements in biology and medicine contribute to our daily lives and are also unveiling the incredibly complex physical mechanisms and human/environment interactions that govern our development and health. The Russian doll analogy appears to hold, with the life sciences uncovering level upon level of complexity and interrelationships in biological structures, mechanisms of influence, and associated health outcomes. While unveiled complexity may confound easy solutions to questions and problems, it also provides an expanding universe of potential for discoveries, applications, and functional possibilities.

Perhaps nowhere has life science research advanced more in the modern age than through insights provided by genetics and genomics. This field is both fundamental in biological research—elucidating the basic code of life, DNA, upon which our form and function depend—and in enabling applied and translational discoveries across most diseases and health disorders.

D. The Genome—Coding Life

Humans are complex. That is true on many levels, and it is certainly true in terms of our biology. The more biologists learn of our biological structure and function, the more complex and intricate the machinery of our biology is revealed to be.

We each comprise approximately 30 trillion individual cells and over 200 different cell types. Our development and ongoing biological function are orchestrated by our DNA (deoxyribonucleic acid), a linear molecule arranged in a double helix (a spiraling ladder) comprising linked base pairs of the nucleotides adenine (A) and cytosine (T), and guanine (G) and thymine (T). Our DNA contains six billion base pairs of these nucleotides, the sum of which is called

11 Tripp, Simon, et al. *Response and Resilience: Lessons Learned from Global Life Sciences Ecosystems in the COVID-19 Pandemic.* TEConomy Partners, LLC for Pfizer, Inc., Jan. 2021.

our “genome,” which is effectively the governing instruction set and a regulatory “code” for our bodies.

The six billion base pairs of our genome are organized into 46 chromosomes (23 inherited from our mother and 23 from our father). A chromosome consists of a long section of DNA containing up to 500 million base pairs of DNA with thousands of genes. An individual “gene” is defined as a grouping of base pairs that together perform a function, encoding the synthesis of a gene product (either RNA or a protein). We each have approximately 20,000 protein-coding genes, which comprise circa 2% of our genome.

We also are each host to trillions of microorganisms found in our gut and other parts of our anatomy, many of which play an important symbiotic role for us in digestion and immune system function. Each of these microorganisms has its own DNA. Collectively, all of those microorganisms and their DNA comprise our “microbiome.” Research in epigenetics has found that microbiome structure can influence and impact human gene expression and regulation.

As genetics and genomics developed as scientific disciplines, for a long time the standard hypothesis was that the part of our genome that “mattered” is the 2% comprising our protein-coding genes, because proteins are the functional biomolecules performing the vast majority of biological activities. It was common to consider the 98% majority of our DNA as “junk DNA”, legacy base pairs accumulated over the huge span of time in our evolution, but no longer relevant or “functional” to our development and predominant biological functioning. We now know that is not the case.

After the publishing of the draft human genome by the Human Genome Project and Celera, an international team of 442 scientists from 32 institutions embarked on a large-scale team research project called ENCODE (the ENCyclopedia of DNA Elements), which used leading edge approaches to measure biochemical activity across the entire human genome, not just protein coding genes. ENCODE revealed that the non-protein coding regions are far from being “junk” and primarily contain DNA with an active biochemistry, even if the preliminary findings did not elucidate the function of that activity. **The results indicate a far more complex and multifaceted functionality to our genome than previously thought.**[12]

Complicating matters further is that while the genome codes for our fundamental life processes it is not deterministic. Other factors also influence our biology and health, including a multitude of interactions with external environmental factors and stimuli, such as the foods we consume and the microbial communities that inhabit us over our life journey.

We are all the same species, homo sapiens, but we are all different. None of us have exactly the same genome sequence nor the same environmental interactions. A large number of the diseases and health disorders we will face across our lifespan will be similarly diverse and complex, many being found to be associated with multiple genes and gene variants in combination with wide-ranging external factors—for example, prior infection with pathogens or differential exposure to mutagenic factors such as chemicals or radiation (causing genotoxic injury).

Numerous diseases and health disorders result from single gene changes (known as monogenic or Mendelian diseases), where a variant within a single gene may code the wrong protein or a dysfunctional gene product. However, many diseases with genetic engagement involve multiple genes and interactions between various genes and additional factors. Many diseases, including most of

12 Parrington, John. *The Deeper Genome: Why There Is More to the Human Genome than Meets the Eye.* Oxford University Press, 2015.

our large-scale chronic diseases, turn out to be a complex soup of genetic, epigenetic, and environmental factors interfacing with one another.

As individuals, our incredible complexity, and the biologically influential environmental factors we each uniquely encounter across our lives, explain why the development of diagnostics and drugs is such an intense, difficult, and expensive challenge. Finding generic solutions to individually variable disease causations and expressions is no small task. It is further complicated by the fact that genetic, and other factors, can influence how we each respond to and metabolize a drug.

Biomarkers are one of the pathways by which the complexity problem is resolved. At the highest level, a "biomarker" is a measurable substance in an organism whose presence is indicative of some phenomenon such as disease, infection, or environmental exposure. When a biomarker is discovered related to a disease, it provides a target for further research, a potential measure to be used in achieving diagnosis, and, if found to be "druggable," a target for a therapeutic. In 1998, the National Institutes of Health Biomarkers Definitions Working Group defined a biomarker as "a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention."[13] Genetics and genomics advancements, including technological advancements such as genome sequencing, have provided an important modern pathway for identifying genetic biomarkers, including diagnostic and therapeutic targets. As extensive collections of whole-genome or partial genome sequences build, scientists can mine these sequences to identify genes and gene variants that stand out differently in people with a disease or condition of interest. These identified genes and variants can then be studied to identify their gene products and the biochemistry involved in their regulation and expression.

E. Human Genetics and Genomics—Applications and Impacts

The applications and impacts of genetics and genomics in human biology and medicine have grown in parallel with advancements in gene sequencing technologies and digital analytics platforms for deriving meaning within the resulting large sequence datasets. The Human Genome Project cost approximately $2.7 billion resulting in publishing of the reference genome.[14] The Human Genome Project and subsequent genomics initiatives sparked the advancement of commercial sequencing instruments and processes that have dramatically increased the speed of sequencing and decreased the price of each sequence. In March 2021, Nebula Genetics was providing whole genome sequencing (WGS)[15] with 30x resolution (meaning that each position is read 30 times to enhance accuracy) for less than $300, a nearly 10,000 fold decrease in price versus the first sequenced human genomes.[16] NHGRI tracks costs associated with DNA sequencing performed at the sequencing centers funded by the Institute, and the most recent NHGRI data (August 2020) place the cost per genome at $689.[17] The pace of advancement in gene sequencing has exceeded even the much-vaunted pace of Moore's Law in computer processors.

13 Strimbu, Kyle, and Jorge A Tavel. "What are Biomarkers?" *Current Opinion in HIV and AIDS* vol. 5,6 (2010): 463-6. doi:10.1097/COH.0b013e32833ed177

14 National Human Genome Research Institute. "The Cost of Sequencing a Human Genome." www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost. Accessed 12 May 2021.

15 Whole Genome Sequencing (WGS) is a term used to describe sequencing that at present does not provide full coverage of the entire genome, where repeat sections in the genome still remain a challenge to resolve. As defined by the National Cancer Institute, whole genome sequencing is a laboratory process that is used to determine nearly all of the approximately 3 billion nucleotides of an individual's complete DNA sequence, including non-coding sequence.

16 Nebula Genomics. "Did You Know that most DNA Tests Decode Only 0.02% of Your DNA?" www.nebula.org/whole-genome-sequencing-dna-test/. Accessed 12 May 2021.

17 National Human Genome Research Institute. "DNA Sequencing Costs: Data." www.genome.gov/about-genomics/fact-sheets/DNA-Sequencing-Costs-Data. Accessed 30 April 2021.

Genomics in the Mainstream of Human Biological Research

The movement of genetics and genomics from a niche in biomedical science into the mainstream underpinning research across most biomedical disciplines is evident in the expansion of genetics and genomics content across the research sphere funded by the National Institutes of Health (NIH).

In 1990, greater than 95% of human genomics research funding flowed, in a concentrated way, through the National Human Genome Research Institute (NHGRI) at NIH. By 2020, the vast majority of human genomics research funding is provided through twenty other individual NIH institutes, indicative of the relevance of genomics across almost every domain of medical science and human life science research.

The current speed and price of sequencing a human's genome has led to the sequencing of patients increasingly becoming a clinical reality in modern healthcare systems. Barriers are less a factor of sequencing cost, and instead relate more to building the capacity needed to analyze the huge volume of data generated, to interpret the meaning of that genetic code for the individual patient, and counsel the patient as to implications for their health.

A developmental tipping point has been achieved in which the utility of genomics and wide-spread use of sequencing is clearly advantageous for significantly enhancing human health outcomes. As discussed below, and in further detail within Chapter III, the functional application of human genetics and genomics to clinical healthcare is now a daily reality in some medical fields (e.g., cancer diagnosis and treatment) and is increasingly front-and-center in neurological, psychiatric, gastrointestinal, immunologic, rheumatologic, dermatologic, pain management, and other application areas of clinical medicine. It is also fundamental to advancements being made in the diagnosis and treatment of a wide range of rare diseases and disorders—helping to end the diagnostic odysseys of millions of patients afflicted with rare diseases that have been difficult to diagnose and sparsely served in terms of available treatments.

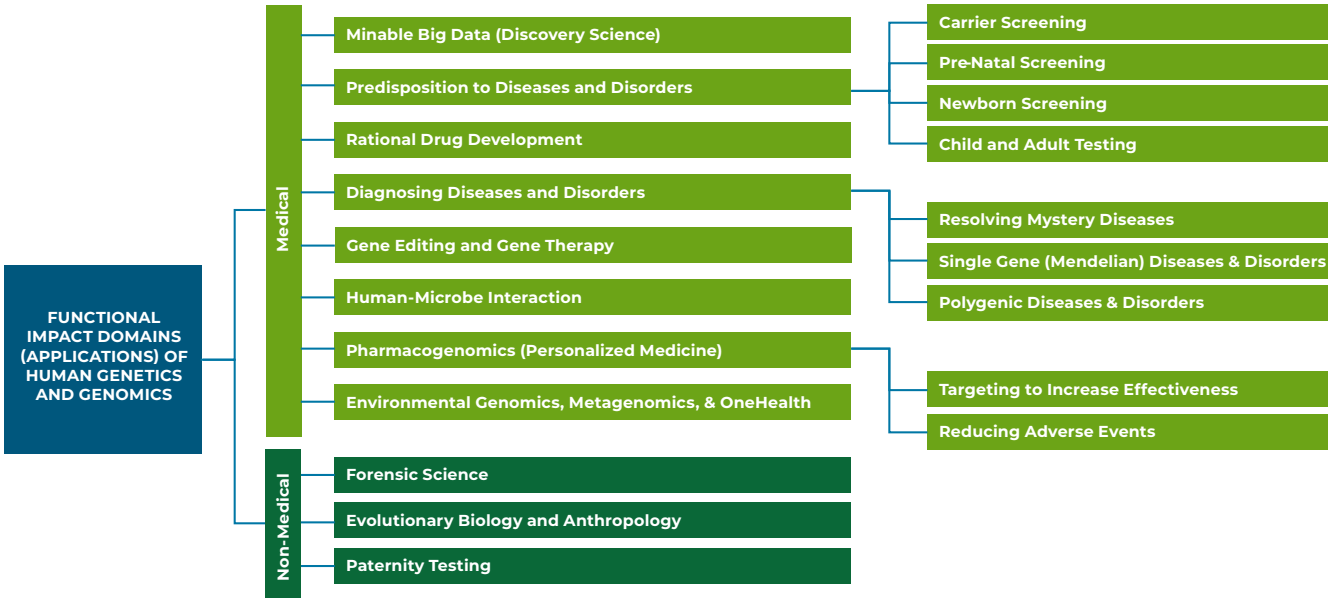
TEConomy has divided this study of the impacts of human genetics and genomics into two macro classes of impacts: **economic impacts** (examining the impact of human genetics and genomics on the U.S. economy) and the **functional** (application) impact domains in which genetics and genomics are **affecting human health and the clinical practice of medicine.**

1. Economic Impacts

The performance of research, the development and manufacturing of commercial genomics technology platforms, the multitude of diagnostics products and therapeutics on the market that is derived from genomics knowledge, and the associated healthcare services provided generate economic activity and support a large volume of jobs across the nation—these are economic impacts, that positively ripple through the U.S. economy. Other economic impacts are associated with lives saved, lives improved, and impacts on people who would otherwise have to be caregivers to loved ones. It is also the case that the application of genetics and genomics to individual healthcare needs costs money, and there is considerable complexity in assessing the comparative costs and benefits of one therapy versus another, or whole-genome sequencing (and what it may elucidate) versus other diagnostic tools and technologies. Because genetics and genomics are such a rapidly advancing field of application and technology in healthcare, cost/benefit equations are continually changing. No health insurer was going to pay $3 billion to sequence a whole genome for a patient, but $689 (less than the average cost of a single MRI)[18] to generate a dataset with lifelong, and increasingly expanding utility for the patient opens up a whole new ballgame.

18 Note: the average cost of an MRI in the U.S. is $1,325. Source: Vanvuren, Christina. "What Can Affect the Cost of an MRI?" *New Choice Health, Inc.*, www.newchoicehealth.com/mri/cost. Accessed 12 May 2021.

Figure 1: Current Functional Impact Domains (Applications) of Human Genetics and Genomics



Source: TEConomy Partners, LLC.

2. Functional Applications

The functional impacts of human genetics and genomics are the impacts resulting from discoveries via the advancement of research and the clinical applications of genomics to benefit human health. The decoding of the human genome was a signature inflection point for science and has been widely acknowledged as a towering achievement for life sciences. What it sparked, however, has been an ongoing expanding universe of advancement in genetics and genomics and incredibly wide-ranging elucidation of biological processes, systems, and outcomes with genetic connectivity that is now fundamental to advancing human health and clinical medicine.

In reviewing applications of human genetics and genomics, functional impacts are generated within eight major domains of activities impacting human health. These are illustrated in Figure 1, and this structure forms the foundation for discussion of functional impacts of human genetics and genomics contained in Chapter III. The study also briefly describes non-medical applications of human genetics and genomics.

F. Purpose of the Study

This study seeks to provide an accessible reporting of the positive impacts for the economy, society, and individual health that are derived from modern human genetics and genomics science and the associated commercial and clinical application of advancements.

In 2011, the authors of this report (while at Battelle Memorial Institute) conducted a detailed impact analysis of the Human Genome Project. The resulting Battelle report[19] highlighted the growth of an emerging industry in the application of genomics that was born from the scientific and technological

19 Tripp, Simon, and Martin Grueber. *Economic Impact of the Human Genome Project.* Battelle Memorial Institute, May 2011.

momentum driven by the Human Genome Project and Celera's work to sequence a reference human genome. It also highlighted some of the early applications occurring through genetics and genomics advancements in healthcare and other life science-related challenges and needs. The 2011 report played a role in highlighting not only the important scientific impacts of federal government investment in "big science" and genomics in particular, but also demonstrated the robust return on public investment that had occurred through economic growth in genomics technologies and emerging applied genomics application domains. The report continues to be linked on the website of NHGRI at genome.gov.[20]

Now, 20 years after the publication of the draft sequence, there has been a significant and large-scale expansion of the human genetics and genomics universe. Technologies for sequencing and for genome analysis have advanced significantly—to the extent that whole-genome sequencing is quite affordable (certainly in comparison to many common medical procedures and tests) and an entire genome may be sequenced in less than one day.[21] Advanced analytics and artificial intelligence systems are now available that can simplify deriving actionable insights from the sequencing data. Research discoveries in human genetics and genomics have compounded, building upon one another in a virtuous network of expanding information, knowledge, and application. Today, this expansion has brought human genetics and genomics to a rather visible inflection point. The speed and affordability of gene sequencing, in combination with deep insights into genomics and -omic sciences more broadly, together with advancing biopharmaceutical, diagnostics, and other medical technologies that can leverage genomic information, have now made genomics a part of the clinical practice of medicine across many medical specializations and medical conditions.

This study seeks to provide a generalized overview of human genetics and genomics achievements and scan the current status of human genetics and genomics in answering human health questions and advancing clinical applications. The study also seeks to highlight the economic contribution of the expanding genetics and genomics sector. The U.S. has been a pioneer in genomic sciences, leveraging both public and private sector investment to build an advanced industry that provides economic expansion and opportunities for further growth while

20 Ibid.
21 Genomics England. "Sequencing a Genome." www.genomicsengland.co.uk/understanding-genomics/genome-sequencing/#:~:text=One%20human%20genome%20can%20be,pieces %2C%20around%20150%20letters%20long. Accessed 12 May 2021.

at the same time advancing human health and well-being. This report characterizes those positive impacts using quantitative analytics in conjunction with a qualitative description of identified functional impact domains and associated benefits.

It is anticipated that this report will be useful to public policymakers and those seeking an understanding of the public and private, and monetary and non-monetary, returns to investments in science broadly, and specific to genomics. It may also be useful for those seeking to gain a broad introduction into the multi-faceted ways genetics and genomics are being used to improve human health and clinical health outcomes—helping to illustrate the power of a rapidly expanding biomedical field that is poised to advance and transform many avenues of clinical medicine. It is also hoped that the reader will be encouraged by the promise of genomic medicine and the evident hope it provides for improved health outcomes for humanity.

G. Limitations of the Study

The economic analysis deployed in this study provides a one-year, point-in-time quantification of the nation's genetics and genomics sector. One of the challenges in measuring genetics and genomics impacts in the economy is that the U.S. industry classification system does not contain a NAICS[22] code that covers the industry specifically. Instead, it is a partial component of many different industry sectors. Without having access to data through NAICS codes, establishing a baseline measure of the genetics and genomics industry in aggregate within the U.S. economy requires the development of a custom dataset (comprising data on individual establishments and companies engaged in human genetics and genomics research, technology development, and application) that effectively builds the data from the ground-up, establishment by establishment (rather than relying on generally available government summary sectoral statistics).

One of the principal challenges in developing establishment-level data in genetics and genomics is that while for some enterprises or organizations genetics and genomics comprises the preponderance of their business or institutional work (for example, gene sequencer manufacturers, genetic testing companies, genetic counselors), for many active in the sector, it is only part of their business or work (for example drug companies, large national diagnostic laboratories, clinical care providers, etc.). Informed estimations have to be made of the portion of revenues, expenditures, and employment at organizations that are related to the application of genetics and genomics. The fact that estimates must be used this way to build the overall dataset that drives the direct effect in the input-output models used is a limitation (discussed further in Chapter II). The study has endeavored to err on the side of being conservative in developing portioning estimates, and thus the resulting measures of impact are likely low rather than high.

The examination of functional impacts is, in many respects, an even greater challenge. Part of this is evident in the very large volume of academic and industry life science research studies in which genetics and genomics are a component or focus. As a fast-moving field, there is ongoing evolution and expansion in the applications of genomics in human healthcare, and it is a significant challenge to do justice to such a wide-ranging field. With thousands of diseases, many hundreds of drugs, and a broad compendium of diagnostics tests having genetic associations, providing full coverage of every application of genomics in healthcare would be an extremely challenging task and out-of-date immediately upon completion, not to mention a rather daunting read. This is not attempted in this study, but rather the functional impact assessment herein works to classify human genetics and genomics advancements by broad application domain (disease diagnostics, pharmacogenomics, gene therapy, etc.) and then uses specific narrative examples of genetics and genomics in action within these domains (together with some measures indicative of scale where readily available). **As such, the functional impact section of this report should be viewed as providing**

22 NAICS is an abbreviation for the North American Industrial Classification System.

an overview of the broad areas in which genetics and genomics are providing benefits to human health, not a formal quantification of these impacts.

This report is also limited in that it focuses on human genetics and genomics only, and this certainly undercounts the wide economic and societal benefits that accrue to advancements in genetics and genomics more broadly. While the application of genomics to medicine is certainly an important area of use, the ubiquity of DNA as the basis for all life on Earth means that genomics finds application in many more fields of science and commerce. Both the science of genomics and the tools and technologies of genomics find application in multiple additional endeavors, including:

- Veterinary medicine
- Agriculture (in applications such as crop improvement, crop protection, animal science, and nutrition)
- Industrial biotechnology (in applications using microbes with engineered genomes to produce biochemical products), and
- Environmental and ecological services (in applications using engineered microbes in industrial waste cleanup, wastewater treatment, and other applications).

These additional areas of genetics and genomics application have significant impacts that are not addressed in this report.

This report is primarily intended to highlight the current status of human genetics and genomics impacts, but it does contain a chapter that briefly discusses the frontiers and potential future advancements that may occur in the foreseeable future. The discussion in that chapter is, of course, speculative, and thus subject to the usual limitations involved when looking towards an uncertain future.

Readers should consider the benefits of genomics from both the economic and functional perspectives, not just one or the other. Examining the field through the hard lens of economics must be tempered by the fact that much that matters in life may not be readily broken down into dollars and cents. The alleviation of pain and discomfort from a medical condition, the ending of a diagnostic odyssey of a patient with a mystery disease, or the lifelong experiences of a child and their parent made possible through that child being effectively treated for a rare genetic condition, are principally humanitarian benefits. The other side of the coin is that novel genetic tests, customized medicines, and highly specialized care can come at a significant cost to individuals and those who pay for healthcare, and it is important to understand the economic implications of emerging clinical frontiers.

In this regard, we caution that because medical genomics is an emerging field, there is relatively sparse literature on the monetary impacts and costs/benefits associated with genomic medicine. It is anticipated that genomic medicine may increase costs in many of its early applications, but these early applications are part of a path that will lead to overall cost savings as medicines are targeted and used more effectively, chronic diseases better managed (or perhaps cured), adverse reactions to medications curbed, and the costs of lifelong care potentially avoided by addressing genetic components in diseases that may be attended to through gene therapies and effective personalized therapeutics.

Chapter II provides an assessment of the economic impact of sectors engaged in human genetics and genomics R&D, the provision of genetics and genomics tools, technologies, and services, and associated economic activity.

Chapter III provides discussion of the multifaceted functional impacts allocable to the key application domains of human genetics and genomics in healthcare. This follows the structure shown previously in Figure 1, and also briefly touches upon some of the non-medical applications of human genetics and genomics.

Chapter IV looks to the future of human genetics and genomics and introduces some of the factors that need to be addressed to increase the positive impacts of the sector.

II. The Economic Impact of Human Genetics and Genomics

A. Measuring Human Genetics and Genomics Economic Impacts

As described in the previous chapter, the development of the economic (expenditure) impacts within this study is focused on and limited to estimating those impacts stemming from the use of genetics and genomics for human biomedical purposes.

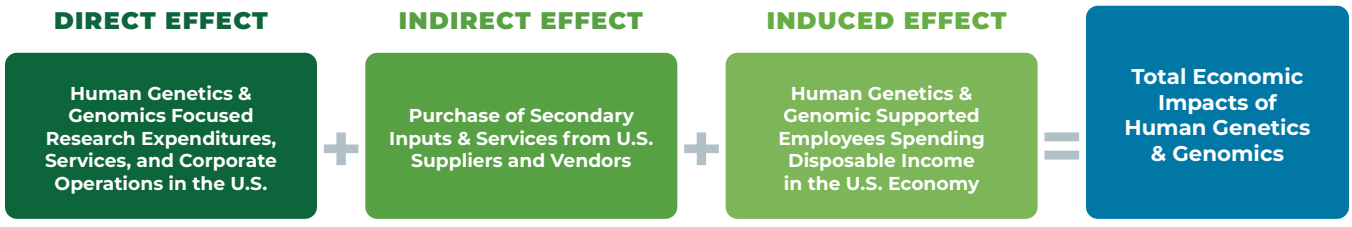
1. Basics of Impact Modeling

The estimation of economic impact makes use of an input-output (I-O) model to represent the interrelationships among economic actors and sectors. Within these models, the flow of commodities (or the value they represent) between industries, consumers, and institutions in an economy are modeled through a matrix representation allowing for the impact of changes in employment, expenditures, or economic output in one sector of the economy to be projected onto other sectors of the economy. These transactions continue under the premise that every dollar spent in the economy (the direct effect) is re-spent on the purchase of additional inputs, goods, or services generating additional economic activity and impacts. I-O analysis is the generally accepted "gold standard" in economic impact measurement and examines the relationships among economic sectors (e.g., institutions or industries) and final consumers.

For the purposes of this study, the I-O analysis models the "ripple effect" that originates from the expenditures made for human genetics and genomics research and direct company operations in the economy, flows through suppliers and vendors as additional inputs are purchased, and through employees, faculty, staff, and related supplier workers who spend their wages in the U.S. economy (Figure 2).

For modeling and estimating expenditure impacts, TEConomy used a 2019 professional IMPLAN I-O economic impact model of the U.S. Originally developed in 1976 by the U.S. Forest Service, IMPLAN

Figure 2: Measuring Economic Impacts of Human Genetics and Genomics



Source: TEConomy Partners, LLC.

