29 Genetics and NCDs

Murielle Bochud, Ambroise Wonkam, Pascal Bovet, Vincent Mooser

The risk of acquiring (or being protected from) NCDs arises from the complex interplay between the environment that an individual lives in (e.g. work and living conditions), patterns of behaviour (e.g. tobacco use, diet, physical activity) and genetic makeup. Unlike physical and behavioural characteristics, an individual's germline genetic makeup does not change over time, which implies that it can be analyzed once in a person's lifetime. Advances in sequencing technologies and IT, associated with a dramatic drop in costs over the last few decades have resulted in sequencing of the first complete human genome in 2001. This spectacular achievement has prompted large efforts which have led to the elucidation at the molecular level of many NCDs and improved risk prediction, as well as better diagnosis and treatment.[1] These discoveries have now implications for developing preventive or treatment strategies tailored to the individual (personalized or precision medicine).[2,3]

BOX 29.1 COMMONLY USED TERMS IN GENETICS

Genomics. The study of the genome of a human or another organism (e.g. bacteria of the gut). Almost every cell of the human body contains a complete copy of the genome. The genome contains all the information needed for a person to develop and grow. The human genome consists of about 3 billion base pairs, with more than 99% of those bases being identically shared across people, and includes ~21,000 protein-coding genes and ~20'000 non-coding genes. Both coding and non-coding genes can be involved in human diseases.

Gene. A DNA sequence that controls the expression of a single protein.

Gene modifications and variants. Most DNA modifications are rapidly repaired but some of them persist and are transmitted across generations. A single nucleotide polymorphism (SNP, pronounced 'snip') refers to a substitution of one single nucleotide (building block of the DNA)

DOI: 10.4324/9781003306689-32

at a specific position in the genome. The vast majority of genetic variants do not lead to diseases. Germline variants are present in sperm, eggs and their progenitor cells and are therefore heritable, while somatic variants occur in other cell types (including tumour cells) and are not inheritable.

Epigenome. Chemical modifications that can affect genetic regions and influence gene expression by turning genes on or off, thereby controlling the production of proteins in cells and tissues during the lifetime of a person in response to environmental exposures or disease processes. Much of the epigenome is reset when parents pass their genomes to their offspring.[4]

Genetics to understand NCD aetiology

All diseases have a genetic component. The contribution of genetics *vs* environment varies between diseases. *Monogenic* diseases are due to one single faulty gene and are rare. Around 7,000 primarily monogenic diseases are known, including familial hypercholesterolemia, cystic fibrosis, sickle cell anaemia, Huntington's disease, polycystic kidney disease or haemophilia A. Common diseases usually result from the cumulative effect of numerous genetic variants with small effects. They are called *polygenic diseases* and include type-2 diabetes (T2D), obesity, ischaemic heart disease (IHD) and several cancers. Common variants can be detected using genotyping and analyzed through genome-wide association studies (GWAS).[5]

Genetics to support the molecular diagnosis of NCDs

For a variety of rare diseases, having a molecular diagnosis greatly helps in shortening the time to definite diagnosis (i.e. 'diagnosis odyssey') and in defining the optimal healthcare, prevention and treatment, as well as family planning. As an example, the presence of deleterious mutations within the gene encoding the LDL-cholesterol receptor (which controls the clearance of LDL particles from the bloodstream) establishes the diagnosis of familial hypercholesterolemia (FH), which is responsible for early-onset IHD. Similarly, assessing mutations in the *BRCA1* or *BRCA2* genes helps diagnose selected breast cancer cases or identify individuals at high risk of developing this condition. Detection of such mutations has implications for preventive measures, diagnosis and treatment.

Genetics to predict the risk of NCDs

Family history is easy, but an often underused tool to point to understand patterns of hereditary conditions in an individual or family. A strong family history