Table 1: Impact Measures Included in the Analysis

Impact Measure	Definition
Output	Also known as production, sales, or business volume, is the total value of goods and services produced in the economy. For public/non-profit entities, such as universities, expenditures are often the truest measure of this economic activity.
Employment	The total number of jobs created; Includes the direct jobs paid for through salary and benefit expenditures and indirect/induced jobs generated through purchase expenditures.
Labor Income	Also known as total compensation, is the total amount of income, including salaries, wages and benefits, received by employees, proprietors, and other supplier workers in the economy;
Value Added	The difference between an industry's total output and the cost of its intermediate inputs; sometimes referred to as the industry's "Contribution to GDP".
Federal and State/Local Government Tax Revenues	Includes the estimated revenues to federal and state/local governments from all sources as a result of the impacts estimated.

is now the most widely used economic impact modeling data and analysis tool in the nation.[23]

IMPLAN models are built upon underlying federal information including the U.S. Bureau of Economic Analysis (BEA) national accounts data, supplemented with state level employment data from the U.S. Bureau of Labor Statistics (BLS) and other economic data from the U.S. Bureau of the Census. Currently, the IMPLAN model reflects and represents 546 sectors of the U.S. economy. The core data and structures developed within an IMPLAN model can be used to analyze the economic impacts of institutions, projects, or entire industries. Employment and expenditure data developed and estimated as part of this study provide the direct impacts to drive the overall economic impact models and estimations. Ultimately, the impact model generates estimates of the additional indirect and induced impacts (also known as the multiplier effects) for employment, personal income, value added, output, and federal and state/local tax revenues.

2. Prior Human Genome Project Economic Impact Results

For context both in terms of size and breadth, values from the authors' prior work on the *Economic Impact of the Human Genome Project* are provided in Table 2. These values reflect the full breadth of genetic and genomic research and the nascent involvement of industry as of 2010.

By comparison, the current study, to be described and discussed in the following pages, focuses solely on the human genetics and genomics domain for a period one decade later, 2019.

23 Note, the authors followed the same I-O methodology by using a prior year's IMPLAN model to estimate the economic impacts of the Human Genome Project.

Table 2. Core Metrics from Economic Impact of HGP Study

Impact	Employment (Jobs)	Output ($M)
Direct Effect	51,655	$22,627.5
Total Impact	310,360	$67,146.0

Source: Economic Impact of the Human Genome Project, Battelle Memorial Institute, May 2011.

B. Drivers of Economic Impacts: Methodology and Assumptions

Three types of economic inputs or "drivers" are used to develop and estimate the impacts of human genetics and genomics—research expenditures, core industry employment, and extended industry sales or employment. For this analysis, 2019 data was captured to the extent practicable using the fiscal year or calendar year. The following specifies the conservative methodology and assumptions used in developing these economic impact drivers.

1. Research Drivers—Investments in Research

A significant driver of human genetics and genomics economic impacts comes in the form of focused research funding from federal agencies and other non-profit organizations. This section describes and captures this funding mechanism for use in driving the economic impact model. It captures the value of human genetics and genomics research funded by these organizations and ultimately performed by universities, research institutes, federal agencies, and other non-profit research organizations.

National Institutes of Health

With the focus of this study specifically on human genetics and genomics impacts, the National Institutes of Health (NIH) becomes the principal governmental agency funding research in this domain. However, even within NIH, the role of human genetics and genomics in its research portfolio is subject to interpretation and perspective—ranging from fundamental research into human genetics and genomics to the use of genetics and genomics as an enabling "tool" in research focused on specific diseases or conditions.

To account for this, TEConomy established three funding scenarios for NIH, each differentiated by size and breadth of funding, developed through an analytical approach that eliminates double counting of funded research efforts and builds upon the previous scenario. The structure of these scenarios and the research funding that is captured within each scenario is built using internal NIH classifications and keywords.[24] These scenarios include:

- **Core NIH Human Genetics and Genomics Funding**
  - Includes all research funding from NHGRI (524 projects, $416.3 million)
  - Research reviewed and approved by specific genetic and genomic-related proposal review study sections (e.g., Genetics of Health and Disease Study Section and others) (1,434 projects, $611.1 million)

24 For the purposes of this study "NIH research funding" was limited to research-oriented grants or contracts (external and intramural) to U.S. researchers. Grants to non-U.S. research performers, U.S. firms (via SBIR/STTR awards), or construction and training awards were excluded from the analysis. For FY 2019, this NIH total research funding reached $28.939 billion. Note: grants to U.S. firms were excluded as individual firms are typically captured in the core industry drivers analysis.

- Research classified in specific and related NIH Research, Condition, and Disease Categories (RCDC) EXCEPT the broadest Genetics category (e.g., Cancer Genomics, Gene Therapy, Gene Therapy Clinical Trials, Genetic Testing, and Human Genome) (5,424 projects, $2,107.3 million)

- **Core + Additional Expanded NIH Funding**
  - Additional research within the specific Genetics RCDC funding category, not captured above (10,493 projects, $3.884 billion)

- **Core + Additional Expanded + Additional Use as Tool Funding**
  - Additional research listing genetics or genomics as a Principal Investigator-provided keyword in the funding information not captured above. A review of these awards showed that most were non-genetics and genomics focused yet were using genetics or genomics as a key analytical approach or tool enabling the research. (13,448 projects; $7.184 billion)

Table 3 shows the value and incremental additions to each scenario and the share of total NIH research funding.

A striking finding from this assessment of NIH research funding is the pervasive nature of genetics and genomics in human biomedical research. Nearly half of all NIH research funding specifies some connections with human genetics and genomics, at least as an investigative tool to support other research.

To maintain a conservative perspective on the role NIH research funding plays in the overall economic impact of human genetics and genomics, the remainder of this chapter focuses on the impacts generated with the inclusion of the smallest, core funding scenario. Economic impact results estimated using the additional more expansive NIH research funding scenarios are included in the Appendix.

Other Federal Agencies

Beyond NIH, other federal agencies provide significant human genetics and genomics-related research funding.

Department of Veterans Affairs

In FY 2019, the Department of Veterans Affairs (VA) Office of Research and Development invested a total of $107.0 million in genomics research and infrastructure. Of this, $26.0 million was towards funding genomics research projects, $44.0 million towards the

Table 3: NIH Research Funding Captured by Each Scenario

NIH Human Genetics and Genomics Funding Scenario	2019 NIH Research Investment Component ($B)	2019 NIH Research Scenario Value ($B)	Scenario Cumulative Share of 2019 NIH Research Funding
Core Funding	$3.135	$3.135	10.8%
Core + Additional Expanded Funding	$3.884	$7.018	24.2%
Core + Additional Expanded + Additional Use as Tool Funding	$7.184	$14.202	49.1%

Source: TEConomy analysis of NIH research awards using the RePORT website's ExPORTER Project file for FY 2019

Million Veteran Program (MVP) core infrastructure for recruitment, enrollment as well as clinical and genomic data curation, and $37.0 million towards genotyping and whole genome sequencing DNA samples from the MVP. With over 830,000 Veterans enrolled to date, MVP is one of the largest healthcare system-based genomic cohorts in the world.[25]

National Science Foundation

The National Science Foundation (NSF) funds substantial research efforts within the broad context of genetics and genomics through research programs in genetic mechanisms, phylogenetic systematics, and the large-scale Plant Genome Research Project. For the purposes of this study, NSF awards that were active at any time within FY 2019 were considered and examined to find those that met the requirement of funding research primarily aimed at human genetics and genomics understanding. Included in these awards are significant research funding for big data and/or bioinformatics research that was primarily aimed at facilitating or enhancing the ability to manage and analyze human genetics and genomics data. The included research reflects 212 awards funded at $19.3 million. [26]

Other Health and Human Services (HHS)

Undoubtedly, other HHS agencies beyond NIH include some level of human genetics and genomics research and/or research funding. However, given the limited detailed information upon which to assess the connections to human genetics and genomics, only within the National Institute for Occupational Safety and Health (NIOSH) was this information as well as funding information available. For FY 2019, TEConomy identified nearly $384,000 in human genetics and genomics-related funding.

Voluntary Health Associations and Other Non-Profit Funding

An important funding stream for human genetics and genomics research is provided by a wide variety of voluntary health associations, patient groups, and other non-profit funders. Likely included within this set of funders are many of the members of National Organization for Rare Disorders that are funding research efforts to understand the genetic and genomic traits of these rare disorders.

The difficulty in including funding from these associations and non-profits is the limited information on human genetics and genomics research within

25 For additional information on the Million Veteran Program see: www.mvp.va.gov.
26 Multi-year awards that extended into FY 2019 were weighted by the number of FY 2019 days as a share of the total project's estimated duration.

their specific research grants and the requirement to gather this information, if it exists, from literally hundreds of organizations across the U.S.

To reflect at least some level of funding from these groups, TEConomy worked to develop a conservative estimate. Building off of prior work, TEConomy first estimated a 2019 value for total research funding from voluntary health associations.[27] An assumption is then made that, at a minimum, these organizations fund human genetics and genomics research at approximately the same "share" that NIH does. To stay most conservative, the "core" percentage of NIH funding, 10.8% (see Table 3), is applied to the estimated FY 2019 voluntary health association total funding level of $1.56 billion to generate a human genetics and genomics research funding estimate of $169.1 million in 2019. The actual level of funding from these organizations is likely to be considerably larger. Anecdotally and via website information, it appears the cutting-edge research funded by these groups is becoming more and more engaged with genetic and genomic exploration.

Research Funding Summary

Combined, human genetics and genomics research funding as captured from funding organizations reaches $3.4 billion under the most conservative estimate of NIH funding. Using the IMPLAN model this level of research funding is estimated to directly employ 13,800 researchers throughout the U.S.

2. Core Industry Drivers—Employment of Core Human Genetics and Genomics Firms

The previous section captures the level of human genetics and genomics research activity stemming from federal and other sources of funding. This section develops an employment-based estimate of the size and scope of the core firms operating primarily, if not exclusively, within the human genetics and genomics domain.

Following a similar approach to the authors' prior work analyzing the genetics and genomics industry, a database of firms with their total employment was developed. The database was initiated by starting with the previous work's 2010 database of firms, excluding those that were primarily in the plant/animal/agricultural domain and determining whether these firms were still in business in 2019, and correcting for considerable mergers and acquisition activity that has occurred within the industry over the past decade. For the purposes of this study, firms that are part of the important instrumentation (e.g., Illumina, ThermoFisher) or bioinformatics subsectors are included in this database even though their products and services may also be used outside of the specific human genetics and genomics domain.

To supplement this existing firm database, lists of firms from a variety of organizations, websites, and market research publications were curated to generate an additional set of U.S. firms for review and inclusion. These firms were then evaluated using web-based research to determine their fit within the human genetics and genomics domain and whether they were still in business. If they met these criteria, an employment value was developed using third-party databases such as Dun and Bradstreet and PitchBook (a provider of angel and venture capital information) and, at times, the firm's website or LinkedIn pages.[28]

The impacts are modeled as an aggregation of IMPLAN sectors, as appropriate, to capture the extent and variety of research and corporate activities by using employment to drive the direct impacts of the economic impact model. These firms and their employment were classified into one of six human genetics and genomics core industry subsectors (Table 4). The employment figures reflect the total employment of firms in each industry subsector.

Within this employment of more than 89,000, some specific caveats and specifications are warranted. The

27 *U.S. Investments in Medical and Health Research and Development 2013 – 2018*. Research!America, Fall 2019.
28 Developing true employment values from these sources can be difficult due to reasons such as outdated data, self-reported data, and the effects of M&A activity. Conservative employment estimates were made, if required.

Table 4: Employment by Human Genetics and Genomics Core Industry Subsectors

Industry Subsector	Est. of 2019 Employment
Core Analytical/Biomedical Instruments and Equipment Manufacturers	26,758
Core Small Biopharma Manufacturers	22,383
Core R&D/Biotech Firms	21,797
Core Medical/Diagnostic Laboratories	14,639
Core Software/IT/Bioinformatics Service Firms	3,729
Core Genetic Counselor and Other Related Services	158
Total Core Firms' Employment	89,464

Sources: Lists of firms developed from various industry and professional websites. Employment estimates from corporate websites and Dun & Bradstreet data. Additionally, LinkedIn information was used, at times, to update and correct some firms' employment levels.

data only include small/mid-sized biopharmaceutical focused firms' employment—the impact of large biopharmaceutical firms is captured via sales estimates in the next section. Dun & Bradstreet and corporate website information was used to determine whether to include firms within the R&D/biotech firms' category versus small biopharma category. The core medical/diagnostic laboratories category does not include employment for Quest Diagnostics and Laboratory Corporation of America (LabCorp); rather their role is estimated in the next section. An assessment of the list of genetic counselors available from the National Society of Genetic Counselors was used to identify U.S. firms. These firms were then classified into medical/diagnostic laboratories or genetic counselor practices as warranted as many medical/diagnostic laboratories include genetic counselor employees. The specific category for genetic counselor and other related services was used to capture, as best as possible, the employment of those firms operating without a direct testing capability. It should be noted that the final economic impact assessment also included 1,752 individual counselors, including those practicing within the VA's Million Veteran Program, that are not captured within the employment figures in Table 4.

Core Industry Drivers Summary

Across six industry subsectors, the economic impact model is driven by direct employment within the core industry drivers of 89,464 employees.

3. Extended Industry Drivers

Two subcomponents of the industry analysis efforts required different approaches to estimating their size and importance in driving the human genetics and genomics economic impacts.

Pharmaceutical Manufacturing

The core industry employment analysis captured small and/or nascent biopharmaceutical firms whose sole focus is the development of pharmacogenetic/genomics drugs (drugs associated with a genetic test) and other genetic and genomic therapies. However, larger pharmaceutical manufacturers with development efforts ranging from small molecule chemistries to large molecule biologics also are a significant component of human genetics and genomics-related economic activities. For these firms, a different approach is used to apportion part of their operations (sales) to the human genetics and genomics domain.

The original *Economic Impact of the Human Genome Project* study also treated these large pharmaceutical firms in a slightly different manner from core industry employment. In the previous study, an estimate of genetic and genomic research expenditures was used to reflect these firms' involvement in the human genetics and genomics domain—one that was just beginning to see the opportunities that genetics and genomics would provide in the development of modern medicines. In the decade since these estimates were developed, the role of human genetics and genomics in biopharmaceutical development is still evolving. However, its use has extended its impact on these large biopharmaceutical firms beyond simply research and further into targeted drug development and the labeling, use, and efficacy of existing products ultimately driving corporate sales.

Lists of the U.S.-based members of the Pharmaceutical Research and Manufacturers Association (PhRMA)[29] and pharmaceuticals by sales for 2019 were examined.[30] Using a sales cut-off value of $1.5 billion to represent the most active biopharmaceutical products generates a list of 102 medicines. Of these, 26 are currently listed in the FDA Table of Pharmacogenomic Biomarkers in Drug Labeling.[31] This subset of biomarker-labeled medicines account for 28% of global sales of these 102 medicines. To generate a U.S. specific estimate, this 28% was then applied to a market study value of 2019 total U.S. prescription drug sales of $177.7 billion yielding an estimated $49.5 billion in U.S. prescription drug sales with one or more genetic biomarkers.[32] This estimate is considered conservative due to the higher usage rate of more expensive pharmaceuticals within the U.S. Total global sales of these 26 medicines reached $96.5 billion, nearly twice the value used to drive the U.S. economic impact estimate.

National Medical Testing Laboratories

Two medical/diagnostic laboratories, Laboratory Corporation of America (LabCorp) and Quest Diagnostics, together account for more than $16.5 billion in total U.S. sales and more than 90,000 U.S. workers.[33] While human genetic and genomic testing is becoming more common in U.S. healthcare, it is still a smaller share of overall medical and diagnostic testing. Using corporate websites and reporting as well as third party market studies, TEConomy developed estimates for these two firms' human genetic and genomic-related sales and employment.[34] Together,

29 PhRMA. "Members." www.phrma.org/en/About/Members. Accessed 12 May 2021.
30 PhRMALive and Outcomes LLC. "Top 200 Medicines Annual Report: Climbing Mount Humira."
31 U.S. Food & Drug Administration. "Table of Pharmacogenomic Biomarkers in Drug Labeling." www.fda.gov/media/124784/download. December 2020.
32 Spitzer, Dan. *Brand Name Pharmaceutical Manufacturing in the U.S.* IBISWorld Industry Report 32541a, August 2020.
33 Data obtained and calculated from corporate 10-K filings reflecting 2019 performance.
34 Curran, Jack. *Diagnostic & Medical Laboratories in the U.S.* IBISWorld Industry Report 62151, November 2020.

these two firms are estimated to account for just under 22,000 workers and more than $3.77 billion in sales within the human genetics and genomics domain.

Extended Industry Drivers Summary

The revenue and employment size of the firms captured within the Extended Industry Drivers warrant special and distinct attention. For both subcomponents, the use of human genetics and genomics technologies and techniques constitutes a relatively small share of overall operations. However, due to the sheer size of these corporate operations, these two subcomponents directly add more than $53.3 billion and nearly 61,000 jobs to the U.S. economy.

C. Economic Impact Analysis of Human Genetics and Genomics in the U.S.

The combination of the three economic impact drivers presented above—Research Drivers, Core Industry Drivers, and Extended Industry Drivers—yields a significant direct economic presence in the U.S. economy. Using the most conservative NIH research funding levels and using the IMPLAN model to estimate employment from sales (or research funding levels) values, it is estimated that **the U.S. human genetics and genomics research and industrial domain employs nearly 166,000 workers** as a direct result of these operations (Table 5). This number includes human genetics researchers, medical geneticists, and genetic counselors, as well as a large number of workers in adjacent, corporate, or operational roles in firms developing lab equipment and software, performing clinical genetics and genomics testing, or manufacturing pharmacogenomic drugs. **Overall, this combined set of drivers directly generate over $108 billion within the U.S. economy.** With this set of direct drivers, the IMPLAN model is used to estimate the total impacts of human genetics and genomics domain on the U.S. economy (Table 5).

From an employment perspective, these impacts show the U.S. human genetics and genomics domain supporting an additional 684,000 jobs (indirect and induced effects) within the U.S. economy for a total employment impact of more than 850,000 workers, reflecting an employment multiplier of 5.12. For every direct job in the human genetics and genomics domain, 4.12 additional jobs are generated throughout the U.S. economy. On average, due to the higher technical and educational requirements,

Table 5: Economic (Expenditure) Impacts — Core NIH Human Genetics and Genomics Funding Scenario (2019)

Impact Type	Employment	Labor Income ($B)	Value Added ($B)	Output ($B)	State/ Local Tax Revenues ($B)	Federal Tax Revenues ($B)
Direct Effect	165,973	$21.66	$52.59	$108.16	$2.90	$5.18
Indirect Effect	288,866	$24.81	$43.59	$87.04	$2.90	$5.38
Induced Effect	395,425	$22.40	$39.44	$70.15	$3.67	$4.95
Total Impacts	850,263	$68.88	$135.62	$265.35	$9.47	$15.51
Multiplier	**5.12**	**3.18**	**2.58**	**2.45**		

Source: TEConomy analysis of Human Genetics and Genomics Input Dataset; IMPLAN 2019 U.S. Impact Model.

the direct jobs generate more than $130,000 in annual compensation (income and benefits) per worker while overall, the total jobs supported by the human genetics and genomics domain still average over $81,000 in compensation per employee.

In terms of value added, or the contribution to U.S. GDP, the human genetics and genomics domain directly adds more than $50 billion to U.S. GDP, and through the economic ripple effects these efforts support, in total, nearly $136 billion of U.S. GDP.

The direct output also leads to considerable additional economic activity in the U.S. For every $1 of output (e.g., sales or research expenditures), $1.45 of additional sales are generated in the U.S. economy leading to an **overall economic impact of U.S. human genetics and genomics of more than $265 billion in 2019.**

The federal tax revenues of $5.18 billion generated by the direct operations of the human genetics and genomics domain alone surpasses the single year federal investment in human genetics and genomics of approximately $3.26 billion across all federal agencies. In the simplest of terms, **from a federal investment perspective, the overall economic impacts of U.S. human genetics and genomics generates a return on investment (ROI) of more than 4.75 to 1.00.**

Progress Over the Past Decade

At the outset of this chapter, the 2010 impact estimates for the Human Genome Project were provided. By comparison, this current analysis shows the dramatic increase in impact over the past decade that human genetics and genomics has had on the U.S. economy.

Even though this current effort is focused solely on "human" genetics and genomics, the size of the workforce directly employed in these endeavors has more than tripled over the decade from just under 52,000 workers in 2010 to nearly 166,000 in 2019. Similarly, direct output has dramatically surpassed the previous estimate as human genetics and genomics developments are leading to actual and significant sales in 2019. Direct output was estimated to be $22.6 billion in 2010 compared to $108.2 billion in 2019.

The economic importance of U.S. human genetics and genomics cannot be denied. The HGP impact study found that the broader genetics and genomics field, for which the human domain is just one component, had an economic impact of $67 billion in 2010. The growth of the human genetics and genomics field over the past decade has been substantial, with this one domain area now representing a total economic impact of more than $265 billion in 2019.

D. Healthcare Costs for Genetic Diseases

The application of human genetics and genomics, as discussed in the next chapter, is providing a wide range of functional benefits in healthcare in terms of identification of patient predisposition to genetic disease, diagnosis of diseases, identification of targets for new drugs, precision drug dosing and limitation of adverse drug events, and development of new approaches to treatments and cures through gene therapy and gene editing. As will be shown, human genetics and genomics are very much at an inflection point where many benefits are now occurring as research discoveries translate into clinical innovations.

Because the large-scale clinical application of human genetics and genomics is a relatively new phenomenon, there is relatively limited literature on the economic impacts of its applications to clinical care. Longitudinal studies tracking impacts over time are particularly scarce and indeed are challenging to interpret given that the cost of gene sequencing, up to and including whole-genome sequencing, has plummeted. Sequencing the genome was, just a few years ago, prohibitively expensive. However, advancements in technology mean that cost is no longer a primary barrier to use in the clinic.

One way to evaluate the economic impacts of human genetics and genomics is to examine the cost burden of disease imposed on the economy by diseases

that are predominantly genetic in their etiology. The majority of what is labelled "rare diseases and disorders" falls into this category, often being single-gene disorders. As such, the cost of these diseases can serve as a surrogate for at least understanding the scale of disease burden related to rare diseases, and by extension, provide intelligence on the kinds of economic costs that may be ameliorated through advancements in genetic and genomic diagnosis and treatments. A recent study conducted by the Lewin Group for the EveryLife Foundation for Rare Diseases makes an important contribution to the literature and provides robust evidence for the very substantial burden imposed by predominantly genetic rare disease.[35] The research program, titled "The National Economic Burden of Rare Disease Study," uses data on medical care costs, together with a detailed survey of rare disease patients and caregivers, to derive estimates for the burden of 379 rare diseases measured in terms of their one year impact (2019). The study is conservative and only estimates the burden for 379 diseases for which survey data were available (versus in excess of 7,000 rare diseases identified[36])—but the results speak to the large-scale costs of such disease to society, to patients, and to patients' families. **The study finds the total cost of the 379 assessed diseases to be $966 billion for 2019.** This total cost burden is shown in Table 6.

The Lewin authors do not choose to extrapolate their findings for 379 rare diseases to the more than 7,000 such diseases that exist because they were unsure of whether the 379 comprise a representative sample of rare diseases extant. The measures shown in Table 6 are thus conservative. What they do provide, however, is a window into the very large-scale cost of these (predominantly genetic) diseases—a cost that in just one year in the U.S. imposed a burden of almost $1 trillion. The direct cost of the 379 rare disease impacts studied by Lewin ($418 billion) can be put in context by examining total healthcare expenditure data maintained

Table 6: Economic Burden of 379 Rare Diseases in the United States (2019)

Cost Element	Description	2019 Impact
Direct medical costs	Includes inpatient hospital or outpatient care, physician visits, Rx medications, durable medical equipment.	$418 billion
Indirect costs: productivity loss	Includes forced retirement, absenteeism, presenteeism, and a reduction in community participation and volunteer service.	$437 billion
Non-medical and uncovered healthcare costs	Includes paid daily care, necessary home and vehicle modifications, and transportation and education costs. Also includes health care services not covered by insurance: experimental treatments, medical foods, and more.	$111 billion
		$966 billion

Source: Lewin Group. "The National Economic Burden of Rare Disease Study." Prepared for: EveryLife Foundation for Rare Diseases. February 25, 2021.

35 35. Yang, Grace, et al. *The National Economic Burden of Rare Disease Study*. Lewin Group for EveryLife Foundation for Rare Disease, 25 Feb. 2021.
36 Liu, Zhichao, et al. "Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic testing in Rare Diseases: Where Are We?" *Trends in Genetics*, vol. 35, no. 11, Nov. 2019.

by the Centers for Medicare and Medicaid Services (CMS). CMS data for 2019 show total healthcare costs for hospital, physician and other provider services, medications, and durable medical equipment being $2.4 trillion in 2019.[37] It is evident, therefore, that the high amount of care required in diagnosing and treating rare diseases comprises a significant portion of overall healthcare costs. At $418 billion (again, just for 379 rare diseases), the direct care costs of these diseases equate to 17.4% of total national direct spending across equivalent expenditure categories.

Of course, many common diseases also have genetic involvement. Cancer is fundamentally a genetic disease, whereby gene mutations cause uncontrolled cell growth. Many more diseases, including cardiovascular diseases, gastrointestinal diseases, autoimmune disorders, psychiatric and neurological disorders, musculoskeletal disorders, etc., have genetic involvement, typically involving many genes (i.e., polygenic). Quantifying the overall cost burden of these diseases, where genetics is part of the equation, is particularly challenging. Still, just referencing cancer, the American Cancer Society reports that "the Agency for Healthcare Research and Quality (AHRQ) estimates that the direct medical costs (total of all health care costs) for cancer in the U.S. in 2015 were $80.2 billion"[38]—with 52% of this cost being for hospital outpatient or doctor office visits, and 38% of the cost for inpatient hospital stays.

While a definitive total cannot be calculated with available data, there should be little doubt that the economic burden of genetic and genomic-related diseases in the U.S. reaches into the trillions of dollars on an annual basis. However, as examined in the next chapter, researchers and clinicians in research institutes, universities, industry, and government labs are making far-reaching contributions to reducing the many burdens associated with genetic and genomic diseases and disorders—making deep progress in the clinical application of genetics science and technology advancements to make a very real difference in the lives of millions of patients.

37 Centers for Medicare & Medicaid Services. "National Health Expenditures 2019 Highlights." www.cms.gov/files/document/highlights.pdf. Accessed 12 May 2021.
38 American Cancer Society. "Economic Impact of Cancer." www.cancer.org/cancer/cancer-basics/economic-impact-of-cancer.html. Accessed 12 May 2021.

III.The Functional Impacts of Human Genetics and Genomics

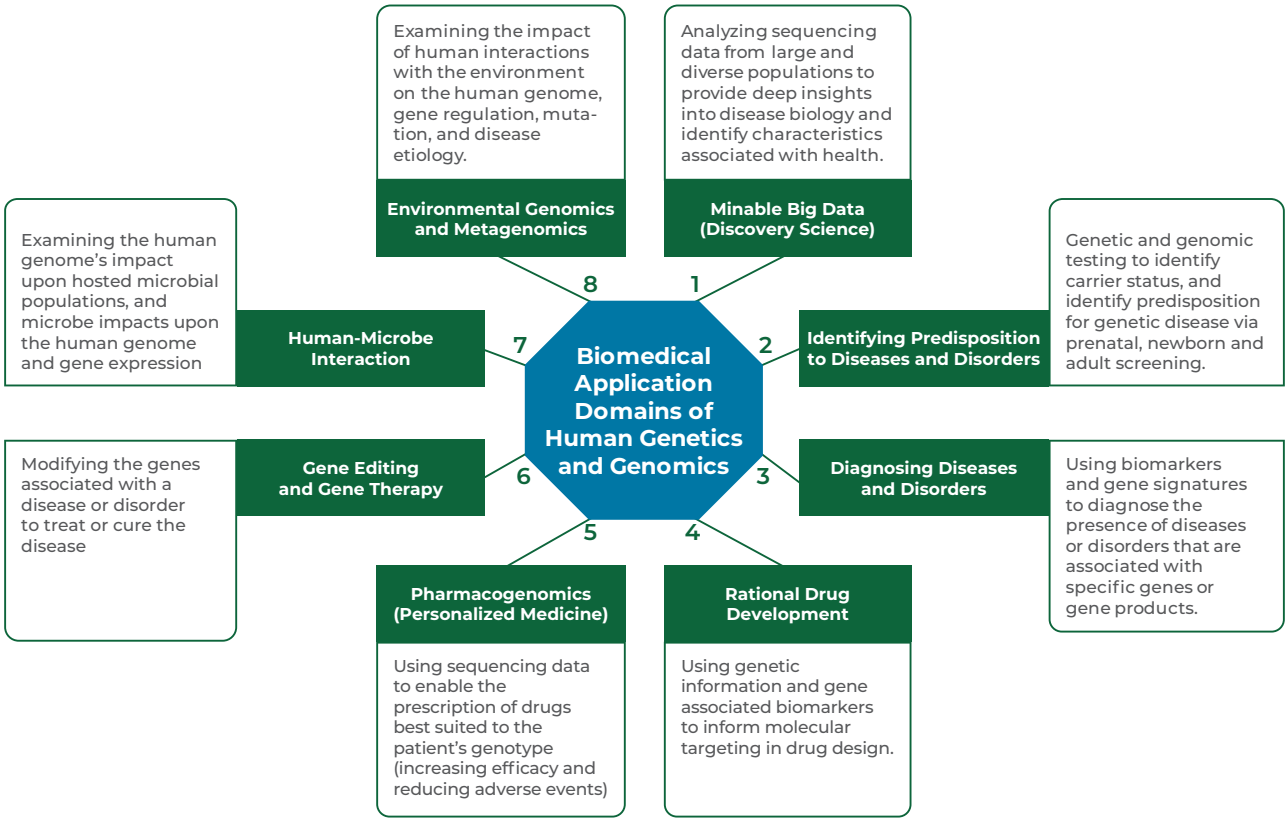
A. The Structure of Functional Impacts (Application Domains of Human Genomics)

The application of genetics and genomics advancements to healthcare can certainly be viewed through an economic lens, and the economic impact of this dynamic science and technology-driven sector is extremely large. However, the generation of business income and jobs is not the *raison d'etre* for medical genomics. Human genetics and genomics are pursued for the scientific and clinical insights they enable, and for their functional impacts, these being the domains of application of the science and technologies of genetics and genomics to preserving and improving human health.

Genomics has become fundamental to advancement of biomedical research, and the insights, tools, and technologies provided by genetics and genomics are now seeing increasingly widespread deployment in clinical healthcare. This chapter seeks to provide a broad overview of the key areas (domains) in which human genetics and genomics are being applied in clinical research and healthcare. The functional impacts are divided into eight medical domains (Figure 3): **minable big data**, whereby large data sets of multi-patient data are providing deep insights into disease biology (and also identifying characteristics associated with health); **identification of genetic predisposition** to diseases and disorders; **diagnosis of diseases** and disorders though genetic signatures; **rational drug development**, whereby genetics information informs molecular targeting in drug design; **pharmacogenomics**, which enables the personalized prescription of drugs best suited to the person's genetics (with a goal of increasing effectiveness and reducing adverse events); **gene editing and gene therapy**, whereby genes associated with disease are modified to treat or cure the disease; and two more "emerging" areas in which **human-microbe genetic interactions** and **human-environmental metagenome interactions** are being examined for association with human gene expression, regulation, mutation, and disease.

In addition to applications to human medicine, there are also several additional human application domains relevant for non-medical uses. These are briefly discussed herein, covering applications in forensic science, anthropology and genealogy, evolutionary biology, and paternity testing.

Figure 3: Functional Biomedical Impact Domains (Applications) of Human Genetics and Genomics



Source: TEConomy Partners, LLC.

B. Fundamental Knowledge Advancement

Before discussing the clinical impact domains of human genetics and genomics, it is important to note that applied and clinical research depends upon a healthy base of discoveries and fundamental insights that are derived through basic research, which may be defined as:

> *Systematic study directed toward greater knowledge or understanding of the fundamental aspects of phenomena and of observable facts without specific applications towards processes or products in mind.*[39]

Genetics can inform us of our past evolution, and genetics and genomics as disciplines have themselves been evolving. Fundamental scientific research, facilitated by advancements in genomics technologies and analytics, has provided a stream of notable discoveries, with just some highlighted below:

- The ENCODE project (ENCyclopedia of DNA Elements) revealed that non-coding DNA, previously termed "junk DNA" because it was thought to be a relic of evolution with little biological function, instead has specific functionality in transcription and translational regulation of protein-coding. In other words, most of it is not

39 Cornell Law School. "32 CFR § 272.3 - Definition of Basic Research." www.law.cornell.edu/cfr/text/32/272.3. Accessed 12 May 2021.

> The importance of basic science derives from its contribution to knowledge deeper within the tree of information and, consequently, its greater potential for integration with other facts. In contrast, the importance of translational science lies in its practicality. Hence, we do not view basic and translational science as one being more important than the other but rather as complementary areas of human endeavor, with the important distinction that basic science findings often precede advances in translational science. We also note that observations in translational or applied science can generate new questions for fundamental research, as illustrated from the fact that vaccination preceded the field of immunology. Hence, the epistemological flow is bidirectional, and investments in both types of science are needed."
>
> **Ferric C. Fang and Arturo Casadevall. "Lost in Translation—Basic Science in the Era of Translational Research."** ***Infection and Immunity*****, vol. 78, no. 2, Feb. 2010.**

junk at all; it is central to life functions (although not fully understood in terms of functionality).

- Basic research into gene silencing led to fundamental discoveries regarding RNA, messenger RNA (mRNA), and development of techniques for RNA interference that enabled human genes to be disabled in a very precise manner to better study their effect and function. Further research has elucidated the presence of multiple types of non-coding RNAs and their impacts on gene expression, one class of which, microRNAs (mnRNAs), may be regulating more than half of all human genes.
- Ongoing refinement to the reference human genome has occurred, closing many of the gaps in the original sequences and uncovering previously uncharacterized parts of the human genome with potential for significant discoveries. [40]
- Recent studies have been elucidating "mosaicism", which is the term used to describe genomic variation among cells (both germline and somatic) within an individual. Much still remains to be studied in this area to better understand "mosaic variation in both nuclear and mitochondrial DNA, the mechanisms that generate mosaicism, and the roles of mosaicism in physiology and human disease."[41]
- Recent work is focused on "studying patterns of gene expression in individual cells, a step that has been driven by new methods for single-cell RNA sequencing and chromatin analysis. Tens of millions of cells have been characterized thus far on route to a complete cell Atlas of the human body. This effort is revealing hundreds of new cell types and characterizing the ways in which cell types differ between healthy people and people with various diseases."[42]
- Basic research has found that "not all genes are expressed in all tissues and that not all genes are expressed during all developmental stages."[43] This has important downstream implications in drug development, where, for example, a research team developing a drug for infantile epilepsy would need to know

40 Miga, Karen H. "Human Genome: Bridging the Gaps." *Nature*, vol. 590, no. 11, Feb. 2021.
41 Green, Eric D., et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 Oct. 2020.
42 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine*, vol.384, no. 1, 7 Jan. 2021.
43 Pharmaphorum Connect. "The Future of Genomic Medicine: Can it Fulfil its Promises?" www.pharmaphorum.com/views-analysis-patients/the-future-of-genomic-medicine-can-it-fulfil-its-promises/. Accessed 12 May 2021.

whether a drug target is "expressed in the brain and also during early development."[44]

It should also be noted that the Human Genome Project, and the ongoing development of genomic tools and datasets, generated a rather seismic shift in the way in which fundamental research is performed. The big data, information mining approach that gene sequencing enables has "transformed the nature of medical discovery, enabling scientists to undertake comprehensive and powerful explorations rather than being confined to testing hypothesis focused on candidate pathways."[45]

What is clear is that the human genome is immensely complex, and there is much still to be discovered through fundamental research into its structure, mechanisms of action, and its interface with biochemical signals from non-genomic origin. In some diseases, hundreds of individual genes are being found to have an effect on disease development and progression, often in concert with multiple environmental and physiological factors. Solving such multigenic and multifactorial challenges is no small task, but distinct progress is being made, aided by tremendous technological advancements in sequencing and data analytics platforms. Any scientist, or group of scientists, embarking on finding solutions to individual human diseases will typically recognize that the path from question to discovery to therapy or cure is long (sometimes spanning their career, if successful at all). Modern genetics and genomics are, however, providing new, more brightly lit paths informed by quantitative datasets that can be mined for insights and therapeutic targets. As noted in NHGRI's strategic vision:

> *...the past decade has brought the initial realization of genomic medicine, as research successes have been converted into powerful tools for use in health care, including somatic genome analysis for cancer (enabling development of targeted therapeutic agents), noninvasive prenatal genetic screening, and genomics-based tests for a growing set of pediatric conditions and rare disorders, among others.*[46]

For some diseases, especially those where a single or only a few genes are involved, real breakthroughs are occurring. The section that follows describes the domains of health sciences and clinical care where these impacts are being felt.

C. Functional Applications for Human Health

As fundamental genomic knowledge has expanded, the enhanced understanding of genetic mechanisms, in concert with access to rich whole exome and genome datasets (and associated reference compendia of human gene variants), has opened the door to a new era of discovery and progress in medicine. The impacts of these advancements are now increasingly reverberating across medicine, a fact highlighted by Eric Green, the Director of the NHGRI, and colleagues who note that:

> *With 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 their initial expectations, with many more anticipated in the next decade.*[47]

Much of the advancement being seen is enabled by dramatic progress in genome sequencing technology performance and cost effectiveness. A virtuous cycle

44 Ibid.
45 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine*, vol.384, no. 1, 7 Jan. 2021.
46 Green, Eric D., et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 Oct. 2020.
47 Ibid.

has occurred, whereby the speed increases and cost decreases in sequencing have facilitated the assembly of exabytes[48] of genomic information that can be mined (assisted by highly advanced and automated analytical systems) for unique insights into genome structure and function.[49] As highlighted by Green:

> *Leading the signature advances has been a greater than one million-fold reduction in the cost of DNA sequencing. 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. These new tools, together with increasingly sophisticated statistical and computational methods, have enabled researchers to create rich catalogs of human genomic variants, to gain an ever deepening understanding of the functional complexities of the human genome, and to determine the genomic bases of thousands of human diseases.*[50]

This leads us to the first functional impact domain of genetics and genomics in human medicine, minable big data, and what it enables, discovery science.

1. Minable Big Data (Discovery Science)

Laurence Hurst highlights the central role that data are playing in advancing functional applications of genomics, noting that:

> *Genomics is in an age of exploration and discovery. Whether we are discovering the genomes of more species, the genomes of more individuals in a species, or more genomes within an individual (at single-cell resolution), we are very much in a phase where we are letting the data lead.*[51]

Studying one person's genome (or in the case of the "reference" human genome, a composite of a few people) has provided valuable information regarding the structure of the human genome and the number of protein-coding genes. It has also allowed for comparison to an expanding library of reference genomes for other organisms, helping to identify regions of similarity and difference that could help illuminate functionality. At a time when whole-genome sequencing cost many millions of dollars per genome, the field was limited to small volumes of sequenced genomes to work with. As sequencing costs declined and sequencing speed increased, the ability to generate data from a large number of individuals started to become realistic, and this has opened new horizons for research.

Currently, whole genome sequencing and whole exome (the part of the genome formed from exons that code proteins) sequencing is fast and affordable (requiring just a day and $689 for whole genome sequencing), and as affordability and speed have increased, the number of sequenced individuals has expanded exponentially. In a recent edition of Science, it is noted that "today, more than 30 million individuals have access to their detailed genomic datasets,"[52] while Ewan Birney of the European Bioinformatics Institute notes that "estimates show that over 60 million patients will have their genome sequenced in a healthcare context by 2025."[53]

Having access to an extremely large volume of sequenced individuals creates a dramatically enlarged platform for discovery. Each of us has a unique genome. While 99.9% of the genome between individuals will be the same,[54] the 0.1% differences can add up to profoundly dissimilar physical characteristics and differential predisposition to

48 An exabyte = 10006 bytes (1,000,000,000,000,000,000 bytes).
49 Stephens, Zachary D., et al. "Big Data: Astronomical or Genomical?" *PLOS Biology*, vol. 13, 7 July 2015. doi:10.1371/journal.pbio.1002195.
50 Green, Eric D., et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 Oct. 2020.
51 Cheifet, Barbara. "Editorial: Where is genomics going next?" *Genome Biology*, vol. 20, no. 17, 22 Jan. 2019. doi:10.1186/s13059-019-1626-2.
52 Zielinski, Dina and Janiv Erlich. "Genetic Privacy in the Post-Covid Word." *Science*, vol. 371, no. 6529, 5 Feb. 2021.
53 Birney, Ewan. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.
54 National Human Genome Research Institute. "Genetics vs. Genomics Fact Sheet." www.genome.gov/about-genomics/fact-sheets/Genetics-vs-Genomics#:~:text=All%20human% 20beings%20are%2099.9,about%20the%20causes%20of%20diseases. Accessed 12 May 2021.

Table 7: Large Population Precision Medicine Initiatives

Country	Project/program name	Expected size	Common diseases	Rare diseases (and cancers)
Australia	Genomics Health Futures Mission	200,000		✔
Canada	Canadian Genomics Partnership for Rare Diseases and Canadian Longitudinal Study on Aging	Nationwide	✔	✔
China	Precision Medicine Initiative	100,000–100 million	✔	✔
Denmark	Danish National Genome Center	60,000	✔	✔
Dubai	Dubai Genomics	Nationwide	✔	
Estonia	Personalised Medicine Programme	150,000	✔	
European Union	1+ Million Genomes Initiative	1,000,000+	✔	
Finland	FinnGen	500,000	✔	
France	Genomic Medicine France 2025	235,000 each year	✔	✔
Hong Kong	Hong Kong Genome Project	50,000		✔
Italy	SardiNIA Project	60,000	✔	
Japan	GEnome Medical alliance Japan	Nationwide	✔	✔
Saudi Arabia	Saudi Human Genome Program	100,000	✔	✔
Singapore (And International)	Genome Asia 100 K	100,000	✔	
Thailand	Genomics Thailand	50,000	✔	✔
Turkey	Turkish Genome Project	100,000–1,000,000	✔	✔
United Kingdom	100,000 Genomes Project	100,000		✔
United Kingdom	Accelerating Detection of Disease	5,000,000	✔	
United States	NHGRI Genomic-Medicine	Nationwide	✔	✔
United States	All of Us Research Program	1,000,000+	✔	

Source: Identified in Chung, B.H.Y., Chau, J.F.T. & Wong, G.KS. “Rare versus common diseases: a false dichotomy in precision medicine.” npj Genom. Med. 6, 19 (2021).

disease, rates of metabolizing drugs, and other factors. Identifying and understanding these differences becomes more feasible the more sequences and data are available. The devil is in the details, and more sequences provide more details.

Large-scale sequencing implementation is enabling the ongoing assembly of robust, evidence-based resources for the identification and classification of genomic variant pathogenicity (variants associated with causation of disease). To-date, the vast majority of larger-scale whole genome sequences have been produced in Western nations, with the result that the available data skew quite significantly in terms of individuals of European ancestry. This bias in the data is being addressed through multiple initiatives worldwide that will contribute greatly to a broader base of represented humanity. Recent research by

Brian Hon Yin Chung, Jeffrey Fong Ting Chau, and Gane Ka-Shu Wong summarizes many of the larger (>20,000 subject genomes) precision medicine projects (for which sequencing is primarily a major element), showing how in forthcoming years, the richness of sequenced populations will be enhanced significantly.[55] The global distribution of these studies (Table 7) holds promise for the development of reliable genomic data on many different populations and sub-populations, helping to build a more inclusive atlas of genome variability across the human species.

The expanding diversity in the base of human genome sequence data is further highlighted by Rotimi, Callier, and Bentley who recently note that:

> *Growing prioritization of diverse populations in genomics research has begun to respond to these gaps. Programs, such as TOPMed, All of Us, International Common Disease Alliance, Human Heredity and Health in Africa (H3Africa), Million Veteran Program, GenomeAsia, and the COVID global consortium, contribute to advances in diversity and inclusion among research participants.*[56]

In the U.S., several large-scale sequencing initiatives are ongoing. Among the largest is the NIH's "All of Us" program, which began in 2018 and is consolidating genetic, health, and environmental data for more than one-million participants, providing a robust resource for evaluating genotype-to-phenotype associations. Another federally initiated program is the "Million Vets Program" (MVP) that, similar to the NIH program, is collecting deep data to allow the study of links between genes, lifestyle, and military exposures and their associated impacts on health and illness. Since launching in 2011, over 825,000 veterans have signed up to participate. The VA notes that "in addition to common health conditions that affect everyone, such as cancer, cardiovascular disease, and diabetes, MVP researchers are also looking at conditions specific to Veterans. This

Increasing Sequenced Population Diversity to Enhance Studies of Genetic Variation.

Up until relatively recently the participants involved in genomic research have largely been of European ancestry.

Multiple initiatives are now underway to substantially increase diversity in genomic datasets. For example, the Human Heredity and Health in Africa (H3Africa) initiative has enrolled more than 60,000 research participants and engaged more than 500 African scientists.

> "Research in genome biology is often descriptive in nature, sequencing genomes and metagenomes, profiling epigenomes and transcriptomes, charting evolutionary history, and cataloging disease linked risk loci. Thanks to major technological advances, we can now generate such descriptive datasets using high throughput platforms."
>
> **Christoph Bock, CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, quoted in Barbara Cheifet. "Editorial: Where is Genomics Going Next?" *Genome Biology*, vol. 20, no. 17, 22 Jan. 2019.**

55 Chung, B.H.Y., et al. "Rare versus Common Diseases: a False Dichotomy in Precision Medicine." *npj Genomic Medicine*, vol. 6, no. 19, 24 Feb. 2021. doi.org/10.1038/s41525-021-00176-x.

56 Rotimi, Charles N., et al. "Lack of Diversity Hinders the Promise of Genome Science." *Science*, vol. 371, no. 6529, 5 Feb. 2021.

includes PTSD [post-traumatic stress disorder], suicide prevention, TBI [traumatic brain injury], and tinnitus."[57]

While advances in the technologies for gene sequencing are at the forefront in generating large and deep datasets from diverse populations, equally important, have been advancements in advanced data analytics comprising the use of analytical computer algorithms and statistical techniques acting upon large-scale sets of structured and unstructured data to derive actionable insight (see sidebar definition for advanced analytics).

Genetic and genomic data, health record data, and environmental data each provide important insights on their own but promise far greater intelligence when examined together. This presents the challenge of analyzing extremely large-scale heterogenous data compiled from multiple sources. Fortunately, as these big data resources have been built, there has been parallel advancement in the science and technology of advanced data analytics, up to and including artificial intelligence (AI) based systems. Advanced analytics provides a pathway forward in terms of mining big genome and genome-phenome datasets to provide functional insights and impacts in broad areas such as: biomarker discovery and identification of druggable targets; multi-gene to disease associations; environmental effects on gene regulation and expression; gene expression effects of prior infectious diseases, etc. The combination of genomics and phenomics big data and advanced analytics provides a powerful pathway forward for modern life science discovery and healthcare improvement. This is highlighted by Liu, Zhu, Roberts, and Tong who note that: "AI is starting to realize its potential in augmenting phenome-wide and genome-wide data profiles to improve clinical utility and diagnostic power."[58]

Evidence for this technological convergence of biomedical big data and AI is seen in evident clusters of new business ventures forming to pursue commercialization of associated opportunities. Research by TEConomy recently used machine learning to identify clusters of U.S. activity focused on advanced analytics applications.[59] Figure 4 illustrates three distinct clusters of venture capital funded enterprises forming in this space, comprising: 1) drug discovery and precision medicine, 2) healthcare analytics, and 3) wearable health monitoring device analytics. The growth and interaction of these clusters builds upon the promise of big data analytics using genome and phenome information to derive clinical health insights and improvements.

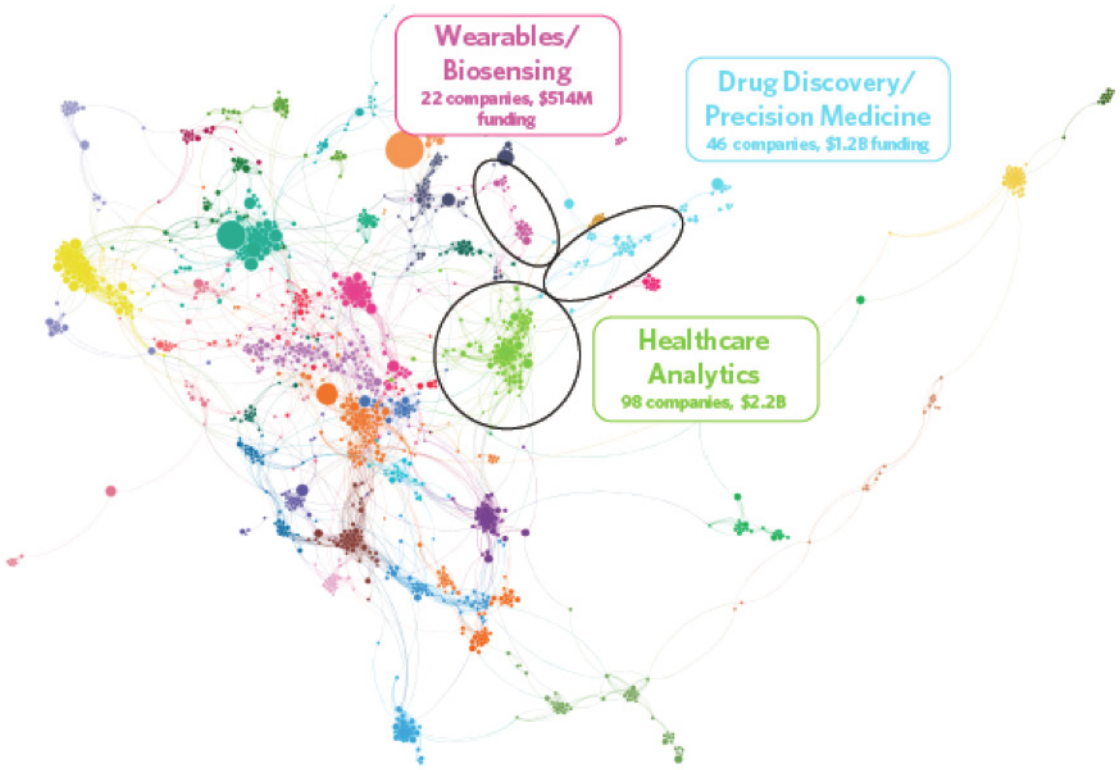
> "Advanced Analytics is the autonomous or semi-autonomous examination of data or content using sophisticated techniques and tools, typically beyond those of traditional business intelligence (BI), to discover deeper insights, make predictions, or generate recommendations. Advanced analytic techniques include those such as data/text mining, machine learning, pattern matching, forecasting, visualization, semantic analysis, sentiment analysis, network and cluster analysis, multivariate statistics, graph analysis, simulation, complex event processing, neural networks."
>
> **Gartner. "Gartner Glossary." www.gartner.com/en/information-technology/glossary/advanced-analytics. Accessed 12 May 2021.**

57 VA Million Veteran Program. "About the Million Veteran Program." www.mvp.va.gov/webapp/mvp-web-participant/#/public/about. Accessed 12 May 2021.