may suggest an underlying genetic cause, its mode of inheritance and related genetic risk, but equally a common exposure to external causes or environment among those affected. Genetic tests can confirm the genetic link to a disease where resources allow. A genetic predisposition (or genetic susceptibility) results in an increased likelihood of developing a particular disease as a result of a person's genetic makeup. *Polygenic risk scores* (PRS, i.e. genetic scores based on thousands, often hundreds of thousands, of variants associated with a specific disease) have been constructed for a variety of NCDs, including IHD, T2D, obesity and several cancers.[6] For example, individuals who have a PRS within the upper decile are exposed to a two-fold increased risk of developing CVD compared to people within the other nine deciles, *independent* of the effect of other conventional risk factors.[7] Similarly, a PRS was found to have the same high and independent impact as an unhealthy diet for T2D risk.[8] Beyond well-defined Mendelian risk for single gene disorders, PRS for complex diseases might be increasingly considered, in the near future, as useful tools to assess risk in complement to conventional risk factors, particularly for CVD (for which treatment is often based on total CVD risk),[9] but also for diabetes and cancer, to guide risk reduction counselling and treatment. As genetic makeup is already present at birth, early risk prediction of NCDs is feasible at a very early age;[10] still, a number of ethical issues need to be considered before embarking upon using this approach.

Genetics to support NCD prevention

The identification of a monogenic condition or a high PRS for an NCD provides the opportunity for early prevention, aiming at delaying the onset of this NCD, reducing its severity and extending years of life without disability, cascade retrospective testing in families and genetic counselling for reproductive options. For example, a person who knows having an increased genetic risk for a particular disease may engage more actively in a healthier lifestyle to mitigate the increased risk,[11,12,13] although the public health utility and cost-effectiveness of this effect need to be assessed further. Preventive mastectomy starting at the age of 25–30 years is an option for carriers of selected *BRCA1/2* variants, as these individuals have a lifetime risk of developing invasive breast cancer as high as 60–85%.[14] Careful counselling is required ahead of undergoing such genetic tests for patients, as well as for their relatives.

Genetics to optimize NCD treatment

Genes encoding proteins involved in the absorption, metabolism, distribution and excretion (ADME) of drugs are called 'pharmacogenes'. The purpose of pharmacogenetics is to understand how genetic variation affects treatment outcomes, with the intent of guiding therapy. Currently, >30 pharmacogenes have been identified that may be useful to adjust treatment (e.g. to adjust the dose of usual medications or to prescribe drugs that act on particular genetic

pathways, such as *CYP2D6*, *CYP3A5*, *G6PD*). There are important differences in the frequencies of variants between ethnogeographic groups,[15] which means that treatment approaches need to be tailored for different populations. It is estimated that genomic variation can account for 20–95% of therapeutic effects.

Examples of public health and clinical implications of genetics in relation to NCDs are summarized in Table 29.1.

Genetics to cure NCDs

The increased knowledge of causes of single genes disorders or factors that affect their severity, paired with the development of gene editing and gene therapy technology, has opened new prospects for curative treatment for genetic conditions. For example, a highly successful gene-editing strategy for treating individuals with sickle cell disease is the induction of foetal haemoglobin through the manipulation of a gene that controls its production (e.g. *BCL11A)*.[16]

Databases and technological advances in genomics

As costs of analyzing genomic variants have decreased dramatically in recent years (e.g. <US$ 500 for whole genome sequencing or <US$ 50 for genotyping hundreds of thousands of common variants), information on the links between genetic data and diseases has accumulated exponentially. Many online population-based genetic databases are widely available, enabling us to aggregate and harmonize genomic variation (e.g. gnomAD), assess genotype-phenotype relationships (e.g. OMIM, GWAS catalogue), examine the functional and clinical relevance of genetic variants (e.g. ClinVar, ClinGen) or assess tissue-specific expression (i.e. how a gene defect translates into altered RNA and subsequent proteins) and epigenomics. Initiatives are also developed in low- and middle-income countries.[17,18] These genetic databases have greatly accelerated the transfer of knowledge from the laboratory into the clinical setting. Tools to analyze large-scale genetic data have also evolved, including machine learning and artificial intelligence (AI). For example, the open-access UK Biobank has generated >90 million testable genetic variants that can be explored for their associations with numerous phenotypes, including linkages with routinely collected data in health medical records.[19]

The molecular mechanisms of several diseases are being uncovered thanks to sharing data from transnational collaborative large-scale population-based cohorts that collect whole genome sequencing and other omics data. An implication of these rapid advances in medical genetics is that it may become possible to fairly inexpensively perform systematic genome sequencing at birth or an early age among all or groups of individuals, for early identification of risk of certain NCDs (and other diseases). Beyond enabling targeted prevention and diagnosis for persons who undergo such tests, data from large genetic databases