58 Liu, Zhichao, et al. "Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic testing in Rare Diseases: Where Are We?" *Trends in Genetics*, vol. 35, no. 11, Nov. 2019.

59 Tripp, Simon, et al. *Artificial Intelligence and Advanced Analytics in Indiana: An Initial Discussion of Industry Needs and University Capabilities.* TEConomy Partners, LLC for BioCrossroads, Jan. 2020.

Figure 4: Biomedical-Related Clusters Identified in the Innovation Landscape Network of U.S. AI Companies Receiving Significant VC Investment, 2014-2018



Source: TEConomy Partners, LLC

New genetic and genomic analytical tools are also in development and coming online, which promise additional high-throughput analytical capabilities and assessment. It is anticipated that new analytical techniques such as CRISPR single cell sequencing and single-cell RNA-seq (scRNAseq) will generate data that provides new insights into biological function.

Case Study: Example of Minable Big Data (Discovery Science)

Geisinger is an integrated healthcare system headquartered in Pennsylvania integrating "primary care and specialists, hospitals and trauma centers, insurance, medical education and research."[60] Geisinger has been at the forefront in terms of recognizing the clinical insights and utility that genomic data and health record data that are mined together provide. Begun in 2007, Geisinger's MyCode Community Health Initiative has seen over 250,000 patients consent to participate in a research program that has been performed in collaboration with Regeneron's genetics research center. More than 100,000 whole exomes had been sequenced (as of reporting in April 2019). Motivated by the MyCode experience, Geisinger launched a "clinical whole-exome population screening program in mid-2018 as part of routine clinical care in a variety of Geisinger clinics, developing an end to end implementation platform—from patient engagement and consenting, to whole exome sequencing in a certified clinical laboratory, to physician education, to genetic counseling at scale, and to integration of clinical results into the electronic health record."[61] The main focus of MyCode is finding and confirming new disease-causing variants (changes) in patient genes, with research programs directed at:

- Searching for changes in genes that protect against disease.
- Targeting new drug development.
- Researching and learning what are the best ways to share medically-actionable genetic results with patient-participants, and then how to facilitate the sharing of that information with other potentially affected family members.
- Translating the results into clinical care.

The CEO of Geisinger, Jaewon Ryu, notes that:

> About 90 percent of the patients we ask let us look at their entire genomes. That's huge trust; a typical rate is 15 percent. Our patients come from an unusually stable population—giving us volumes of data from more than 15 years of electronic health records. MyCode therefore provides unprecedented opportunity for early diagnosis and developing new and tailored treatments, or precision medicine. MyCode is already letting us help participants and their families prevent or mitigate the impacts of some identified genetic risk factors, including cancer and heart disease.[62]

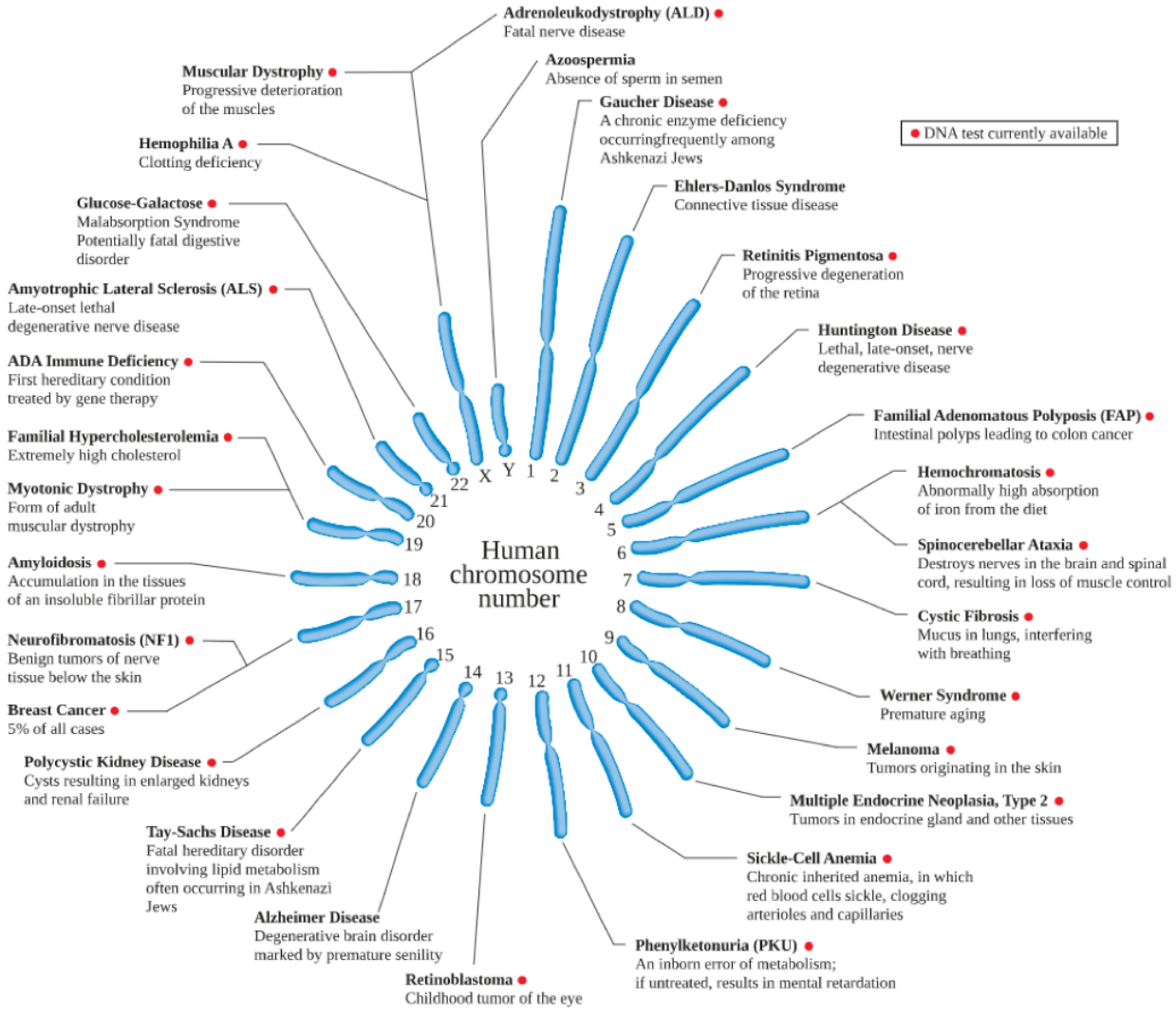
Some of the results emanating from MyCode in terms of advancements in breast cancer genetics have been called out as among the most significant applications of medical genetics in the "Genomic Medicine Year in Review." *BRCA1* and *BRCA2* are genes associated with hereditary breast/ovarian cancer syndrome. They have a high degree of sequence variation, but the population prevalence of "pathogenic" and "more likely pathogenic" variance (P/LP) has not been known. The research team used 50,000+ whole exome sequences from MyCode to examine the frequency of P/LP variants, finding a frequency significantly higher than other estimates. Importantly, it was found that "almost half of all variant carriers did not meet current criteria for clinical testing, and of those meeting testing criteria, nearly half had not undergone clinical testing. Thus, 3/4 of at-risk women were not identified as such and are not benefiting from evidence based interventions; this is a significant care gap with implications for population health."[63] This represents a prime example of how the assembly and analysis of big data enables robust functional impacts to be generated in clinical care.

60 Geisinger. "About Geisinger." www.geisinger.org/about-geisinger. Accessed 12 May 2021.
61 Willard, Huntington. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.
62 Geisinger. "About Geisinger." www.geisinger.org/about-geisinger. Accessed 12 May 2021.
63 Manolio, Teri A., et al. "Genomic Medicine Year in Review: 2019." *The American Journal of Human Genetics*, vol. 105. 5 December 2019. Reporting on findings from Manickam, K. et al. "What We're Missing: Most BRCA1 and BRCA2 Variant Carriers are Undetected." *JAMA Network Open.*

2. Identifying Predisposition to Diseases and Disorders

One of the primary research and clinical applications of human genetics and genomics is its use in identifying potential predisposition to developing diseases and health disorders. The *BRCA* gene example from Geisinger is an example of this, where the identification of these genes in the patient identifies risk for breast cancer and guides clinical decision making. The library of gene-disease associations has expanded rapidly, and as Figure 7 illustrates, there are genes within every human chromosome associated with disease. Notably, all but three of the diseases or disorders listed in Figure 5 now have a genetic test associated with them.

Figure 5: The human chromosome set, indicating examples of locations for pathogenic gene variants causing hereditary diseases



Note: Conditions that can be diagnosed using DNA analysis are indicated by a red dot.
Source: By Ігор Петков - Own work, CC0, https://commons.wikimedia.org/w/index.php?curid=57928376

Figure 5 illustrates only a small sampling of the diseases and health disorders that have genetic associations. The number identified is growing, and at the time of writing 4,395 genes are noted by the Online Mendelian Inheritance in Man (OMIM) project as being identified with a disease or disorder-causing mutation (phenotype-causing mutation), and 6,828 diseases or disorders have a known genetic basis.[64]

One of the key applications of this knowledge, and an extremely valuable one, is the development of genetic testing for predisposition to disease. Genetic testing is:

> *The use of a laboratory test to look for genetic variations associated with a disease. The results of a genetic test can be used to confirm or rule out a suspected genetic disease or to determine the likelihood of a person passing on a mutation to their offspring. Genetic testing may be performed prenatally or after birth.*[65]

The above definition from NHGRI implies two principal uses for genetic tests: 1) diagnosis of present disease (which was discuss previously), and 2) identification of the presence of a gene variant that may predispose an individual or their offspring to the development of a disease associated with that gene. This latter use of genetic testing is becoming increasingly deployed as more gene-disease associations are established. Genetic testing for predisposition to diseases or health disorders may be divided into three categories: 1) carrier screening, which tests a prospective parent for the presence of gene variants that have been shown to be associated with risk of passing down a hereditary disorder (thereby helping to inform family planning and associated decisions); 2) pre-natal and post-natal testing, which focuses on testing for genetic predisposition to disease in the fetus or in newborns; and, 3) child and adult testing. Each application is introduced below.

a. Carrier Screening

As noted by NHGRI:

> *Carrier screening is a type of genetic testing performed on people who display no symptoms for a genetic disorder but may be at risk for passing it on to their children. A carrier for a genetic disorder has inherited one normal and one abnormal allele for a gene associated with the disorder. A child must inherit two abnormal alleles in order for symptoms to appear. Prospective parents with a family history of a genetic disorder are candidates for carrier screening.*[66]

Carrier screening has generated significant impacts on decision making for potential parents. This particularly comes into play when two individuals each carry a single copy of a disease allele that, while harmless to them, has a serious risk of manifesting into a serious genetic disease in their offspring. Carrier screening can be used by individuals, in advance of marriage or long-term partnering, to assess genetic risks, and by couples contemplating starting a family. The application of carrier screening in some communities and populations has had profound effects, limiting, for example, the birth of children with devastating, often fatal, diseases such as Tay-Sachs disease and highly debilitating disorders such as Sickle Cell Anemia and Beta-thalassemia. The rise of low cost, high accuracy whole genome sequencing is greatly expanding the potential to perform DNA-based carrier screening across a broad range of autosomal recessive, single-gene disorders. Carrier screening may also be used by at-risk couples pursuing in vitro fertilization (IVF), allowing their clinician to test for potential genetic abnormalities before implantation of fertilized eggs (this is termed "preimplantation genetic diagnosis") or to select embryos with normal chromosomes for implantation ("preimplantation genetic screening").

64 OMIM®, Online Mendelian Inheritance in Man®. OMIM is a comprehensive, authoritative compendium of human genes and genetic phenotypes that is freely available and updated daily. OMIM is authored and edited at the Johns Hopkins University School of Medicine. For more information see: www.omim.org/statistics/geneMap.

65 National Human Genome Research Institute. "Genetic Testing." www.genome.gov/genetics-glossary/Genetic-Testing. Accessed 12 May 2021.

66 National Human Genome Research Institute. "Carrier Screening." www.genome.gov/genetics-glossary/Carrier-Screening. Accessed 12 May 2021.

b. Pre-natal and Post-natal Testing

Not every potential parent is in a position to access carrier screening and, of course, a great many pregnancies are unplanned. As a result, many (indeed most) pregnancies occur without prior carrier screening. Once a pregnancy is underway, pre-natal screening provides the ability for clinicians to evaluate the healthy development of the baby (through established diagnostics such as ultrasounds and maternal blood tests) and can also include genetic tests to screen for whether the baby may be born with certain genetic conditions and chromosomal disorders (such as Down's syndrome). Such genetic testing has traditionally been reserved for mothers with a certain risk profile (such as family history of genetic disease, mothers who are older, or persons who know they carry certain monogenic alleles that confer risk). Such genetic prenatal testing has required invasive procedures, such as amniocentesis, that carry a measure of risk to the pregnancy. Increasingly, however, physicians are able to order non-invasive prenatal screening (or "cell-free DNA screening") that uses cell-free placental DNA fragments in maternal blood to screen for fetal genetic conditions, such as the common trisomies[67] (e.g., Down syndrome) and deletion or duplication syndromes. A non-invasive prenatal screening (NIPS) is more accurate, with fewer false positives for the most common trisomies, than other screening tests—leading to fewer invasive procedures.[68]

NIPS can now test for:

- Trisomy 21 (Down's syndrome)
- Trisomy 18 (Edwards syndrome)
- Trisomy 13 (Patau syndrome)
- XXY chromosome (Klinefelter syndrome)
- XO chromosome (Turner syndrome)
- Microdeletions in chromosomes
- Rh factor (positive or negative determination)
- Multiple other less common trisomies and single-gene disorders.

At the present time, these NIPS tests are not considered fully diagnostic, and follow-up testing is recommended using other procedures if a positive result is achieved through NIPS. **Ultimately,**

67 A "trisomy" is a condition in which an extra copy of a chromosome is present in the cell nuclei, causing developmental abnormalities.
68 Dondorp, Wybo, et.al. "Non-invasive Prenatal Testing for Aneuploidy and Beyond: Challenges of Responsible Innovation in Prenatal Screening." *European Journal of Human Genetics*, vol. 23, no. 11, 18 Mar. 2015.

pre-natal diagnostics can help mothers make informed decisions regarding their pregnancy and discuss options for care with their health-care provider, help families prepare for a potential challenge to their baby's health, and help ensure that clinicians are ready to support any special health needs of the resulting newborn.

Finally, post-natal testing is a suite of genetic and genomic tests that are employed to **evaluate newborn health and diagnose present or emerging health issues.** As noted by Holm, "the greatest opportunity for lifelong impact of genomic sequencing is during the newborn period."[69] Having a whole-genome sequence, or at least a whole-exome sequence, completed as soon as possible after birth provides a broad spectrum of genetic information for significant immediate use and expanding clinical utility across the lifespan. As noted above, 6,828 diseases or disorders currently have a known genetic basis, and a whole-genome sequence provides an increasingly accessible pathway to evaluating common or rare genetic mutations associated with immediate health challenges or the development or emergent health issues over a life span. The most immediate benefit of newborn screening is as a contributor to achieving a diagnosis of a genetic disease or disorder in the newborn—the advantages of which are discussed in the diagnosis section.

Newborn genetic screening represents a highly visible and successful approach to identification of inherited health conditions.

c. Child and Adult Testing

The clinical reality of whole genome, or whole exome, sequencing is a relatively new phenomenon. Thus, the vast majority of the present U.S. population did not benefit from access to this valuable screening and diagnostic tool at birth. **Having sequencing performed at any stage in life will, however, still have potentially significant clinical utility**, with utility obviously maximized the earlier in life the sequencing is performed. With the expanding library of identified gene-disease linkages and the assurance that this library will continue to grow as more research findings accumulate, it is only a matter of time before full sequencing of everyone is a clinical reality and considered standard of care—with our genome sequence ideally connected to a lifelong electronic health record. The cost/benefit ratio of such a data structure has been consistently shifting in its favor as the cost of sequencing dramatically declines and information on gene-disease associations increases. Indeed, the promise of universal sequencing in clinical application is being realized in some health systems, as shown in the Geisinger MyCode example. Multiple other health systems have large-scale pilot projects underway, including examples such as Mayo Clinic, Intermountain

69 Holm, Ingrid A., et al. "The BabySeq Project: Implementing Genomic Sequencing in Newborns." *BMC Pediatrics*, vol. 18, no. 225, 9 July 2018. doi:10.1186/s12887-018-1200-1.

Healthcare, Mount Sinai Healthcare System, Kaiser Permanente, and the Veterans Administration. It is also increasingly the case that "centers for personalized medicine" have been established at leading healthcare centers that offer whole-genome or whole-exome sequencing to selected patients, often with an initial focus on cancers or rare disorders.

Today, with thousands of genes associated with thousands of diseases, it is perhaps not surprising that genetic and genomic tests are becoming increasingly applied in medicine across the lifespan. A key application is in determining the "risk" for patients in developing a disease that is associated with particular alleles. The previously mentioned *BRCA* gene tests for risk of hereditary breast cancers are one example, with the *BRCA1* and *BRCA2* genes accounting for 20-25% of hereditary breast cancers. Testing positive for these gene variants allows a patient to enter into informed discussions with their physicians regarding potential risk reduction approaches, such as increased screening frequency or prophylactic breast removal surgery. Multiple cancers now have genetic tests associated with them for risk evaluation, including breast, colorectal, cutaneous melanoma, gastric, ovarian, pancreatic, prostate, renal cell, thyroid, and uterine cancers. Cardiovascular-related tests are also available to evaluate risks for developing aortopathies, arrhythmias, cardiomyopathies, genetic forms of high blood pressure and high cholesterol, and thrombophilia. The above are just some of the areas of disease in which genetic testing for risk is seeing application, and the library of available tests will continue to expand.

Because most common diseases have been found to have associations with many individual genes (i.e., they are polygenic, as opposed to monogenic), a new area of science and practice is in development focused on determining potential risk based on the presence of multiple gene variants. This testing results in the generation of "polygenic risk scores" (PRS), which are an "emerging technology for aggregating the small effects of multiple polymorphisms across a person's genome into a single score."[70] It has been noted, "In medicine and public health, PRSs could, in the future, be used for initiating additional risk screening or motivating behavior change."[71] It is an area of research interest and potential promise. Writing in early 2019, Liz Worthy notes that:

> *Over the last 18 months we are seeing increased application of polygenic risk score analysis making use of large GWAS [genome wide association studies] and WGS [whole genome sequencing] data. PRS seeks to estimate an individual's propensity towards particular phenotype… These methods have a variety of uses including human disease risk assessment in research settings and there is increased focus on their application within healthcare settings.*[72]

For the patient, the advantages afforded through the application of genetic tests for disease predisposition are potentially significant—providing a pathway for adopting risk reduction lifestyle or medicinal approaches, more frequent use of early disease detection screenings, and prophylactic surgeries in selected instances. **It is important to note that genetic and genomic testing for the predisposition of disease is best performed in consultation with a patient's physician and with genetic counselors**—skilled personnel who can interpret the results and provide recommendations for health strategies rooted in evidence-based clinical practice. While there are direct-to-consumer tests, there is risk attached to non-professional test result interpretation that could lead to unnecessary anxiety or pursuit of unnecessary/unproven interventions.

70 Ossorio, Pilar N. "Polygenic Risk in a Diverse World." *Science*, vol. 371, no. 6529, 5 Feb. 2021.
71 Ibid.
72 Worthy, Liz. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.

Case Study in Genetic Screening: Detecting At-Risk Carriers Not Detected Through Other Methods

Results from a large cohort study conducted at Renown Health in Nevada show how genetic carrier testing results in the identification of risk in patients where it was not previously suspected. Describing the study, Manolio noted that:

> *The value of genetic screening in an unselected population for identifying individuals carrying P/LP genomic variants for HBOC [Hereditary Breast and Ovarian Cancer], Lynch syndrome, and FH [Familial Hypercholesterolemia] has not been widely explored. The Healthy Nevada Project at Renown Health performed exome sequencing in 26,906 participants with available electronic medical records and analyzed genomic variants in nine risk genes for these conditions. Roughly 1.3%, 90% of whom had not been previously identified, carried P/LP variance. Among carriers, 22%, 70% of whom were diagnosed before age 65, were diagnosed with clinically relevant disease. Less than 20% of carriers had medical record documentation of inherited genetic disease risk or relevant family history.*[73]

The researchers conclude that: "this suggests that genomic screening for inherited cancer and cardiovascular risk conditions can identify a significant number of at-risk carriers who are not detected by standard medical practice and who may benefit from earlier clinical risk screening."[74]

73 Manolio, Teri A., et al. "Genomic Medicine Year in Review: 2020." *The American Journal of Human Genetics*, vol. 105, 5 December 2019. Reporting on findings from Grzymski, J.J. et al. "Population Genetic Screening Efficiently Identifies Carriers of Autosomal Dominant Diseases." *Nature Medicine*, vol. 26, 27 July 2020.

74 Ibid.

3. Diagnosing Diseases, Rare Diseases, and Disorders

Genetic or genomic tests can not only determine potential risk for developing a disease, they can also be highly informative in guiding the diagnosis of a present disease or disorder. As noted by the precision medicine program at Duke University:

> *Whole genome and whole exome sequencing are increasingly being used in the clinic to aid in the diagnosis of rare congenital disorders and solve diagnostic dilemmas. One of the first and most high-profile examples of using clinical sequencing to end a diagnostic dilemma is the case of 6 year old Nic Volker at the Medical College of Wisconsin, who suffered from a mysterious severe bowel disease of unknown origin. Searching for an explanation and treatment, doctors turned to next-generation DNA sequencing technology. They identified a mutation in the XIAP gene as the likely cause of his illness, knowledge that suggested a course of treatment that led to his recovery, ending his diagnostic odyssey. Early results from some clinical sequencing programs estimate the success rate of disease gene identification at about 25-30%, offering hope to thousands of individuals with previously undiagnosed or untreated rare disorders, while recognizing that sequencing will not provide all of the answers.*[75]

75 Duke Center for Applied Genomics and Precision Medicine. "Clinical Whole Genome Sequencing." https://precisionmedicine.duke.edu/researchers/precision-medicine-programs/clinical-whole-genome-sequencing. Accessed 12 May 2021.

Genetic and genomic tests for disease diagnosis are being deployed across a broad range of rare and more common diseases and disorders.

a. Diagnosis of Rare Diseases and Disorders

Rare diseases, by their inherent nature, present diagnostic challenges because so few physicians have encountered them. Often these diseases may present symptoms seen in other, more common diseases resulting in an understandable misdiagnosis and inappropriate treatment strategies being adopted. Patients and their families may embark on long "diagnostic odysseys", seeing dozens of practitioners, undergoing multiple tests and procedures, enduring fruitless attempts at treatment over many years without ever getting a definitive, accurate diagnosis. **Genetic and genomic testing has been a pathway to solving this dilemma in multiple diseases and disorders impacting many thousands of patients.**

While individual rare diseases are, by definition, rare, they collectively impact a large global and domestic population. Liu, Zhu, Roberts, and Tong estimate that:

> *Approximately 7000 rare diseases have been recognized, a substantial number of which are life threatening or chronically debilitating. Around 80% of rare diseases are genetic in origin. A single rare disease affects a small number of the population (defined as <1/15,000 in the US and <1/2000 in Europe)... Most rare disease patients (50 to 75%) show onset at birth or in childhood. As many as 30% of rare diseases patients die before the age of five years. Furthermore, each rare disease patient has been estimated to cost a total of $5 million throughout their lifespan.*[76]

The Genetic and Rare Diseases Information Center reports that 25-30 million people in the U.S. have a rare disease, and over 350 million people worldwide are afflicted.[77] **Approximately 1 in 10 individuals has a rare disease, so collectively rare diseases have a significant population impact.**

Liu, Zhu, Roberts, and Tong further report that:

> *An incomplete knowledge of the natural history of each rare disease can make a substantial proportion (~60%) of rare diseases intractable and undiagnosable. Panel-based NGS or targeted sequencing tests are designed to reveal causal mutations for genes known to be associated with a specific rare disease. Since the NGS gene panel is predesigned or expert-selected, ultradeep, uniform coverage allows for high sensitivity and also for specific variant calling for rare genetic variants.*[78]

By deploying genetic and genomic testing, up to and including whole genome sequencing, diagnostic odysseys may be ended for many patients—not only providing a pathway to appropriate treatment but also reducing significant waste in the healthcare system and the associated costs of incorrect diagnosis. Even if no treatment is available, peace of mind can result through simply having an "answer" and being able to end the costly hunt for diagnosis. In discussing rare diseases and the application of sequencing, Chung, Chau, and Wong report on impressive results from sequencing adoption:[79]

> *Affected individuals often endure years of diagnostic odyssey, which is not only fruitless but more expensive than sequencing their*

76 Liu, Zhichao, et al. "Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic testing in Rare Diseases: Where Are We?" *Trends in Genetics*, vol. 35, no. 11, Nov. 2019.
77 National Center for Advancing Translational Sciences, Genetic and Rare Diseases Information Center. "FAQs About Rare Diseases." https://rarediseases.info.nih.gov/diseases/pages/31/faqs-about-rare-diseases. Accessed 12 May 2021.
78 Liu, Zhichao, et al. "Toward Clinical Implementation of Next-Generation Sequencing-Based Genetic testing in Rare Diseases: Where Are We?" *Trends in Genetics*, vol. 35, no. 11, Nov. 2019.
79 Chung, B.H.Y., et al. "Rare versus Common Diseases: a False Dichotomy in Precision Medicine." *npj Genomic Medicine*, vol. 6, no. 19, 24 Feb. 2021. doi:10.1038/s41525-021-00176-x.

> Whole genomic sequencing and whole exome sequencing are eliminating the phenomenon of the diagnostic odyssey for rare genetic disease: it's realistic today to have a genome or exome test ordered at first subspecialist outpatient visit and to have a diagnosis by the time of the second visit . This is clearly the most powerful diagnostic tool ever developed for the millions of children with rare diseases."
>
> **Stephen Kingsmore. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.**

genomes upfront[80,81]. For infants admitted to intensive care within the first 100 days of life, sequencing produced diagnostic yields of 36.7%; and in 52.0% of the diagnosed, medical management was affected.[82] Results improved to 50.8% and 71.9%, respectively, when trio sequencing was conducted. Other studies have given similar results.[83]

Ranging from individual genetic tests through to complete whole-genome sequencing, the full range of genetic and genomic tools and technologies are now being deployed in clinical diagnostic testing. "Besides targeted sequencing, there are increasing applications of whole-genome sequencing/whole-exome sequencing (WGS/WES) to detect complex genetic variants and provide complete genetic information in support of rare disease diagnosis."[84] This statement recognizes that while many rare diseases may be associated with a single gene (i.e., Mendelian), there are also many challenging diseases where multiple genes come into play (polygenic).

b. Diagnosing Single Gene (Mendelian) Disease and Disorders

Mendelian, single-gene diseases represent an extremely large compendium of diseases and disorders. As noted by Collins, Doudna, Lander, and Rotimi, "The discovery of genes responsible for more than 5000 rare mendelian diseases has facilitated genetic diagnostics for many patients, pregnancy-related counseling, new drug treatments, and in some cases, gene therapies."[85] A number of these single gene-associated diseases and disorders have a substantial impact across populations and within certain sub-populations. Some of the more widely known examples include:

- **Cystic Fibrosis (CF)**—a progressive, genetic disease that causes persistent lung infections and limits the ability to breathe over time. In people with CF, mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene cause the *CFTR* protein to become dysfunctional. CF affects approximately 30,000 Americans.[86]
- **Alpha- and beta-thalassemias**—inherited genetic blood disorders causing the body to

80 Tan, T. Y., et al. "Diagnostic Impact and Cost-effectiveness of Whole-exome Sequencing for Ambulant Children with Suspected Monogenic Conditions." *JAMA Pediatrics,* vol. 171, Sept. 2017. doi:10.1001/jamapediatrics.2017.1755.

81 Farnaes, L., et al. "Rapid whole-genome sequencing decreases infant morbidity and cost of hospitalization." *npj Genomic Medicine*, vol. 3, no. 10, 4 Apr. 2018. doi:10.1038/s41525-018-0049-4.

82 Meng, L., et al. "Use of Exome Sequencing for Infants in Intensive Care Units: Ascertainment of Severe Single-gene Disorders and Effect on Medical Management." *JAMA Pediatrics*, vol. 171, no. 12, 4 Dec. 2017. doi:10.1001/jamapediatrics.2017.3438.

83 Wright, C. F., et al. "Pediatric Genomics: Diagnosing Rare Disease in Children." *Nature Reviews Genetics,* vol. 19, no. 325, 5 Feb. 2018. doi:10.1038/nrg.2017.116.

84 Liu, Zichao, et al. "Editorial: Advancing Genomics for Rare Disease Diagnosis and Therapy Development." *Frontiers in Pharmacology*, vol. 11, no. 598889, 25 September 2020.

85 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine*, vol.384, no. 1, 7 Jan. 2021.

86 Centers for Disease Control and Prevention. "Sickle Cell Disease." www.cdc.gov/ncbddd/ sicklecell/data.html#:~:text=In%20the%20United%20States&text=It%20is%20estimated%20that%3A,every%2016%2C300%20Hispanic%2DAmerican%20births. Accessed 12 May 2021.

make fewer healthy red blood cells and less hemoglobin than normal. Beta-thalassemia affects at least 1,000 people in the U.S.; however, the exact prevalence is not known.[87]

- **Sickle cell disease**—an inherited group of disorders in which red blood cells are misshapen into a sickle shape. The cells die early, leaving a shortage of healthy red blood cells (sickle cell anemia), and can block blood flow. The Centers for Centers for Disease Control and Prevention (CDC) estimate that the disease affects approximately 100,000 Americans.[88]
- **Fragile X Syndrome (FXS)**—a genetic disease that causes mild to severe intellectual disability, with typical associated symptoms including delays in talking, anxiety, and hyperactive behavior. The exact number of people who have FXS is unknown, but a review of research studies estimated that about 1 in 7,000 males and about 1 in 11,000 females have been diagnosed with FXS.[89]
- **Huntington's Disease**—a fatal genetic disorder that leads to progressive breakdown of nerve

"The undiagnosed diseases network (UDN), a multidisciplinary collaboration evaluating patients who have with complex presentations and have remained undiagnosed despite extensive clinical investigation, performed in depth clinical evaluations along with exome and whole genome sequencing, metabolomics testing, and studies in model organisms. From 2015 to 2017, the UDN accepted 601 of 1519 patients referred for evaluation. Of the first 382 patients with a completed evaluation, 132 (35%) received a diagnosis; these included 31 (11%) with new syndromes. Among diagnosed patients, the majority (58%) had medical care changes, such as changes in therapy or shortening of the diagnostic odyssey."

Teri A. Manolio et al. 2019. "Genomic Medicine Year in Review: 2019." The American Journal of Human Genetics, vol. 105, 5 Dec. 2019. Reporting on original research by Splinter, K., et al. *The New England Journal of Medicine*, vol. 379, 2131-2139, 2018.

87 Challenge TDT. "Beta-Thalassemia Overview." www.challengetdt.com/beta-thalassemia-overview. Accessed 12 May 2021.
88 Centers for Disease Control and Prevention. "Sickle Cell Disease." www.cdc.gov/ncbddd/ sicklecell/data.html#:~:text=In%20the%20United%20States&text=It%20is%20estimated%20that%3A,every%2016%2C300%20Hispanic%2DAmerican%20births. Accessed 12 May 2021.
89 Centers for Disease Control and Prevention. "Fragile X Syndrome." www.cdc.gov/ncbddd/fxs/ data.html. Accessed 12 May 2021.
90 Yohrling, George, et al. "Prevalence of Huntington's Disease in the US." *Neurology,* vol. 94, no. 15, 14, Apr. 2020.

Case Study in Whole Genome Sequencing for Diagnosis: Project Baby Bear, California

As noted in *ClinicalOmics*, a pilot program in California funded by the state "showed that precision medicine for critically ill babies enrolled in California's Medicaid program reduced their suffering and yielded better health outcomes, while decreasing the cost of their healthcare, saving the Golden State $2.5 million."[91] The initiative, named Project Baby Bear, deployed rapid Whole Genome Sequencing (rWGS) as the core approach. Stephen Kingsmore, the President and CEO of Rady Children's Institute for Genomic Medicine, which led the project, notes that:

> *For seriously ill children who are hospitalized in intensive care units, the most significant advance has been ultra-rapid whole-genome sequencing. It's routinely possible now to examine nearly every genetic disease and either make a diagnosis or rule out genetic disease, in 36 hours. That's fast enough to guide weighty management decisions in even the most seriously ill children. Where rapid whole-genome sequencing is absolutely transformative is in seriously ill children in whom a genetic disease was not suspected at test order. Those children were, with the best intentions in the world, being treated for the wrong diagnosis.*[92]

In describing the results of the pilot project, Kingsmore reports that the project (which used WGS for Medi-Cal-enrolled infants in intensive care units at five California children's hospitals) had compelling results:

> *In 720 infants in intensive care units tested so far, one in three received the genetic disease diagnosis, and in about 1/3, we are able to exclude genetic disease as the course of illness. One in four infants has a change in care as a result of rapid whole genome sequencing. One in five has a change in outcome.*[93]

A recent review report on the results of Project Baby Bear shows that over 23 months, the project:

- Completed rWGS on 178 babies and families.
- *Provided diagnoses for 76 babies (43%).*
- *Led to a change in the management of 55 babies (31%) that resulted in fewer hospital days, fewer procedures or new therapies.*
- *Diagnosed 35 rare conditions that occur in less than one in one million births.*
- *Achieved a three-day turnaround time for provisional results.*[94]

The study also demonstrated that this clinical application of whole genome sequencing "reduced healthcare costs and downstream spending, primarily by empowering doctors to eliminate unnecessary procedures and discharge babies sooner."[95] In a retrospective analysis of the program's economic impacts, it is concluded that:

> *By introducing Medi-Cal babies into a coordinated system of care that included physicians trained in identifying babies likely to benefit from whole genome sequencing, lab interpretation scientists, genetic counselors and others, the state of California saved millions of dollars in healthcare expenses due to avoided procedures and shorter hospital stays… The avoided procedures and reduced hospital time amounted to $2.5 million in cost savings. These cost savings stemmed from changes in the medical management of just 29 babies who received significant benefit from genome sequencing.*[96]

---

91 "Rady Children's Helps California's Project Baby Bear Improve Outcomes, Save $2.5M." *Clinical Omics Magazine*, 19 June 2020.
92 Kingsmore, Stephen. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.
93 Ibid.
94 Rady Children's Hospital – San Diego. "Project Baby Bear Final Report: Period Covering July 1, 2018 – June 1, 2020." Report to the State of California.
95 Ibid.
96 Ibid.

c. Diagnosing Complex (Polygenic) Diseases and Disorders

Most common debilitating diseases are not caused by mutation of a single gene, rather they are influenced by combinations of mutations in many genes each having a small effect (but combining to have the potential for large effects). Complicating the situation is that the expression, regulation, or products of these genes may be a result of interactions with a multiplicity of environmental factors. There is thus a complex "soup" of genetic and environmental factors at play in many common diseases. While these diseases have a complex etiology that does not mean that progress cannot be made. Indeed, with the sophisticated tools of next generation sequencing and advanced computational analytics, significant biological and mechanistic insights are being produced. As noted in a recent perspective in the New England Journal of Medicine:

> *The discovery of more than 100,000 robust associations between genomic regions and common diseases has pointed to new biologic mechanisms, such as the role of microglia in Alzheimer's disease, autophagy in inflammatory bowel disease, and synaptic pruning in schizophrenia. It has also enabled the development of polygenic risk scores to identify patients at increased risk for heart disease, breast cancer, and other conditions, although additional rigorous testing of such scores is needed, including evaluation of clinical outcomes.*[97]

Multifactorial inheritance disorders are caused by a combination of environmental factors and mutations in multiple genes. For example, different genes that influence breast cancer susceptibility have been found on chromosomes 6, 11, 13, 14, 15, 17, and 22. Some common chronic diseases are multifactorial disorders, with examples including:

- heart disease
- high blood pressure
- Alzheimer's disease
- arthritis
- diabetes
- cancer, and
- obesity.

Eric D. Green, et al. "Perspective: Strategic Vision MedicineNet. "Genetic Diseases (Disorder Definition, Types, and Examples). www.medicinenet.com/genetic_disease/article.htm. Accessed 12 May 2021.

For many common diseases, the challenge of identifying and characterizing genetic effects is not insignificant. In the case of epilepsy, for example, analysis of the ClinGen Epilepsy Gene Curation Expert Panel indicates 2,702 genes associated with epilepsy, a disorder that affects approximately 50 million people worldwide.[98]

> Building on the recent successes of unraveling the genetic underpinnings of rare and undiagnosed diseases, the field is poised to gain a more comprehensive understanding of the genetic architecture of all human diseases and traits. However, myriad complexities can be anticipated. For example, any given genomic variant may affect more than one disease or trait; can confer disease risk or reduce it; and connect additively, synergistically, and/or through intermediates."
>
> **Eric D. Green, et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 October 2020.**

97 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine*, vol.384, no. 1, 7 Jan. 2021.

98 Helbig, I. et al. "The ClinGen Epilepsy Gene Curation Expert Panel—Bridging the Divide Between Clinical Domain Knowledge and Formal Gene Curation Criteria". *Human Mutation*, vol. 39, no. 11, Nov. 2018.

4. Rational Drug Development

Pharmaceuticals (drugs) are a fundamental part of the armamentarium of medicine, providing the means to treat and ameliorate the symptoms of disease, and in some cases, cure the disease. Pharmaceuticals bring relief to millions worldwide and have greatly extended the average human lifespan and quality of life across that lifespan. Traditionally, drug discovery has used a trial-and-error approach whereby a library of chemical substances is tested on cultured cells or animals and evaluated for its effects. Molecules generating an apparent positive effect are then brought forward into clinical trials to evaluate effectiveness on a disease in humans. Rational drug development, however, takes a different approach, one in which biomarkers or preidentified druggable targets that are present in, or generated by, a disease may serve as molecular targets for purposefully designed drugs designed to bind to the target. Human genetics and genomics assist in this approach in multiple ways:

- Helping to identify biomarkers, protein targets, etc. through comparative analysis of disease affected patients versus healthy individuals.
- Identifying genetic variations across individuals impacted by the disease that may influence the effectiveness of a designed drug (often related to differences in metabolism) and potential adverse side effects.
- As noted by Dugger, Platt and Goldstein,[99] sequencing can inform understanding of "the phenotypic effects of a spectrum of rare mutations ranging from loss-of-function to gain-of-function mutations within a single gene" and can provide "information on the putative efficacy and/or toxic effects resulting from the modulation of that particular gene product in humans. This knowledge thereby builds confidence in the rationale for targeting that gene product for the treatment of a more common human disease, rather than relying on information gained from less predictive animal or cellular models."[100]

The application of genetics and genomics to drug development has resulted in multiple clinical successes. Monoclonal antibodies for the treatment of various immune-mediated conditions (targeting the protein interleukin-23 for example, produced by the *IL23R* gene) are already used clinically. Similarly, mutations in the *PCSK9*[101] gene were identified in families with autosomal dominant hypercholesterolemia, and pursuing *PCSK9* as a drug target resulted in the FDA "approving two monoclonal antibodies (alirocumab and evolocumab) for the treatment of high cholesterol not adequately controlled by statins or diet."[102] Put simply, since genes code for proteins, and proteins (and nucleic acids) are typical biomolecular targets for drugs, understanding the relationship between genes and disease provides potential for rationally identifying drug targets. Kalydeco, a targeted drug for cystic fibrosis, approved by FDA in 2012, resulted from rational drug development informed by genomics. Cystic fibrosis is characterized by physical responses to the abnormal flow of salt and fluids in and out of the cells in different parts of the body. Kalydeco specifically "acts on the gating defect associated with the *CFTR* protein [coded by the *CFTR* gene], helping to open up the blocked chloride channels."[103]

It is interesting to note that, in some regards, genetics and genomics advancements have helped to rebalance pharmaceutical research in terms of work on chronic diseases versus rare diseases. Given the prior trial-and-error model, it was in the best interests of the industry to concentrate resources on major chronic conditions in search of blockbuster drugs. Costly and with a high failure rate, drug companies had to triage their funds towards areas with the most promise for financial return. Modern genomics, however, and the large-scale identification

99 Dugger, Sarah A., et al. "Drug Development in the Era of Precision Medicine." *Nature Reviews Drug Discovery*, vol. 17, no. 3, 8 Dec. 2017.
100 Ibid.
101 The PCSK9 gene provides instructions for making a protein that helps regulate the amount of cholesterol in the bloodstream.
102 Dugger, Sarah A., et al. "Drug Development in the Era of Precision Medicine." *Nature Reviews Drug Discovery*, vol. 17, no. 3, 8 Dec. 2017.
103 Ciriello Pothier, Kristen. *Personalizing Precision Medicine. A Global Voyage from Vision to Reality.* John Wiley & Sons, Inc., 2017.

of gene-disease associations has provided a more target-rich environment to address rarer, typically monogenic, disorders[104]—while at the same time showing that many of the more common diseases are highly complex genetically, with sometimes hundreds and even thousands of genes engaged. The result is noted in a recent 2021 study by the Biotechnology Innovation Organization (BIO), which reports that:

> *Throughout the last decade, industry investment and drug development have pivoted towards rare, congenital diseases. Specific examples of clinical and commercial successes have encouraged this transition. Drivers of these successes include targeting molecularly defined causes of disease, regulatory incentives, and favorable reimbursement environments.*[105]

The BIO report provides a review of two longitudinal datasets on drug development, and notes that:

> *One large difference between this 2011–2020 dataset and the previous 2006–2015 iteration is the intensifying focus on rare diseases. Our latest analysis includes 1,256 phase transitions within rare diseases, a considerable increase over the 521 noted in the previous study. This spans 685 different lead developers (not including those listed solely as partners). This indicates that companies view pivoting to rare disease clinical development as a sound strategy.*[106]

The BIO authors note that drug development for rare disorders has had "notably more successful than industry averages—and in particular chronic, highly prevalent diseases." They also note that:

> *A greater understanding of human disease—whether at the molecular or genomic level—ultimately leads to the investigation of personalized medicine. Indications are increasingly segmented by biomarkers in order to match patients with the treatments most likely to show the greatest benefit, according to the underlying drug mechanism and disease pathophysiology. We identified 767 phase transitions out of 12,728 (6%) that incorporated patient preselection biomarkers in their corresponding clinical trial design. This was accomplished by mapping Informa Pharma Intelligence's Biomedtracker and Trialtrove databases, to provide the supplemental level of clinical trial detail.*[107]

This later statement is important—showing that pharmaceutical companies are now able to use genetic and genomic information to target the trials of their biopharmaceutical molecules to patients who have been preselected through the presence of biomarkers (often genetic). This has the potential to advance more drugs successfully towards market since they are more likely to demonstrate efficacy in their trials by virtue of being rationally targeted. The BIO authors conclude that the data they have reviewed "builds confidence in the pursuit of drug development programs targeted at biomarker-enriched patient populations. Such assets are likely to advance through clinical development with lower levels of attrition and should in theory improve patient outcomes via the advent of increasingly personalized medicine."[108] **What is evident is that advancements in human genetics and genomics are not only delivering definitive diagnoses for patients, guiding their care; they are also highly contributory to the development of new therapeutics to treat identified conditions.**

---

104 Rare diseases have often been termed "orphan diseases" in that they represent a class of disease that had not been "adopted" by the pharmaceutical industry.
105 Thomas, David. *Clinical Development Success Rates and Contributing Factors 2011–2020.* BIO, Informa Pharma Intelligence, and QLS, Feb. 2021.
106 Ibid.
107 Ibid.
108 Ibid.

Case Study: Genomics and Rational Drug Development in Cancer

As noted by Collins, Doudna, Lander, and Rotimi: "Studies of cancer genomes have revealed hundreds of genes in which somatic mutations propelled tumor initiation and growth, information that has fueled the development of new drugs."[109] Two long-standing cancer drugs, developed with the help of genetics and genomics, are illustrative:

- *Approved by the FDA in 1998, Herceptin is a drug that targets metastatic breast cancer cells that overproduce the HER2 protein (a product of the HER2 gene). It was the first approved targeted drug based on an individual's genetics, and also came with development of an FDA approved companion diagnostic called HercepTest. Herceptin has had robust results in achieving improved outcomes in patients. Pothier reports that an October 2014 study showed the overall 10-year survival rate for patients to be 84% for those treated with Herceptin and chemotherapy, versus 75% for patients treated with chemotherapy only.*[110]

- *Another example of rational drug development for cancer is Gleevec (imatinib). Chronic myeloid leukemia (CML) is fundamentally a genetic disease caused by mutations in a patient's DNA that translocate genes between chromosomes 9 and 22. The translocated genes generate a hybrid gene that produces a novel protein that causes greatly accelerated production of white blood cells in bone marrow (i.e., a cancer, as defined by uncontrolled cell proliferation). According to the American Cancer Society, CML comprises circa 15% of leukemias in adults. Without a cure, a diagnosis of CML used to mean a best-case survival of 5 years post diagnosis through aggressive cytotoxic chemotherapy with severe side effects. The discovery of the gene translocation causation of CML provided scientists with a target for therapeutic development. Oncologist Brian Druker developed a targeted drug, Gleevec (imatinib) and collaborated with Novartis to bring the drug forward through clinical trials. The drug was a success, effectively blocking the effect of the translocated genes. For the great majority of CML patients, Gleevec is a highly effective treatment—providing a pathway to normal life expectancy through a pill taken once a day with only mild side effects. Gleevec is not a "cure", because the mutated genes are still present, but the drug is highly effective at stopping these genes from causing leukemia. In up to 30% of cases, patients start to develop resistance to Gleevec, but ongoing research has resulted in the development of two alternative drugs aimed at the hybrid gene, Sprycel (dasatinib) and Tasigna (nilotinib) enabling a patient to switch between drugs if resistance develops. Through increasingly sensitive diagnostic tests of BCR-ABL gene expression and blood analysis, clinicians are able to monitor the effect of the administered drug, modify dosage, and switch drugs if resistance development is detected. In effect, the treatment of CML became one of the pioneers in a personalized approach to genetic medicine.*

Other examples are Tarceva (erlotinib) and Iressa (gefitinib) both of which restrict activation of a protein (epidermal growth factor, or EGFR) which is abnormally active in a subset of lung cancers due to mutations in the protein.[111]

The National Cancer Institute (NCI) reports that "as a result of research into the genomic changes associated with cancer, drugs have been developed to fight the disease in several ways:

- *Inhibiting enzymes that trigger the abnormal growth and survival of cancer cells.*
- *Blocking aberrant gene expression characteristic of cancer cells.*
- *Halting molecular signaling pathways that are in overdrive in cancer cells.*[112]

The NCI notes that "these "targeted therapies" specifically combat characteristics of cancer cells that are different from normal cells of the body. This makes them less likely to be toxic for patients compared to other treatments such as chemotherapy and radiation that can kill normal cells."[113]

---

109 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine*, vol.384, no. 1, 7 Jan. 2021.
110 Ciriello Pothier, Kristen. *Personalizing Precision Medicine. A Global Voyage from Vision to Reality.* John Wiley & Sons, Inc., 2017.
111 National Cancer Institute. "Cancer Genomics Overview." www.cancer.gov/about-nci/organization/ccg/cancer-genomics-overview. Accessed 12 May 2021.
112 Ibid.
113 Ibid.

> The ability to tailor a drug regimen to a specific genetic code that is truly personalized to that specific DNA double helix has been a dream of researchers, physicians, and patients alike. Advances in precision medicine, specifically around the genome…are making this dream a reality."
>
> **Kristen Ciriello Pothier. 2017. *Personalizing Precision Medicine. A Global Voyage from Vision to Reality*. John Wiley & Sons, Inc., 2017.**

5. Precision Medicine and Targeted Therapeutics (Pharmacogenetics)

Having an ability to sequence a patient's whole genome rapidly and cost-effectively has opened the door to a new paradigm in healthcare termed "precision medicine" whereby an individual's genetic profile is used to guide decisions made in regard to the prevention, diagnosis, and treatment of disease. As noted by NHGRI, "knowledge of a patient's genetic profile can help doctors select the proper medication or therapy and administer it using the proper dose or regimen."[114]

The discipline of "pharmacogenomics" (also "pharmacogenetics") has grown to be able to deploy genetic and genomic knowledge and tools to help physicians select the "right drug and the right dose" for a patient based on their genome (assuming there is statistically significant clinical information linking a drug to specific gene variants in terms of efficacy and side effects). Pharmacogenetics is an area of research and, increasingly, clinical practice, that addresses the genetically determined variation in how individuals respond to specific drugs in terms of differences in dose requirement, efficacy, and the risk of adverse drug reactions (ADRs).

David Khan, in the Journal of Allergy and Clinical Immunology, notes that:

> *At its most basic, the term pharmacogenetics describes any influence that genetics can have on drug therapy. The newer term pharmacogenomics is often used interchangeably with pharmacogenetics, but there are some subtle differences. Pharmacogenetics mainly deals with single drug-gene interactions. In contrast, pharmacogenomics incorporates genomics and epigenetics to look at the effect of multiple genes on drug responses. Pharmacogenomics is considered the future of drug therapy and is a rapidly growing field in the area of precision (personalized) medicine.*[115]

Pharmacogenetics and pharmacogenomics enable three principal pathways to the improvement of clinical outcomes:

1. Selection of the therapeutic (among multiple choices) that is likely to prove most efficacious based on the patient's genome and a drug's proven efficacy for their specific genotype.
2. Ruling-out a therapeutic (among multiple choices) based on the patient's genome and a drug's potential for unacceptable adverse side effects given their specific genotype.
3. Development of an optimized drug dosage for a patient based on their genotype's influence on the rate at which they will metabolize the drug.

In terms of this last benefit, Namandie Bumpus notes that:

> *Owing to genetics, people can be categorized as poor, intermediate, extensive, or ultrarapid*

114 National Human Genome Research Institute. " Glossary of Genetic Terms: Personalized Medicine." www.genome.gov/genetics-glossary/Personalized-Medicine. Accessed 12 May 2021.

115 Khan, David A. "Pharmacogenomics and Adverse Drug Reactions: Primetime and Not Ready for Primetime Tests." *Journal of Allergy and Clinical Immunology,* vol. 138, no. 4, Oct. 2016. doi:10.1016/j.jaci.2016.08.002.

metabolizers of certain drugs. For example, in the case of a drug that is pharmacologically active and its metabolites inactive, a drug may accumulate in the body of a poor metabolizer and toxicity could occur as a result. By contrast, someone who is an ultrarapid metabolizer of the same drug may not achieve concentrations of the drug in their blood that are high enough to be effective. For prodrugs, where the parent drug is inactive or substantially less active than its metabolite, genetically encoded variation in drug metabolism could affect the ability of a person to activate the drug.[116]

In regard to adverse drug reactions, David Khan points out that:

> *Medications are a cornerstone of the therapeutic armamentarium for most clinicians. The goal of pharmacotherapy is to cure or control a specific condition or disease without causing adverse effects. Unfortunately, adverse drug effects are common and not always predictable. Adverse drug reactions (ADRs) have been defined as reactions that are noxious and unintended and occur at doses normally used in human subjects. ADRs can be related to a number of factors, including known pharmacologic activity of a drug, drug interactions, drug toxicity, and drug hypersensitivity. ADRs are a relatively common cause of morbidity and mortality.*[117]

He further notes that "genetic factors can play a role in pharmacokinetics, pharmacodynamics, and susceptibility to hypersensitivity responses. The degree to which genetics contributes to ADRs is not entirely clear and varies by drug, as well as the type of ADR."[118]

Keeping up with the research literature in regard to pharmacogenetics and pharmacogenomics findings regarding specific drugs has been an expanding challenge for clinicians. However, several leading organizations have been collaborating, and reliable peer-reviewed compendia resources for recommendations have come online. This is highlighted by Mary Relling, who reports that the application of pharmacogenomics to improving healthcare is being codified through "a number of users converging on key peer reviewed, nonprofit curated genomic medicine resources to guide clinical actions, such as ClinVar, ClinGen, the Clinical Pharmacogenetics Implementation Consortium (CPIC), and PharmGKB."[119]

In cancer, tumor profiling is typically performed for patients: with cancer with an unknown primary; with cancer that has not responded to standard treatments; and, to help guide decision-making when there are multiple treatment options. Depending on the clinical situation, the testing method may be a single gene, a gene panel, whole exome, or less commonly, whole genome sequencing.

a. Targeting to Increase Effectiveness

When a healthcare provider is considering prescribing a drug the knowledge of the patient's genotype can now be used in many cases to guide the therapeutic strategy, identify the most effective drug, determine appropriate dosing, and assess the risk of toxicity or other negative side effects. The U.S. Food and Drug Administration (FDA) provides a list of pharmacogenetic associations that it has evaluated and for which it considers there to be "sufficient evidence to suggest that subgroups of patients with certain genetic variants, or genetic variant-inferred phenotypes, are likely to have altered drug metabolism,

116 Bumpus, Namandje N. "For Better Drugs, Diversity Clinical Trials." *Science*, vol. 371, no. 6529, 5 Feb. 2021. doi:10.1126/science.abe2565.
117 Khan, David A. "Pharmacogenomics and Adverse Drug Reactions: Primetime and Not Ready for Primetime Tests." *Journal of Allergy and Clinical Immunology,* vol. 138, no. 4, Oct. 2016. doi:10.1016/j.jaci.2016.08.002.
118 Ibid.
119 Relling, Mary. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.

Table 8: Analysis of FDA Recognized Pharmacogenetic Associations

A1 =	Data support therapeutic management recommendations
A2 =	Data indicate a potential impact on safety or response
A3 =	Data demonstrate a potential impact on pharmacokinetic properties only

Macro Category	Disease or Condition	A1	A2	A3	Sum
Cancer	Cancer - Breast	2	1		**3**
Cancer	Cancer - Leukemia	2	1		**3**
Cancer	Cancer - Colon	2			**2**
Cancer	Cancer - Lung	2			**2**
Cancer	Cancer - Peripheral T-Cell Lymphoma	1			**1**
Cancer	Cancer - Rectal	1			**1**
Cancer	Cancer - Bladder	1			**1**
Cancer	Cancer - Skin	1			**1**
Cancer	Cancer - Kidney		1		**1**
Cancer	Cancer - Soft Tissue Sarcoma		1		**1**
Cardiovascular	Heart Attack or Stroke (blood thinners)	2			**2**
Cardiovascular	Hypertension (high blood pressure)	1	1		**2**
Cardiovascular	Cardiac Arrythmia	1	1		**2**
Cardiovascular	Hypercholesterolemia		1		**1**
Dermatologic	Eczema	1			**1**
Gastroenterological	Nausea and Vomiting (antiemetics)	3			**3**
Gastroenterological	GERD/Acid Reflux	2		1	**3**
Gastroenterological	Ulcerative Colitis	1	1		**2**
Hematologic	Thrombocytopenia			1	**1**
Immunologic	Transplant Rejection (immune suppression)	1			**1**
Infectious Disease	HIV/AIDS	2	1	1	**4**
Infectious Disease	Tuberculosis		1		**1**
Infectious Disease	Bacterial Infections		1		**1**
Metabolic	Obesity	1			**1**
Metabolic	Gaucher's Disease	1			**1**
Muscle	Rare Muscle Disorders	2			**2**
Muscle	Muscle Relaxation for Surgery/Intubation	2			**2**
Neurologic	Convulsions/Seizures	3	2	1	**6**
Neurologic	Pain (treated by anti-inflammatories, narcotics, etc.)	3	1	1	**5**
Neurologic	Nerve Pain (treated by antidepressants)			4	**4**

Macro Category	Disease or Condition	A1	A2	A3	Sum
Neurologic	Tourette's Syndrome	2			**2**
Neurologic	Huntington's Disease (Chorea)	2			**2**
Neurologic	Narcolepsy	1			**1**
Neurologic	Opioid Withdrawal	1			**1**
Neurologic	Multiple Sclerosis	1			**1**
Neurologic	Tardive Dyskinesia	1			**1**
Neurologic	Alzheimer's Disease/Dementia			1	**1**
Neurologic	Insomnia			1	**1**
Psychiatric	Depression	5		5	**10**
Psychiatric	Schizophrenia	6	1		**7**
Psychiatric	ADHD	2			**2**
Psychiatric	Bipolar Disorder	2			**2**
Psychiatric	Obsessive Compulsive Disorder			1	**1**
Psychiatric	Anxiety			1	**1**
Psychiatric	Female Hypoactive Sexual Desire Disorder	1			**1**
Rheumatologic	Rheumatoid Arthritis	2	1		**3**
Rheumatologic	Osteoarthritis	2			**2**
Rheumatologic	Gout		1		**1**
Rheumatologic	Dry Mouth in Sjogren's Syndrome		1		**1**
Urologic	Overactive Bladder/Incontinence		1	1	**2**
Urologic	Kidney Stones		1		**1**

Note: An individual drug may have more than one disease or condition for which it has application. For example, the drug Irinotecan is used for both colon cancer and small cell lung cancer, and the UGT1A1 gene is associated with the metabolism of the drug and thus is of effect for both types of cancers for which Irinotecan may be prescribed.
Source: TEConomy Partners analysis of FDA Table of Pharmacogenetic Associations.

and in certain cases, differential therapeutic effects, including differences in risk of adverse events."[120]

While cancer is perhaps the most well recognized cluster of disease for which genetic tests may impact drug selection and dosing, analysis of FDA data (Table 8) shows that pharmacogenetic associations are in place for multiple chronic diseases and conditions, covering applications in major categories such as cardiovascular disease, gastroenterological diseases and disorders, infectious diseases, neurological diseases and disorders, psychiatric conditions, and rheumatologic disease. Pharmacogenetic associations now span a range from relatively rare diseases such as Tourette's syndrome and Tardive dyskinesia to common conditions such as hypercholesterolemia and depression (with over 50 listed in Table 8). There are more than 100 drugs for which the associations are now listed by the FDA.

120 U.S. Food & Drug Administration. "Table of Pharmacogenomic Associations." www.fda.gov/ medical-devices/precision-medicine/table-pharmacogenetic-associations. 25 Feb. 2020.

Case Study: Pharmacogenetics/Pharmacogenomics and Cancer Treatment Efficacy

Cancer is a genetic disease. As noted by the National Cancer Institute (NCI):

> *Cancer is a group of diseases caused by changes in DNA that alter cell behavior, causing uncontrolled growth and malignancy. These abnormalities can take many forms including DNA mutations, rearrangements, deletions, amplifications, and the additional removal of chemical marks. These changes can cause cells to produce abnormal amounts of particular proteins or make misshapen proteins that do not work as they should. Oftentimes a combination of several genomic alterations work together to promote cancer . Genetic alterations can be inherited from one's parents, caused by environmental factors, or occur during natural processes such as cell division. The changes that accumulate over one's lifetime are called acquired or somatic changes and account for 90 to 95% of all cases of cancer. The field of cancer genomics is a relatively new research area that takes advantage of recent technological advances to study the human genome, meaning our full set of DNA. By sequencing the DNA and RNA of cancer cells and comparing the sequences to normal tissues such as blood, scientists identify genetic differences that may cause cancer. This approach, called structural genomics, may also measure the activity of genes encoded in our DNA in order to understand which proteins are abnormally active or silenced in cancer cells, contributing to their uncontrolled growth.*[121]

The NCI reports that "genomic information about cancer is leading to better diagnosis and treatment strategies that are tailored to patients' tumors, an approach called precision medicine. As a result of research into the genomic changes associated with cancer, drugs have been developed to fight the disease in several ways:

- inhibiting enzymes that trigger the abnormal growth and survival of cancer cells
- blocking aberrant gene expression characteristics of cancer cells, and
- halting molecular signaling pathways that are in overdrive in cancer cells."[122]

It should also be noted that terms such as breast cancer, pancreatic cancer, or lung cancer, for example, denote the location in which the cancer presents in the body, but that does not mean that every occurrence of these cancers is the same. Each form of cancer is typically characterized by there being many cancer subtypes that can vary considerably between patients. The NCI notes that genetics and genomics are proving to be an effective tool in characterizing cancer types and subtypes and elucidating their comparative response to different therapeutic approaches. The NCI reports, for example, that:

> *Cancer genomics research also contributes to precision medicine by defining cancer types and subtypes based on their genetics. This molecular taxonomy of cancer can provide patients with more precise diagnosis, and therefore a more personalized treatment strategy. There are several ways in which the molecular definition of cancer already benefits patients:*
>
> - *Breast cancer is classified based on molecular characteristics into distinct subgroups—Luminal A, Luminal B, triple-negative/basal-like, and HER2 type—that vary in their aggressiveness and respond differently to therapies. Breast cancer patients may benefit from diagnosis and treatment guided by knowledge of their tumor's molecular subtype.*
> - *Diffuse large B cell lymphoma can be divided into the ABC and GCB subtypes by genomic profiling, identifying patients who respond differently to current chemotherapy regimens and molecularly targeted therapies.*

---

121 National Cancer Institute. "Cancer Genomics Overview." www.cancer.gov/about-nci/organization/ccg/cancer-genomics-overview. Accessed 12 May 2021.
122 Ibid.

- *In 2013, The Cancer Genome Atlas project identified four subtypes of endometrial cancer—POLE ultramutated, microsatellite instability (MSI) hypermutated, copy-number (CN) low, and CN high—that correlate with patient survival. This research has already given rise to new clinical trials and investigation of how these subtypes can improve the future of endometrial cancer care.*

- *Lung cancer patients who have a gene fusion involving the ROS1 gene often respond well to treatment with a targeted therapy called crizotinib. In these cases the disease is best defined and treated based on its unique genetic change.*[123]

Elaine Mardis reviewing major advancements in omics sciences notes that:

> *In my opinion, the most significant advances are the increasing numbers of FDA-approved targeted and immunotherapies in cancer, most of which can be correlated to genomic aspects of cancers including specific genes/mutations of known cancer driver genes, and immunogenomic metrics such as increased neoantigen load, microsatellite instability in the setting of mismatch repair defects that predict sensitivity to immune checkpoint blockade inhibitors, or sensitivity to PARP inhibitors in the setting of homologous repair defects.*[124]

Because of its nature as a "genetic disease", cancer has long been on the frontlines in the clinical application of advanced genetics and genomics. This has paid off considerably, such that "today, biomarkers directly connected to drugs, or to crucial outcomes in the human body, allow physicians to identify drugs that are most likely to help a patient, and those drugs can be used to target cancerous cells only, which reduces the side effects that the patient experiences."[125]

123 Ibid.
124 Mardis, Elaine. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.
125 Ciriello Pothier, Kristen. *Personalizing Precision Medicine. A Global Voyage from Vision to Reality*. John Wiley & Sons, Inc., 2017.

b. Reducing Drug Side Effects

According to the U.S. Department of Health and Human Services (DHHS), an "adverse drug event (ADE) is an injury resulting from medical intervention related to a drug. This includes medication errors, adverse drug reactions, allergic reactions, and overdoses."[126] In inpatient settings, the DHHS reports that adverse drug events account for an estimated 1 in 3 of all hospital adverse events, affecting about 2 million hospital stays each year and prolonging hospital stays by 1.7 to 4.6 days. In outpatient settings, adverse drug events each year account for over 3.5 million physician office visits, an estimated 1 million emergency department visits, and approximately 125,000 hospital admissions.[127]

Variations in individual genomes have been found to have significant impact on risk for adverse drug reactions, and there is increasing evidence helping to guide physician decisions regarding which drugs to prescribe, which to avoid, and what dosing should be used to lessen the risk of an adverse event. The previously cited data for FDA pharmacogenetic associations lists several drugs for which genetic tests can identify patients at higher risk for adverse drug reactions. Examples are shown on Tables 9 and 10.

126 U.S. Department of Health and Human Services. "Adverse Drug Events." https://health.gov/our-work/health-care-quality/adverse-drug-events. Accessed 12 May 2021.
127 Ibid.

Table 9: Examples of Drugs with Pharmacogenetic Associations with Adverse Reactions

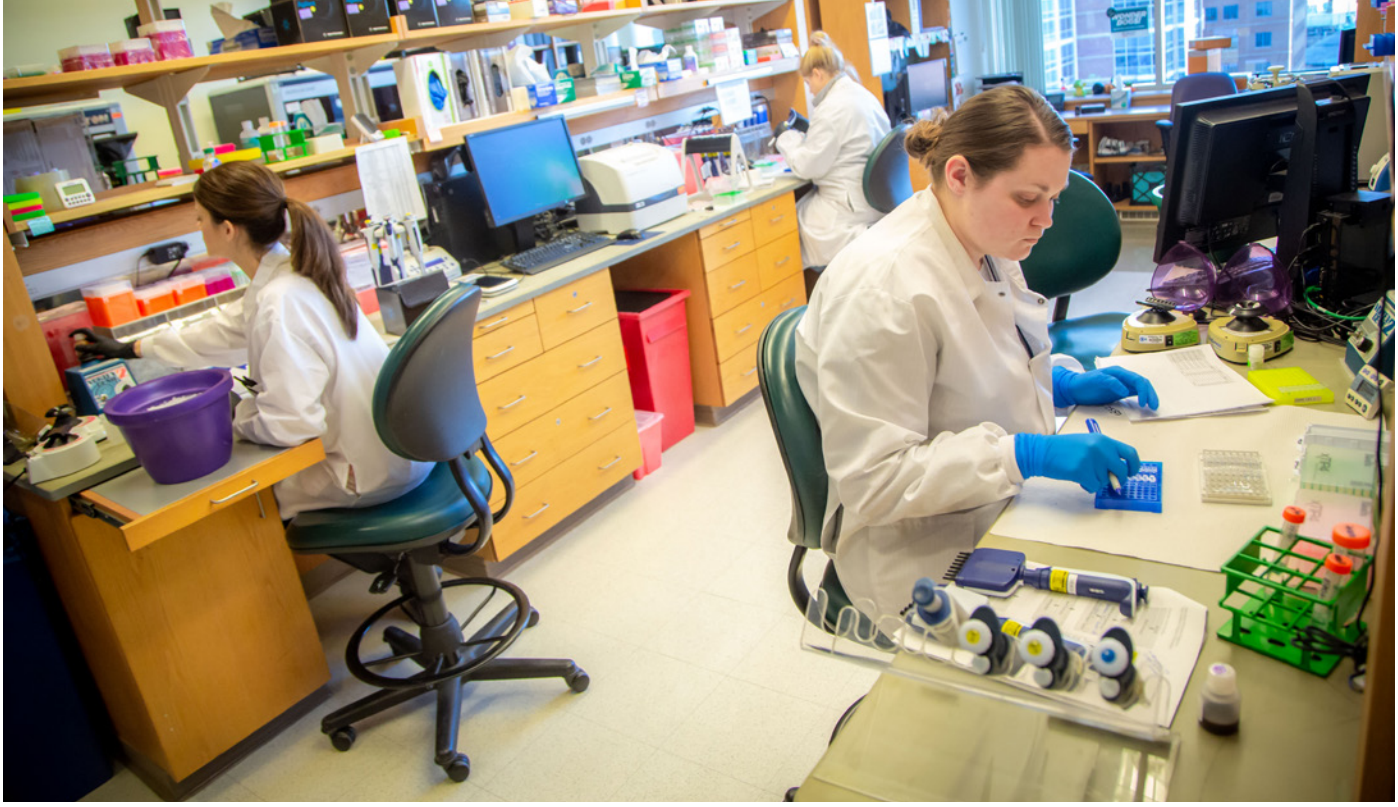
Drug	Disease or Disorder Treated	Gene	Affected Subgroups
Abacavir	HIV/AIDS	HLA-B	*57:01 allele positive
Carbamazepine	Seizures (anticonvulsant), also used in peripheral neuropathy and bipolar disorder	HLA-B	*15:02 allele positive
Lapatinib	Breast cancer	HLA-DRB1	*07:01 allele positive
Simvastatin	Hypercholesterolemia and high triglycerides	SLC01B1	521 TC or 521 CC (intermediate or poor function transporters)
Warfarin	Blood thinner used in patients with risk of heart attack or stroke	CYP4F2	V433M variant carriers

Source: TEConomy Partners analysis of FDA Table of Pharmacogenetic Associations.

Table 10: Examples of Drugs with Pharmacogenetic Associations with Poor Drug Metabolism which may Result in High Systemic Concentration and Associated Adverse Reactions

Drug	Disease or Disorder Treated	Gene	Affected Subgroups
Amifampridine	Rare muscle diseases	NAT2	Poor metabolizers
Amphetamine	ADHD, Narcolepsy, Obesity	CYP2D6	Poor metabolizers
Fluorouracil	Skin cancer and actinic keratosis	DPYD	Intermediate and poor metabolizers
Iloperidone	Schizophrenia	CYP2D6	Poor metabolizers
Tolterodine	Overactive bladder	CYP2D6	Poor metabolizers

Source: TEConomy Partners analysis of FDA Table of Pharmacogenetic Associations.



Case Study: Avoiding Adverse Drug Events in HIV Treatment

Ziagen (abacavir) is a frequently prescribed antiviral for HIV patients. It is an example of an important medication that, unfortunately, has significant adverse effects for a proportion of patients taking it. Approximately 3-5% of patients taking Ziagen are hypersensitive to it and may have significant reactions (including potentially fatal reactions). In the early 2000s, a family of genes were discovered to be associated with Ziagen hypersensitivity (those with the *HLA-B*57:01* gene variant). Genetic testing definitively identifies whether an HIV positive patient has the gene variant, and it has been found that those patients without the variant will be free of the hypersensitivity.

As noted by David Khan:

> *Approximately 3% to 5% of patients treated with abacavir have a hypersensitivity syndrome that typically appears within the first 6 weeks of therapy and rarely can be fatal. Multiorgan symptoms, including fever, rash, gastrointestinal symptoms, respiratory symptoms, and hypotension, can occur along with liver and renal involvement."... "A landmark study was performed to determine whether screening patients with HIV-1 for HLA-B*5701 before treatment with abacavir would reduce the incidence of the hypersensitivity reaction. This study was the first to show in a very rigorous manner the benefits of pharmacogenetic testing in reducing the risk of hypersensitivity reactions. In 2008, the US Food and Drug Administration issued an alert recommending that all patients should be screened for the HLA-B*5701 allele before starting or restarting therapy with abacavir or abacavir-containing medications.*[128]

128 Khan, David A. "Pharmacogenomics and Adverse Drug Reactions: Primetime and Not Ready for Primetime Tests." *Journal of Allergy and Clinical Immunology,* vol. 138, no. 4, Oct. 2016. doi:10.1016/j.jaci.2016.08.002.

6. Gene Editing and Gene Therapy

Much of the important work of genetics and genomics has come in the form of identifying gene variants that are associated with a disease. The identification of a gene variant that codes for a malformed protein, or fails to produce an important protein, or otherwise effects disease etiology, provides potential biomarkers, or targets, for developing diagnostics and, hopefully, therapeutic drugs—very useful tools in the clinical toolkit. But what if instead of treating the effects of a miscoded gene, we could instead correct (edit) the gene itself? If we could do that, then the application of genetics and genomics would not just be to treat the symptoms of a disease, it would potentially cure it (particularly in the case of a monogenic disease or disorder). This is the basis of the development of the field of gene therapy, which the FDA describes as:

> *A technique that modifies a person's genes to treat or cure disease. Gene therapies can work by several mechanisms: replacing a disease-causing gene with a healthy copy of the gene; inactivating a disease-causing gene that is not functioning properly, [or] introducing a new or modified gene into the body to help treat a disease.*[129]

Ultimately, gene editing and gene therapy represent new pathways to the treatment and curing of diseases, but these approaches are still in the early stages of clinical application. Part of the caution in clinical application currently arises from a need for further study of the potential for off-target gene edits (mutagenesis) to occur in non-targeted genes and for unintended mosaicism to occur. Researchers note, however, that the utility of gene editing is not only in its potential clinical application, but also as a powerful tool for investigating gene variation, function, and linkages to diseases and disorders:

> *Advances in DNA synthesis and genome editing allow the field of genomics to progress from a largely observational ('reading DNA') to more experimental ('writing' and 'editing' DNA) approaches. Enabling true synthetic genomics (that is the synthesis modification and perturbation of nucleic acid sequences at any scale) will allow for more powerful experimental testing of hypothesis about genome variation and function and improve opportunities for linking genotype to phenotypes. Genome editing is increasingly being used for practical applications in medicine such as in gene therapy, biotechnology and agriculture.*[130]

There are a variety of types of gene therapy products, including:

- **Plasmid DNA:** Circular DNA molecules can be genetically engineered to carry therapeutic genes into human cells.
- **Viral vectors:** Viruses have a natural ability to deliver genetic material into cells, and therefore some gene therapy products are derived from viruses. Once viruses have been modified to remove their ability to cause infectious disease, these modified viruses can be used as vectors (vehicles) to carry therapeutic genes into human cells.
- **Bacterial vectors:** Bacteria can be modified to prevent them from causing infectious disease and then used as vectors (vehicles) to carry therapeutic genes into human tissues.
- **Human gene editing technology:** The goals of gene editing are to disrupt harmful genes or to repair mutated genes.
- **Patient-derived cellular gene therapy products:** Cells are removed from the patient, genetically modified (often using a viral vector) and then returned to the patient.

U.S. Food & Drug Administration. "What is Gene Therapy." www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/what-gene-therapy. Accessed 12 May 2021.

129 U.S. Food & Drug Administration. "What is Gene Therapy." www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/what-gene-therapy. Accessed 12 May 2021.
130 Green, Eric D., et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." *Nature*, vol. 586, 29 Oct. 2020.

For research, the key advantage of modern gene editing is the ability to precisely change a gene to experimentally examine the effect of the change. Modern technology using CRISPR-Cas9 provides a highly refined way to silence genes and insert genes, and even modify a single letter in the genome. In describing it, Nessa Carey highlights that:

> *A breakthrough in 2012 ripped open the genetic fabric of every organism on this planet, from humans to ants and from rice to butterflies. It's giving every biologist in the world the tools to answer in a few months questions that some scientists have spent half their careers trying to address.*[131]

As a result of both technological and scientific advancement, the application of gene editing has progressed from the benchtop, through animal models, and onwards into human clinical trials and approved clinical therapeutics. Gene editing is still in its early days in terms of clinical use, with issues remaining to be resolved in terms of risk of unintended off-target changes that may happen along with intended changes. More work needs to be done before the full promise of this technology can be realized, but emerging application areas point to gene editing and gene therapies' potential to become important additional clinical tools for addressing genetic diseases and disorders. Collins, Doudna, Lander, and Rotimi capture this optimism, writing that:

> *After years of ups and downs, some dramatic successes of gene therapy are emerging, such as for spinal muscular atrophy and hemophilia. The pace of this research could increase dramatically in the future; precisely targeted genome editing technologies (e.g., CRSPR-Cas9) now provide new avenues to therapeutics... as these technologies continue to mature, it will become increasingly possible to alter cellular genomes efficiently and accurately.*[132]

A Key Consideration

It should be noted that the discussion of gene editing and gene therapy pertains to changing non-hereditable (somatic) genes—that is changes to an individual's genes that will impact the individual but not then be passed to any children they may subsequently have.

There is also ongoing discussion and public debate about the potential use of gene editing for making heritable genetic changes (changes to the germline). Such genome edits would result in changes to an individual's DNA being passed to their progeny and subsequent generations. **At the present time, the general consensus of leading organizations in medical genetics, genetics research, and genetic counseling is that genome editing that culminates in human pregnancy should not be currently undertaken and that further research is required into the scientific, clinical, and ethical implications of germline editing.**

The thousands of rare monogenic diseases and disorders may particularly lend themselves to gene therapy approaches. However, the development of such therapies will, at least early on, be limited to those diseases and disorders that meet some fairly stringent criteria, which are noted by Carey as follows:

> *There is a whole list of key factors. Can you be 100% certain that patients you have diagnosed with the condition all have the same disease? This rules out disorders like schizophrenia where there are probably many different forms of the illness. Do you know exactly how the disease is caused in your patients? This rules out type 2 diabetes where it isn't clear which is the key step in the development of the condition. Do you know what genetic change you need to create? This rules out multiple sclerosis, where we think multiple minor genetic variations interact with the environment to trigger the condition. Can you be sure that making the specific edit you have in mind will prevent or reverse pathology? This rules out Alzheimer's disease. Drug trials*

131 Carey, Nessa. *Hacking the Code of Life. How Gene Editing Will Rewrite our Futures.* Icon Books, Ltd., 2020.
132 Collins, Francis S., et al. "Perspective: Human Molecular Genetics and Genomics – Important Advances and Exciting Possibilities." *The New England Journal of Medicine,* vol.384, no. 1, 7 Jan. 2021.

targeting what we thought was the key pathway failed spectacularly recently, and the companies involved have probably lost billions of dollars as a consequence. Can you get the gene editing reagents to the tissues where they are most needed, in high enough levels? This probably excludes Parkinson's disease as the brain is quite a difficult tissue to access. … Many of the most common and debilitating conditions aren't likely to be good candidates for gene editing anytime soon, because they are too challenging in one or more of these problem areas.[133]

Even with the above challenges, biomedical researchers and clinicians have identified many diseases and disorders that meet the criteria. Significant progress has been made in advancing gene therapies through clinical trials, with examples including:

- **Metachromatic leukodystrophy (MLD)**—a rare storage disorder caused by mutations in the gene coding for arylsulfatase A that results in affected children failing to develop motor skills, typically leading to their death by age 10.
- **Lipoprotein lipase deficiency**—a rare disorder in which "patients as young as two have extremely high levels of cholesterol and suffer recurring, life-threatening bouts of pancreatitis."[134] The drug Glybera resulted and is the first approved gene therapy product in clinical use.
- **Childhood X-linked adrenoleukodystrophy (CCALD)**—is a genetically determined metabolic disorder. Those affected typically experience "normal development until they reach 4–10 years of age, at which time behavioral changes including memory impairment and emotional instability manifest to varying degrees, followed by progressive deterioration of vision, hearing, and motor function."[135]
- **Leber congenital amaurosis (LCA)**—is a rare genetic eye disorder. Affected infants are often blind at birth, and other symptoms may include crossed eyes; rapid, involuntary eye movements; unusual light sensitivity; cataracts; and/or, keratoconus. In 2017, the gene therapy Luxturna (voretigene neparvovec-rzyl) was approved by the FDA to treat children and adults with two mutations in the *RPE65* gene which includes a type of LCA called LCA2.[136]
- **Beta-thalassemia**—is a rare genetic blood disorder that reduces the production of hemoglobin, leading to severe anemia and the need for transfusions.
- **Spinal muscular atrophy**—is a genetic degenerative disorder that starts in the central nervous system and progressively affects all the muscles in the body. The therapy Spinraza is an FDA-approved synthetic antisense oligonucleotide that binds to RNA, which corrects splicing errors in the causative genes. Gene therapy with Zolgensma adds a functional version of the gene.
- **Hemophilia B**—is a rare genetic bleeding disorder in which affected individuals have insufficient levels of a blood protein called factor IX. "The gene therapy etranacogene dezaparvovec substantially increased production of the blood-clotting protein factor IX among 52 patients in the largest and most inclusive hemophilia B gene therapy trial to date."[137]

Gene-based therapies are also being successfully applied in the treatment of selected cancers. An approach proving successful is CAR-T therapy, which

133 Carey, Nessa. *Hacking the Code of Life. How Gene Editing Will Rewrite our Futures.* Icon Books, Ltd., 2020.
134 Reilly, Philip R. Orphan: *The Quest to Save Children with Rare Genetic Disorders.* Cold Spring Harbor Laboratory Press, 2015.
135 Kim, Ji Hyung and Hyon J. Kim. "Childhood X-linked Adrenoleukodystrophy: Clinical-Pathologic Overview and MR Imaging Manifestations at Initial Evaluation and Follow-up." *RadioGraphics*, vol. 25, no. 3, 1 May 2005.
136 National Organization for Rare Disorders. "Rare Disease Database." www.rarediseases.org/rare-diseases/leber-congenital-amaurosis/. Accessed 12 May 2021.
137 The American Society of Hematology. "Gene Therapy for Hemophilia B Found Safe and Effective in First Phase III Trial." *American Society of Hematology,* 8 Dec. 2020, www.hematology.org/newsroom/press-releases/2020/gene-therapy-for-hemophilia-b-found-safe-and-effective-in-first-phase-iii-trial. Accessed 12 May 2021.

is an abbreviation for chimeric antigen receptor (CAR) T-cell therapy. The process is described as follows:

> *First, T cells, a type of immune cell, are taken from a person's blood. Then, in the laboratory, gene replacement therapy is used to add a new gene to T cells. This new gene adds a special receptor, called a chimeric antigen receptor (CAR), to T cells to make CAR-T cells. CAR-T cells are able to bind to and attack certain cancer cells. Large numbers of the CAR-T cells are made in the laboratory, and once a sufficient amount has been produced, the cells are put back into the body to fight certain cancers.*[138]

At its heart, CAR-T cell therapy uses genetics to change a person's own immune cells to recognize and fight cancer cells inside the body. Currently, there are two FDA approved CAR-T therapies used in clinical oncology:

- CAR-T Lentiviral vector, ex vivo, used for acute lymphoblastic leukemia.
- CAR-T Retroviral vector, ex vivo, used for relapsed or refractory large B-cell lymphoma.

7. Human-Microbe Interactions

None of us go through life truly alone. Each of us is host to communities of trillions of microbes, an amount that is considerably larger than the total count of human cells in our bodies. Ed Yong highlights this fact quite effectively in the title of his book *I Contain Multitudes* and notes: "how ubiquitous and vital microbes are: they sculpt our organs, defend us from disease, break down our food, educate our immune systems, guide our behavior, bombard our genomes with their genes, and grant us incredible abilities."[139] As Yong's quote highlights, there is significant biological interaction between the human genome and microbes.

a. The Human Microbiome

Even though our microbiome has a considerable effect on our health, for a report focused on human genetics and genomics it might be asked why one would consider microbes to be within the bounds of this study.

138 Explore Gene Therapy. "Get to Know the Different Types of Gene-Based Therapies." AveXis, Inc. www.exploregenetherapy.com/how-gene-replacement-therapy-is-different. Accessed 3 Feb. 2021.
139 Yong, Ed. *I Contain Multitudes: The Microbes Within us and a Grander View of Life*. Ecco, 2016.

Each microbe has its own unique genome, but this report is focused on the human genome. It is a fair point, but two very interesting findings from recent research into our microbial passengers points to a distinct human genome effect—with impacts going in two directions (microbes impacting our genes and gene expression, and human genotype impacting the make-up of the microbial communities humans' host). In effect, it has been found that humans are, loosely speaking, genetic symbiotes. It is a difficult subject matter to research, requiring access to human sequencing and metagenomic sequencing of the human microbiome, but the data collected by the Human Microbiome Project (HMP) can now be referenced to completed human genome sequences to make important findings. Providing a signpost to potentially interesting genetic interactions is the fact that the gut microbiomes of monozygotic ("identical") twins are found to be significantly more similar than those of dizygotic (non-identical) twins. The findings in human twins confirm findings in mouse models that show that the host (the mouse) genome influences microbiota composition and that host genotype explains a significant proportion of variation in the gut microbiome.[140]

Additional research on whole genome sequencing and microbiome metagenome sequencing of participants in the HMP are compelling, showing that "most microbes are correlated to genetic principal components, especially in stool, but also in oral samples."[141] The authors note that they "identified associations between high level genetic features and various microbiome features; however, the mechanistic forces of those associations remain unclear."[142] It is a very interesting, albeit nascent area of human genomics research, but given the sheer complexity of microbiota and the impact on health of human microbiomes, interesting and clinically relevant future findings are to be anticipated.

b. Infectious Diseases

While many microbes serve important positive life functions for humans, for example aiding digestion and the breakdown of micronutrients, many are pathogenic—the viruses and bacteria that cause infectious diseases. For 2020 and into 2021, COVID-19 has been very much on the minds of all. Clearly, understanding the genetic structure of the SARS-CoV-2 virus has been highly important in the global mission to control the pandemic, and genetics and genomics as disciplines have been on the front-lines contributing to many areas (see sidebar).

In terms of advancing understanding of the effect of the human genome on infectious disease response and susceptibility, it has been serendipitous that the

Genomics in the COVID-19 Pandemic

"Genomics rapidly assumed crucial roles in COVID-19 research and clinical care in areas such as: (1) the deployment of DNA and RNA sequencing technologies for diagnostics, tracking of viral isolates, and environmental monitoring; (2) the use of synthetic nucleic acid technologies for studying SARS-CoV-2 virulence and facilitating vaccine development; (3) examination of how human genomic variation influences infectivity, disease severity, vaccine efficacy, and treatment response; (4) the adherence to principles and values related to open science, data sharing, and consortia based collaborations; and (5) the provision of genomic data science tools to study COVID-19 pathophysiology. The growing adoption of genomic approaches and technologies into myriad aspects of the global response to the COVID-19 pandemic serves as another important and highly visible example of the integral and vital nature of genomics in modern research and medicine."

Eric D. Green, et al. "Perspective: Strategic Vision for Improving Human Health at the Forefront of Genomics." Nature, vol. 586, 29 Oct. 2020.

140 Benson, Andrew K., et al. "Individuality in Gut Microbiota Composition is a Complex Polygenic Trait Shaped by Multiple Environmental and Host Genetic Factors." *Proceedings of the National Academy of Sciences*, vol. 107, no. 44, Oct. 2010. doi:10.1073/pnas.1007028107.

141 Kolde, Raivo, et al. "Host Genetic Variation and its Microbiome Interactions within the Human Microbiome Project." *Genome Medicine*, vol. 10, no. 6, 29 Jan. 2018. doi: 10.1186/s13073-018-0515-8.

142 Ibid.

pandemic hit at a time when widespread human genetic sequencing is relatively inexpensive. While sequencing the virus itself is not "human" genomics, the discipline has been contributing important research that may influence future approaches to diagnosis and treatment of infectious disease. The COVID-19 Host Genetics Initiative, for example, has brought together multiple stakeholders in the human genetics' community to "generate, share, and analyze data to learn the genetic determinants of COVID-19 susceptibility, severity, and outcomes."[143] The Initiative formed to help advance research that could lead to potential discoveries that "could help to generate hypotheses for drug repurposing, identify individuals at unusually high or low risk, and contribute to global knowledge of the biology of SARS-CoV-2 infection and disease."[144] Worldwide engagement in the Initiative has been considerable, with participation of "over 2000 scientists from over 54 countries working collaboratively to share data, ideas, recruit patients and disseminate findings."[145]

Another example is the COVID Human Genetic Effort, which is an international consortium aiming to "discover the human genetic and immunological bases of the various clinical forms of SARS-CoV-2 infection."[146] In particular, the COVID Human Genetic Effort is directing work to search for:

- "Monogenic or digenic inborn errors of immunity (IEI), rare or common, underlying severe forms of COVID-19 in previously healthy individuals, including severe pneumonia, multisystem inflammatory syndrome in children (MIS-C), Long COVID, COVID Toes, etc.
- Phenocopies of these monogenic IEI, such as auto-antibodies neutralizing gene products of loci whose variants underlie these IEI (e.g., auto-antibodies to type I IFNs mimicking inborn errors of type I IFNs).
- Single-gene variants, rare or common, which make certain individuals resistant to the infection by the SARS-CoV2 itself, despite repeated exposure, or resistant to the development of clinical manifestations despite infection."[147]

Multiple U.S. research centers and labs are active participants in both of the above referenced international consortia.

Some interesting research contributions have also been made as a result of the growth of direct-to-consumer genetic testing, where the service of AncestryDNA has built a deep repository of DNA information for its participants. There have been separate genome-wide association studies (GWAS), performed in hospitals with COVID-19 patients that indicated hereditary genetic associations with higher levels of disease impact and infection. However, as noted in a news article by Lakshmi Supriya,[148] these studies have the inherent bias of examining those who are infected, typically with a more severe case (since they are hospitalized). AncestryDNA took a different approach. Recognizing that their deep resource of customer DNA may contain clues to COVID-19 genetic protection or susceptibility associations, they surveyed their customers to capture self-reported information on "exposure, risk factors, symptoms and demographics, most of whom had mild disease."[149] With a large sample of over 700,000 respondents, the AncestryDNA research team had a deep resource to work with, ultimately finding gene associations related to immunity and others associated with susceptibility and disease severity—some of which represent new findings, and some confirming associations identified by other researchers. As Supriya notes, "the study provides a complementary analysis to studies focusing on severe disease, which is promising

143 COVID-19 Host Genetics Initiative. "About." www.covid19hg.org/about/. Accessed 12 May 2021.
144 Ibid.
145 Ibid.
146 COVID Human Genetic Effort. "Our Mission." www.covidhge.com/. Accessed 12 May 2021.
147 Ibid.
148 Supriya, Lakshmi, Ph. D. "Study Finds Protective Genetic Associations with Mild COVID-19." *News Medical,* 29 Jan. 2021, Life Sciences sec.
149 Ibid.

for finding new genetic associations, in particular those that provide protection against the virus."[150]

The degree of susceptibility or immunity to various infectious disease organisms has shaped the human genome throughout our evolution as a species. Historic pandemics, involving large-scale deaths due to smallpox, plague, or influenza, for example, have been "natural selection" events, favoring ongoing reproduction of genotypes with disease resistance traits and down selecting genotypes with high susceptibility to the disease. It is also the case that some of the protective genotypes may serve to help an individual resist certain pathogens, but then be related to negative repercussions also.[151] This has been found to be the case with sickle cell disease where the genes that are associated with the disease also appear to be associated with a positive immunity to Malaria.

Another area of important research at the interface between the human genome and infectious disease is the affect that infection can have on the human genome and upon gene regulation and expression. As noted by the National Institute of General Medical Sciences (NIGMS):

> *When viruses infect us, they can embed small chunks of their genetic material in our DNA. Although infrequent, the incorporation of this material into the human genome has been occurring for millions of years. As a result of this ongoing process, viral genetic material comprises nearly 10 percent of the modern human genome.* [152]

Interestingly, recent research at the University of Colorado Cancer Center by Sharon Kuss-Duerkop and Dohun Pyeon finds that viruses are not just cutting and pasting code within DNA; they are also engaged in "suppressing gene expression via DNA methylation, specifically by targeting DNA methyltransferases (DNMTs)."[153] They note that "viruses elect to turn off genes like *interferon-b* that the immune system regularly enlist to fight foreign pathogens which allows them to replicate quickly and run rampant. This could lead to cancer development."[154]

Human genetics and genomics as a discipline has much more to discover regarding the complex relationship between the human genome and the pathogenic, and potentially positive, effects of microbial infections.

8. Metagenomics and Environmental Genomics

All the above discussion of genetics and genomics research, clinical application, and impacts will have hopefully resulted in an appreciation for the incredible complexity of not just the human genome but the vast network of interfaces between the genome and other biological and environmental systems. Each of us walks a slightly different path through life, experiencing different influences upon our physiology in terms of the food we eat, the amount of sun we expose ourselves to, the environments we are exposed to in our jobs, the pathogens that we by chance encounter, etc. Any and all of these and more may be subtly changing (mutating) letters in our genome or periodically influencing gene regulation or expression. If you want a challenging career, working on unravelling genome-environment interactions and effects would have to be high on the list. There is a specific area of research inquiry in genetics and genomics that provides perspective on the subject—it is termed "metagenomics"

Rob Knight, in considering the future of human genetics and genomics, notes that:

> *Genomics is a key underpinning for metagenomics. This is the case because reference-based approaches are dramatically*

150 Ibid.
151 Pittman, Kelly J., et al. "The Legacy of Past Pandemics: Common Human Mutations That Protect against Infectious Disease." *PLOS Pathogens,* vol. 12, no. 7, 21 July 2016.
152 NIH, National Institute of General Medical Sciences (NIGMS). "Our Complicated Relationship with Viruses." *ScienceDaily,* 28 November 2016.
153 University of Colorado Anschutz Medical Campus. "Here's How Viruses Inactivate the Immune System, Causing Cancer." *ScienceDaily*, 2 March 2018.
154 Ibid.

> Metagenomics. Also known as environmental genomics or community genomics, metagenomics investigates the communal genome contained within an environmental sample. It enables the study of the symbiosis and interactions of organismal genomes and genetic products as a biological system."
>
> **Simon Tripp and Martin Grueber. Economic Impact of the Human Genome Project. Battelle Memorial Institute, May 2011.**

faster and more accurate than reference-free approaches whenever the reference database is complete and correct. However, with a few exceptions (such as bacteria in the human gut of healthy western adults), we are far from having adequate reference data. Sequencing efforts such as the Genomic Encyclopedia of Bacteria and Archaea (GEBA) projects have been extremely valuable in filling in missing branches of the tree of life, but projects such as microbial earth which seeks to sequence all type strains, and 1000 fungal genomes project remain under-resourced. Building these references and augmenting them with new clinical isolates and with isolates from remote human populations and from a panel of diverse environmental samples, such as those provided by the Earth Microbiome Project, could dramatically accelerate progress in all metagenomic studies, whether targeted at human or animal health or at the environment. The benefits that could be achieved would greatly outweigh the modest investment required to complete these studies.[155]

There are countless research questions that metagenomics will be used to pursue, and each will benefit from the speed, accuracy, and affordability of gene sequencing in combination with ongoing advancement in bioinformatics and artificial intelligence-based approaches to the mining of genomic and metagenomic big data.

9. Non-Medical Applications of Human Genomics

Each of the functional impact domains of human genetics and genomics discussed above (domains 1 through 8) have been viewed through the primary lens of medical science—the application of genetics and genomics to understanding genomic structures and mechanisms and their effect on human health and pathology. There are, however, multiple other areas of scientific research and functional application of human genetics and genomics that are not principally directed at medical questions. Three such applications are highlighted briefly below, focused on: forensic science, anthropology and evolutionary biology, and paternity testing.

a. Forensic Science

Genetics, and more recently genomics, has become an essential tool for forensic scientists in criminal justice systems. In the late 1980s, genetic analysis entered forensic use through examination of non-coding region repeats in DNA that are highly variable among individuals. This became known as DNA "fingerprinting," helping to identify crime suspects. The technology was revolutionary for forensic science in that, as noted by Melissa Lee Phillips, "for the first time, forensic scientists could create genetic profiles so specific that the only people who share them are identical twins."[156] Phillips notes that "DNA fingerprint techniques evolved subtly over the next several years, until the polymerase chain reaction (PCR), developed by Kary Mullis, was introduced into forensic work. By

155 Cheifet, Barbara. "Editorial: Where is genomics going next?" *Genome Biology,* vol. 20, no. 17, 22 Jan. 2019. doi:10.1186/s13059-019-1626-2.
156 Phillips, Melissa Lee. "Crime Scene Genetics: Transforming Forensic Science through Molecular Technologies." *BioScience*, vol. 58, no. 6, June 2008. doi:10.1641/B580604.

allowing the selective amplification of any desired stretch of DNA, PCR ushered in unprecedented sensitivity in low-level DNA detection at crime scenes."[157]

Analysis of DNA provides a pathway to definitively identify the individual associated with DNA evidence at a crime scene, and also may be used to establish the identity of human remains. Forensic genetics is an evolving discipline, with new technologies enabling varied use in criminal justice applications:

The U.S. Department of Justice on Advancing Justice Through DNA Technology

DNA can be used to identify criminals with incredible accuracy when biological evidence exists. By the same token, DNA can be used to clear suspects and exonerate persons mistakenly accused or convicted of crimes. In all, DNA technology is increasingly vital to ensuring accuracy and fairness in the criminal justice system.

For example, in 1999, New York authorities linked a man through DNA evidence to at least 22 sexual assaults and robberies that had terrorized the city. In 2002, authorities in Philadelphia, Pennsylvania, and Fort Collins, Colorado, used DNA evidence to link and solve a series of crimes (rapes and a murder) perpetrated by the same individual. In the 2001 "Green River" killings, DNA evidence provided a major breakthrough in a series of crimes that had remained unsolved for years despite a large law enforcement task force and a $15 million investigation.

DNA is generally used to solve crimes in one of two ways. In cases where a suspect is identified, a sample of that person's DNA can be compared to evidence from the crime scene. The results of this comparison may help establish whether the suspect committed the crime. In cases where a suspect has not yet been identified, biological evidence from the crime scene can be analyzed and compared to offender profiles in DNA databases to help identify the perpetrator. Crime scene evidence can also be linked to other crime scenes through the use of DNA databases.

The United States Department of Justice. "Advancing Justice Through DNA Technology: Using DNA to Solve Crimes." www.justice.gov/archives/ag/advancing-justicethrough- dna-technology-using-dna-solve-crimes. Updated 7 March 2017.

Identifying the "Golden State Killer"

For decades, police sought to identify the individual responsible for 12 murders and 45 rapes across California between 1976 and 1986. The police had DNA evidence from crime scenes, but the DNA did not match any individuals in existing criminal DNA databases, and without a suspect there was no way to identify the offender.

The recent growth of ancestral DNA databases provided a pathway to a breakthrough in the case. Police analyzed one of the databases and were able to narrow the DNA to a particular family. Standard investigative procedures were then able to be used to narrow the family members down to one suspect The U.S. Department of Justice – resulting in a confession and conviction.

Forensic genetics is slowly transitioning into forensic genomics... Genomic, transcriptomic, and epigenomic principles, data, and technologies are applied to identify and analyze useful DNA and RNA markers to address various forensic questions that cannot be answered, or only in a limited way, via genetic or other approaches. Human genome data produced with SNP microarray technologies, and increasingly whole exome and whole genome data established via massively parallel sequencing (MPS) technologies, are used to identify DNA markers for individual identification, as well as for appearance and ancestry prediction. The latter is forensically relevant for finding unknown perpetrators of crime who are unidentifiable with standard DNA profiling. Human transcriptome data of various tissues generated with expression microarray technologies, and increasingly with whole transcriptome sequencing via MPS technologies, are used to identify RNA markers to determine the cellular source of crime scene sample. This is forensically relevant for reconstructing the course of events that may have happened at the

157 Ibid.

> Information about the relationships amongst species or populations within species and the time of their divergence from each other can be found in the DNA. It is the job of the evolutionary geneticist to interpret this information from DNA. A subset of anthropology—anthropological genetics—uses the evolutionary geneticist's tool kit to infer human evolutionary history from our and our closest relative's DNA."
>
> **Jason Hodgson & Todd Disotell. "Anthropological Genetics: Inferring the History of Our Species Through the Analysis of DNA."** ***Evolution: Education and Outreach*****, vol. 3, Sept. 2010.**

scene of crime and to support the use of DNA at the activity level of evidence interpretation.[158]

Genetics and genomics are also being used by researchers in the field of criminology to examine genetic correlates to offenders, drawing upon advancements in understanding genetic factors that influence human behavior. An example of this is the work of Eric Connolly and Kevin Beaver that incorporated behavioral genetic methods to "assess gene-environment interplay between anger, family conflict, and violence using a subsample of kinship pairs drawn from the Child and Young Adult Supplement of the National Longitudinal Survey of Youth."[159] Their analysis reveals "a significant shared genetic liability for anger and exposure to family conflict indicating gene-environment correlation" and they conclude that findings from the study "underscore the importance of using genetically informed methodologies to identify underlying risk factors involved in both exposure and response to different forms of strain."[160]

b. Anthropology and Evolutionary Biology

Our DNA codes for us as individuals in the present, but it is also a molecular historical record of our ancestry—providing coded documentation of our lineage (ancestry) and our evolution as a species. Advancements in human genetics and genomics have provided important scientific capabilities that have enabled researchers in evolutionary biology, physical anthropology, archaeology, and associated disciplines to answer many questions regarding our evolutionary biology, our genetic linkages to other species, our population migrations, and our genealogy. Modern genetics and genomics have, for example, enabled substantial advancement in anthropology such that Judith Fridovich-Keil writes: "comparative DNA sequence analyses of samples representing distinct modern populations of humans have revolutionized the field of anthropology."[161]

While genetics and genomics are proving fundamental to advancements in the above-cited academic research disciplines, they have also enabled the development of commercial services that offer genetic ancestry testing (genetic genealogy) to the general population—providing insights regarding ancestry that supplement traditional methods of review of family records and historical documentation. Interest levels have been high, to the extent that two private companies 23andMe and Ancestry now have among the largest repositories of human genetic data in the world. 23andMe, for example, reports that it has "more than 12,000,000 customers."[162] Consumer interest in these services has created a rather rich resource of genetic data that is being used now to advance research, with 23andMe, for

158 Kayser, Manfred and Walther Parson (Editors). "Special Issue: Forensic Genomics." *Gene*, 2017.
159 Connolly, Eric and Kevin Beaver. "Assessing the Salience of Gene–Environment Interplay in the Development of Anger, Family Conflict, and Physical Violence: A Biosocial Test of General Strain Theory." *Journal of Criminal Justice*, vol. 43, no. 6, November–December 2015.
160 Ibid.
161 Fridovich-Keil, Judith. "Human Genome Project." Encyclopedia Britannica Scientific Project.
162 23andMe. "Your DNA is Amazing!" www.23andme.com/about. Accessed 12 May 2021.

example, noting that it has published more than 150 peer-reviewed studies in scientific journals.[163]

c. Paternity testing

While connecting a baby to his or her mother is rather easily accomplished, by the obvious nature of birth—the question of paternity is less obvious. Before genetic tests were available, blood tests and other methods were deployed, but they were less than fully conclusive. Today, however, as noted by the Cleveland Clinic:

> *A DNA paternity test is nearly 100% accurate at determining whether a man is another person's biological father. DNA tests can use cheek swabs or blood tests. You must have the test done in a medical setting if you need results for legal reasons. Prenatal paternity tests can determine fatherhood during pregnancy.*[164]

Determining paternity may be performed simply to inform a father and parenting decisions, but it may also be required as part of a legal process—providing a determinative path to legal rights in child support cases and child custody cases, and also for determining legal rights to Social Security benefits and inheritance. Some applications of paternity testing could also be listed under the previous discussion of medical applications of human genomics because there is utility in establishing paternity for identification of links to genetic conditions and for determining potential compatibility for organ or tissue donation.

D. Summary

Whether for medical or non-medical applications, it is clear that human genetics and genomics advancements provide extremely large-scale benefits across a broad variety of functional impact domains. Genetics and genomics are considered fundamental within modern biological science, providing answers to basic biological research questions, and they underpin a diverse range of applied innovations and applications that are greatly enhancing human health and well-being.

163 Ibid.
164 The Cleveland Clinic. "DNA Paternity Test." https://my.clevelandclinic.org/health/diagnostics/ 10119-dna-paternity-test. Accessed 12 May 2021.

IV. Into the Future

The primary purpose of this report is to provide a current overview, or point-in-time snapshot, of the economic impact of human genetics and genomics in the U.S., and to provide a useful overview of the principal application domains of human genetics and genomics that are generating positive advancements in human health and well-being.

Since the authors first wrote the impact study for the Human Genome Project on its 10th anniversary, genetics and genomics technologies, and the cost-effectiveness of their application, have advanced astonishingly quickly. The literature of fundamental and applied research discoveries in genetics and genomics has expanded equally rapidly, and the application of human genetics and genomics is evident in almost every branch of human medicine and modern biological science.

While the accomplishments of genetics and genomics scientists have been many and varied, with key categories of application highlighted herein, an overused analogy still applies. We are probably just looking at the "tip of the iceberg" in terms of the future of human genomics. There are still multiple large outstanding areas to pursue, including:

- Understanding of the full structure of the human genome and the biological activity it produces or influences.
- Developing additional knowledge regarding the functional relationships between genotype and phenotype and the influence that environmental interactions have on gene expression, regulation, and mutation over lifespans.
- Advancing knowledge of gene-disease relationships, especially (but not exclusively) in regard to common complex, polygenic diseases and disorders.
- Overcoming the current skewing of genomic data towards northern European genotypes by supporting the concerted effort to build more diversity of data across humanity worldwide. This is an important effort to help ensure health disparities are better understood and addressed, and an equitable future secured in the application of genetic medicine advancements.
- There is significant need to translate the research and innovation advancements already being made into much more widespread clinical application, and a distinct need to connect patient genome data to medical records and family history.

Generating predictions for the future of fields of science, technology, and their application is no easy task, especially in areas as large and diverse as those driven by human genetics and genomics. The frontiers of genetics knowledge are constantly expanding, and it is impossible to predict in advance the breakthroughs that may occur that will open-up new pathways to discovery, innovation, and application. CRISPR is a recent example that quite suddenly

and unexpectedly is making available to scientists an exquisitely precise and flexible tool for gene editing that is greatly accelerating progress in both research inquiry and practical application. Similarly, the continued convergence of genetics and genomics with the rapidly advancing field of advanced analytics and artificial intelligence promises a dynamic future.

Researchers have regularly unveiled increasing complexity in human genome functionality and its interactions with other biological systems, and that tendency makes it somewhat challenging to fully predict future status. That said, there are certain observable trends—in sequencing speed, in emerging areas of inquiry, in expanding clinical applications—that point to near-term directionality with some degree of confidence. Several anticipated future areas of advancement are highlighted below.

A. Ongoing Fundamental Discovery

The population of fully sequenced human genomes already encompasses millions of individuals, and further rapid expansion of this universe is to be anticipated. As this report highlights, there are many large-scale sequencing projects presently being conducted around the world. As these datasets are built, they increase opportunities to identify interesting variation across human genomes enabling fundamental questions to be pursued with higher resolution regarding gene functions and the impact of genetic variation on diseases and disorders (many of which are particularly endemic in certain geographic regions and associated regional population subtypes).

One of the key advances to be anticipated is development of an enhanced understanding of genomic variation among diverse population groups spread across the world. The limitations imposed by current data being skewed to genotypes associated with European ancestry will be overcome as sequencing programs in Africa (where the most genetically diverse population is located) and Asia, for example, expand significantly. It may be anticipated that more diversity in the data will reveal interesting findings, as was the case in genes for sickle cell disease being associated with protection against malaria. More diversity in sequenced populations will also be important for further advancing precision medicine, likely unveiling many more mutations associated with drug efficacy and adverse drug reactions.

Computational data sciences are also progressing rapidly, with significant progress being made in advancing automated analysis tools rooted in artificial intelligence advancements. It is likely that computational and analytical sciences will be as important as biological sciences in contributing to a rich discovery environment. A challenge area that will need to be

> One of the biggest challenges is most health care systems are not built to prevent adverse events, but mostly to treat adverse events. Another is the lack of a centralized, organized health care system designed to support life-long results such as genomic testing. For example, we can conduct a preemptive screen for pharmacogenetic tests in a single test, and these results are more likely to be applicable as the patient grows older and is exposed to more high-risk drugs. But we don't have a good system for making these results available when needed. There is no uniform electronic health record. At any one time, patients may have multiple prescribers and pharmacies with little to no coordination, much less with common access to genetic test results that can inform prescribing and capitalize on the availability of clinical decision support. The lack of logical prescribing based on laboratory tests is just one small example of the disconnectedness and lack of computationally informed medicine that impacts all levels of our healthcare."
>
> **Mary Relling. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.**

addressed, given the importance of data analytics, is data access and data interoperability. The ability to analyze data and make discoveries can only occur if scientists have access to data, and there is still much that needs to be done to provide access to health records and other non-genomic data that can be matched to genomic data for analysis. There are also issues in terms of data needing to be formatted and archived in ways that facilitate analysis.

It should also be anticipated that significant advancements in fundamental knowledge will be advanced though studies at a single-cell resolution, assisted by recent advancements in technology and techniques. As noted by Olivier Harismendy:

> *DNA sequencing at deep coverage or at single-cell resolution is revealing a vast genetic heterogeneity of normal or dysplastic tissues. At present these insights are mostly at the stage of observations, but future studies will address the consequences of such heterogeneity in tissue homeostasis and function. The new information that is provided will provide a better understanding of diseases and conditions associated with aging, genotoxic injuries, and the accumulation of such mosaic mutations.*[165]

This is also echoed by Jernej Ule, noting that "we will be able to move beyond the static picture of genomic data towards studies of the dynamic transitions that cells make on a genomic scale in response to external and internal cues."[166]

It should also be anticipated that fundamental and applied research in genetics and genomics will increasingly uncover linkages not just between gene variations and disease but also gene variations and health. Some institutions are already focusing on this opportunity. For example, the Institute for Systems Biology (ISB) in Seattle is a proponent of medicine that can be predictive, preventive, personalized, and participatory (termed P4 medicine), and is working on quantifying wellness—integrating genomics as a component in helping individuals improve their health, longevity, and quality of life. As Lee Hood, the cofounder of ISB notes:

> *Systems biology will revolutionize the practice of health care in the coming decades. Today, medicine is largely reactive. It waits until a person is sick and then treats a disease with*

165 Cheifet, Barbara. "Editorial: Where is genomics going next?" *Genome Biology*, vol. 20, no. 17, 22 Jan. 2019. doi:10.1186/s13059-019-1626-2.
166 Ibid.

varying levels of success. The revolution will emerge from the convergence of systems biology and the digital revolution's ability to create consumer devices, generate and analyze "big data" sets and deploy this information through business and social networks. By providing an understanding of disease at the molecular level, systems medicine will eventually be able to predict when an organ will become diseased or when a perturbation in a biological network could progress to disease.[167]

B. Expanding the Clinical Application of Genomics

It is safe to predict that the application of genetics and genomics in clinical medicine will continue to expand substantially. Currently, there is a large observable difference between genetics and genomics having ubiquitous use in biological research versus a far less uniform application of discoveries and advancements into actual clinical practice.

In the U.S., a key challenge is imposed by the heterogeneous structure of the nation's healthcare system. Rather than a single system, the U.S. comprises a patchwork quilt of individual health systems, together with intermediate insurance organizations and third-party payers, that influence the adoption of established genomics tools and practice advancements. In some locations (for example, central Pennsylvania within the Geisinger health system), genomics is becoming integral to the management of the healthcare of covered patient populations; however, this is far from the norm. There is a long way to go before all patients across the nation have access to state-of-the-art genomics and the personalized medicine and the improved outcomes they enable. The disconnected nature of the U.S. health system structure also hampers effective cascade screening since patients move and their records do not follow-them, and family members may reside in different health systems across the nation. As the National Academies note "the benefits of screening will be multiplied if systems for affective cascade screening can be implemented, but there is currently no roadmap for such testing."[168]

It should also be anticipated that new classes of medicines, developed through synthetic biology and gene editing, will expand in clinical use. At the present time, gene therapies and gene editing see limited clinical application. However, the degree of editing precision provided by CRISPR technology will lead to more widespread therapeutic applications developed using gene editing procedures, which themselves cause editing within the patient's genome.

> "Now that the sequencing technology has advanced to the point where delivering high quality, cheap, and rapid genomes is a commodity – and where robust, reproducible, and accessible methods exist for analysis and interpretation of these datasets – it is clear that the largest hurdle relates to the integration of these types of methods into clinical practice. Tools, methods, and processes need to be developed in order to deliver this information in ways that care providers can digest it and use it for the treatment of their patients."
>
> **Liz Worthy. "Luminaries Share Their Thoughts on Advances in 'Omics Over the Past Five Years." *Clinical Omics Magazine*, vol. 6, no. 2, March-April 2019.**

167 Institute for Systems Biology. "Scientific Wellness." https://isbscience.org/research/ scientificwellness/. Accessed 12 May 2021.
168 National Academies of Sciences, Engineering, and Medicine. I*mplementing and Evaluating Genomic Screening Programs in Health Care Systems: Proceedings of a Workshop.* Washington, DC: The National Academies Press, 2018. doi:10.17226/25048.

C. Educating and Updating Providers

Part of the widespread clinical adoption challenge with genomic medicine relates to the education of medical professionals. Genetics and genomics are complex and fast-moving fields, rendering it difficult to keep physicians up-to-speed in the latest findings and clinical practice implications and recommendations.

It remains to be seen what model for the practice of genomic medicine will predominate. For example:

- There may be development and adoption of computational clinical decision support tools that assist primary care and other physicians in interpreting the results of genetic and genomic tests and guide clinical decision making.
- Physicians may simply be expected to adapt and to educate themselves regarding genomics similar to how they have to access information on new drugs and new practice procedures. Education here would occur via continuing professional education courses, through visits by company representatives, and other traditional pathways.
- Genetics and genomics counselors may become increasingly embedded in large clinical practices as a localized resource, working to keep pace with expanding genetics and genomics advancements and to provide consultation with physicians and other clinical providers.

Because the field is moving quite fast, there will be a need to not only relay new discoveries and advancements to clinicians, but also to revise their knowledge since it is likely that reinterpretation of variant results will occur under evolving evidence and study.

More widespread use of testing for predisposition for disease, and the growth in polygenic risk score systems and other tools, will require physicians to become comfortable in working with patients to interpret results and develop preventive care regimens.

D. Ethical Considerations

Particularly in biological sciences, the frontiers of science may raise ethical considerations. Such is certainly the case in human genetics and genomics where abilities to edit the genome, up to and including hereditary germline DNA, present challenges requiring ethical debate. Should we edit carrier genomes to prevent passing down of genetic disease to progeny, and if so, which diseases should qualify? Some diseases are relatively easily managed with medicines, while some have no treatments at all. Should we reserve gene editing as a last resort for these currently intractable diseases, or should it only be used in diseases that dramatically shorten lives, involve great pain in those afflicted, or impose large-scale economic burdens on society? We will not presume to guess how such issues will be resolved, but it is clear that they are presenting and will require addressing.

Since germline editing impacts our evolution as a species, it is a particularly contentious issue, and it needs to be globally addressed. Currently, more than 40 countries ban germline editing, but there are 195 countries in the world.[169] Somatic gene editing holds significant promise for helping people who are sick and presents less of an ethical challenge (although there are still issues relating to the potential for unforeseen consequences in editing a not fully understood genome).

The widespread collection of individual genetic data also presents privacy issues and potential for genetic discrimination (for example, the risk of a person being treated differently by their employer or insurance company because they have a gene variant that causes or increases the risk of an inherited disorder). The NHGRI sought to identify potential ethical and legal issues pertaining to human genetics as a component of the original Human Genome Project, establishing the Ethical, Legal, and Social Implications of Human Genetics Research (ELSI) program to "examine these issues and assist in the development of policy recommendations and

169 Center for Genetics and Society. "Other Countries." www.geneticsandsociety.org/topics/other-countries. Accessed 12 May 2021.

guidelines to ensure that genetic information is used appropriately."[170] The ELSI program continues to study and address these issues, and NHGRI notes that:

As the ELSI program has evolved, four high priority areas have emerged from its work that serve to categorize domains of ethical, legal, public policy, and societal education that will need to be further addressed. These include:

1. Privacy and Fairness in the Use and Interpretation of Genetic Information

- *Privacy*
- *Discrimination/Stigmatization*
- *Philosophical/Conceptual Assumptions*
- *Public Policy Issues*

2. Clinical Integration of Genetic Technologies

- *Clinical Ethical Issues*
- *Genetic Testing/Counseling*
- *Professional Issues and Standards*

3. Issues Surrounding Genetics Research

- *Informed Consent*
- *Other Philosophical and Ethical Issues*
- *Legal Issues*
- *Ethnocultural Issues*
- *Other*

4. Education

- *Professional-Health*
- *Professional-Other*
- *Public-K through 12*
- *Public-College*
- *Public-Consumer*
- *Combination Professional/Public.*

Equitable access to the benefits of modern human genetics and genomics advancements also represents a challenge to be addressed. As noted in this report, there are concerted efforts underway to increase the diversity of sequencing human sub-populations with multiple large-scale genome sequencing projects taking place in areas of the globe whose populations have been underrepresented in current genomic data. Increasing diversity in the data is important to advancing equitable research and, ultimately, for equity in the clinical applications of genetic and genomic medical innovations that can benefit the full spectrum of humanity.

E. Conclusion

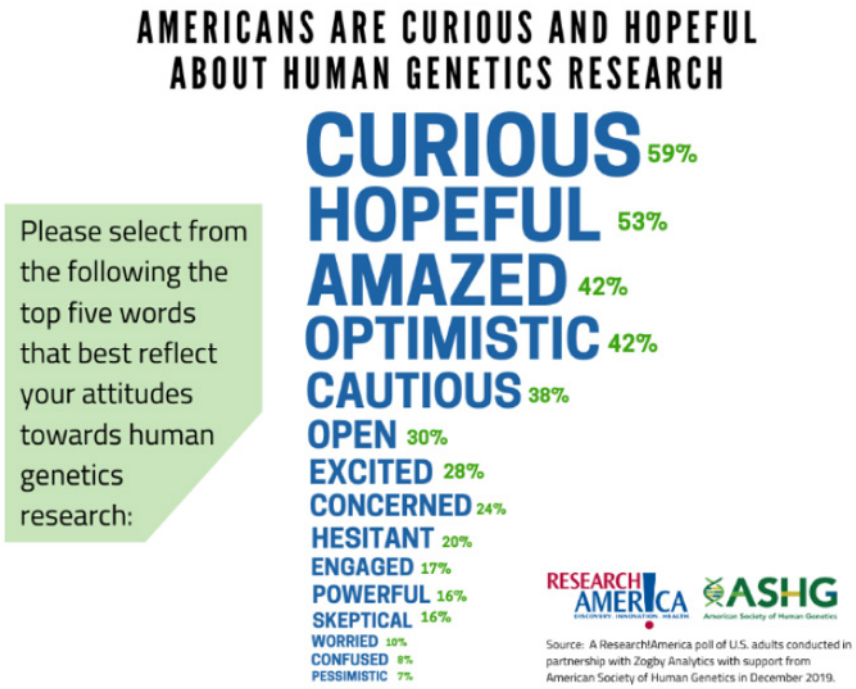
What is absolutely clear is that human genetics and genomics will be in the vanguard in terms of contributing to advancements in medical science and enhancing the practice of clinical medicine. Expanding understanding of the human genome, variations in genomes, external factors that interface with the genome, and genetic relationships to health and disease will provide improved health, quality of life, and large-scale benefits to society (both economic and social). Indeed, as this report finds, it already has.

It is heartening to note that a survey of the general public in the U.S. found that people generally share an optimistic view of genetics and genomics and the promise it holds for a better future. A study released by the American Society of Human Genetics in partnership with Research!America found that the "large majority of Americans agree that genetic knowledge will be important to their own health and their families' health."[171] The study also found that "a strong majority of Americans (77%) indicate positive feelings about human genetic research". Figure 6 summarizes other topline survey findings.

170 National Human Genome Research Institute. "Review of the Ethical, Legal and Social Implications Research Program and Related Activities (1990-1995)." www.genome.gov/ 10001747/elsi-program-review-19901995#:~:text=The%20original%20issues% 20identified %20 in,impact%20of%20genetic%20information%20on. Accessed 12 May 2021.

171 *ASHG Survey Finds Americans Strongly Support Human Genetics Research.* Research!America and American Society of Human Genetics, 29 Jan. 2020.

Figure 6: U.S. Adults Attitudes Regarding Human Genetics Research – Survey Findings



Source: American Society of Human Genetics and Research!America 2019. "ASHG Survey Finds Americans Strongly Support Human Genetics Research."

Based on the economic and functional impacts of human genetics and genomics detailed herein, the public is right to feel optimistic about human genetics research. The fields of human genetics and genomics are having profound positive impacts both in terms of biomedical discovery as well as within the clinical practice of medicine—working to improve outcomes for millions of patients and demonstrating great promise for future highly positive contributions to human health and well-being worldwide. While generating these positive functional impacts is the *raison d'etre* for pursuing the advancement of human genetics and genomics, the fields have also had the very positive spillover effect of building a powerful science- and technology-based economic sector for the U.S.—a sector that supports 850,263 jobs across the nation and generates $265.4 billion in economic output. Human genetics and genomics innovation is expanding the stock of knowledge upon which the nation's continued advancement depends and shows great promise to continue to do so long into the future.

Glossary of Terms

Term	Definition
Allele	A variant of a gene inherited from one parent. An allele's frequency in a population can change due to four forces: mutation, natural selection, random genetic drift, and gene flow.
Biomarker	Any substance, structure, or process that can be measured in the body or its products that influence or predict the incidence of outcome or disease
Cancer Genomics	The study of the DNA sequence and gene expression in tumor cells as they compare to normal host cells.
Clinical	The observation and treatment of actual patients rather than theoretical or laboratory studies.
DNA	Deoxyribonucleic acid, the molecule that carries genetic instructions in all living things. The DNA molecule consists of two strands that wind around one another to form a double helix. Each strand has a backbone made of alternating sugar (deoxyribose) and phosphate groups. Attached to each sugar is one of four bases: adenine (A), cytosine (C), guanine (G), and thymine (T). The sequence of the bases along the backbones serves as instructions for assembling protein and RNA molecules.
Functional Genomics	The study of how elements in the genome contribute to biological processes.
Gene	The basic physical and functional unit of heredity. Technically a distinct sequence of nucleotides forming part of a chromosome, the order of which determines the order of monomers in a polypeptide or nucleic acid molecule which a cell (or virus) may synthesize.
Genotype	The genetic constitution of an individual organism.
Germline Cells	Germline cells pass on their genetic material to the progeny. These include sperm and egg cells.
Microbiome	The collective genomes of the microbes (composed of bacteria, bacteriophage, fungi, protozoa, and viruses) that live inside and on the human body
Mosaicism	Mosaicism is when a person has two or more genetically different sets of cells in his or her body
Pharmacogenomics	The study of the role of the genome, or multiple genes, in predicting drug metabolism and response.

Term	Definition
Protein	A molecule made up of amino acids. Proteins are needed for the body to function properly. They are the basis of body structures and of other substances such as enzymes, cytokines, and antibodies.
Somatic Cells	All cells in the body that are not germline cells
Synthetic Biology	An interdisciplinary field that involves the application of engineering principles to biology. It aims at the (re-)design and fabrication of biological components and systems that do not already exist in the natural world.
Systems Biology	The holistic study of the interactions and behavior of the components of biological entities, including molecules, cells, organs, and organisms.
Whole Exome Sequencing (WES)	Identification of the sequence of base-pairs in the protein-coding regions of the genome.
Whole Genome Sequencing	Identification of the sequence of base-pairs across the full genome, including protein-coding and regulatory regions.

Appendix—
Additional Economic Impact Information

Table 11: Listing of NIH Genetic and Genomic-Related Proposal Review Study Sections

- Behavioral Genetics and Epidemiology Study Section [BGES]
- Cancer Genetics Study Section [CG]
- Gene and Drug Delivery Systems Study Section [GDD]
- Genes, Genomes, and Genetics [F08]
- Genetic Variation and Evolution Study Section [GVE]
- Genetics of Health and Disease Study Section [GHD]
- Genomics, Computational Biology and Technology Study Section [GCAT]
- Molecular Genetics A Study Section [MGA]
- Molecular Genetics B Study Section [MGB]
- Molecular Neurogenetics Study Section [MNG]
- Therapeutic Approaches to Genetic Diseases Study Section [TAG]

Table 12: Economic (Expenditure) Impacts— Additional Expanded NIH Funding Scenario ($7.018 billion)

Impact Type	Employment	Labor Income ($B)	Value Added ($B)	Output ($B)	State/ Local Tax Revenues ($B)	Federal Tax Revenues ($B)
Direct Effect	181,595	$23.31	$54.78	$112.04	$2.95	$5.48
Indirect Effect	303,139	$25.83	$45.15	$89.86	$3.00	$5.59
Induced Effect	418,167	$23.69	$41.71	$74.18	$3.88	$5.23
Total Effect	902,902	$72.84	$141.64	$276.08	$9.83	$16.31
Multiplier	*4.97*	*3.12*	*2.59*	*2.46*		

Source: TEConomy Partners analysis of Human Genetics and Genomics Input Dataset; IMPLAN 2019 U.S. Impact Model

Table 13: Economic (Expenditure) Impacts— Additional Use as Tool NIH Funding Scenario ($14.202 billion)

Impact Type	Employment	Labor Income ($B)	Value Added ($B)	Output ($B)	State/ Local Tax Revenues ($B)	Federal Tax Revenues ($B)
Direct Effect	210,493	$26.37	$58.85	$119.23	$3.05	$6.04
Indirect Effect	329,541	$27.72	$48.03	$95.07	$3.16	$5.98
Induced Effect	460,234	$26.07	$45.91	$81.64	$4.27	$5.76
Total Effect	1,000,268	$80.16	$152.78	$295.93	$10.48	$17.78
Multiplier	*4.75*	*3.04*	*2.60*	*2.48*		

Source: TEConomy Partners analysis of Human Genetics and Genomics Input Dataset; IMPLAN 2019 U.S. Impact Model

ASHG and the project authors wish to thank the following organizations for their generous support of this study.

INVITAE

REGENERON
SCIENCE TO MEDICINE®

TECONOMY
PARTNERS LLC

ASHG
American Society of Human Genetics