Table 29.1 Examples of public health and clinical implications of genetics in relation to NCDs

	Domain	*Examples*	*Current applications*	*Future applications*
Understand NCD aetiology.	Analyses of genetic variants across the genome in large populations for research.	Identification of rare coding variants responsible for monogenic forms of NCDs and common variants contributing to polygenic forms of NCDs (e.g. 240 regions in the genome are associated with T2D).	Essential to understand disease mechanisms at the molecular level and to identify and validate new drug targets.	Develop new therapeutics. Develop new diagnostic tests.
Establish molecular diagnosis of an NCD.	Specific genetic analyses in the clinic.	FH (e.g. *LDLR, PCSK9, APOB)*, breast cancer (e.g. *BRCA*), diabetes (e.g. *GCK, HNF1A*). Familial adenomatous polyposis (e.g. *AFAP*, *HNPCC*).	Molecular diagnosis in familial/atypical forms of selected NCDs.	Systematic genome sequencing at birth.
Predict risk of an NCD.	Family history.	Breast cancer, FH.	May point to familial forms of NCDs and to indication for genetic analyses.	Systematic genome sequencing at birth or early ages.
	Analyses of selected genes.	FH (e.g. *LDLR, APOB, PCSK9*), breast cancer (e.g. *BRCA*), colon cancer (e.g. *AFAP*, *HNPCC*).	Test relatives in cascade screening for clinically actionable mutations.	Systematic genome/candidate genes sequencing in people diagnosed with NCD to identify those with a monogenic form (e.g. MODY, FH).
	Polygenic risk scores.	IHD, T2D, obesity, breast or colon cancer.	None yet.	Systematic analysis of PRS for NCDs in everybody or otherwise high-risk individuals for targeted prevention.

NCD prevention.	Analysis of common and rare genetic variants.	*BRCA* for breast cancer; *LDLR, APOB, PCSK9* for FH; *AFAP, HNPCC* for colon cancer.	Preventive mastectomy in carriers of specific mutations. Intensive lipid-lowering therapy in FH. Early screening of cancer.	Cutting/replacing gene(s) (e.g. CRISPR). Vaccination (tobacco smoking, dyslipidaemia, hypertension).
Optimize NCD treatment.	Genetic analysis of pharmacogenes (genes that modify the impact of medicines).	*CYP3A5* (alters the metabolism of drugs, e.g. nifedipine). *BRCA* (implying specific immunotherapy), etc.	Helps titrate medicines and/or select specific medicines or assess adverse drug reactions and drug interactions.	Systematic analysis of pharmacogenes, embedded in electronic medical records and prescriptions.

CRISPR: a genetic engineering technique in molecular biology by which the genomes of living organisms can be modified.

will advance knowledge on disease diagnosis and classification, the development of new diagnostic tests and new treatments. For example, a systematic analysis of a patient's pharmacogenes, embedded through algorithms or AI procedures in patients' electronic medical records, could help guide prescriptions for some patients with an NCD.

Implications of integrating genomics into healthcare

As benefits from genomics are increasingly demonstrated, the demand for integrating genomic medicine into routine healthcare (for NCDs and other diseases) will increase – and health systems will need to adapt accordingly, including: (i) develop appropriate facilities for data processing, storage and analysis; (ii) train personnel accordingly; (iii) develop regulatory frameworks including standards around ethics and informed consent ('genethics'); (iv) data sharing, community engagement, protection of privacy (where and how information is stored and used); (v) develop protocols (e.g. which tests should be done in which circumstances); (vi) establish adequate quality control (e.g. analytics, how results are communicated to individuals); (viii) train health professionals and educate the public ('genome literacy'); (viii) provide guidance to appraise and set directions on how to use genomics in healthcare (e.g. board made of experts from different areas, such as ethicists, civil society, etc.); and (ix) ensure universal access.

Implications for individuals

While genomics and its application for clinical medicine and public health are still in their infancy, a number of tests are already used for NCDs in some countries, as part of 'personalized' or 'precision' medicine and this area will continue to develop. The implications of genetic testing are very significant and need to be considered carefully with appropriate counselling both before and after testing.[20]

Nowadays, people can easily get information on their genome (based on hundreds of thousands of genetic markers on many traits, including risks of NCDs and other diseases) through a variety of direct-to-consumer genetic testing kits (using a swab of saliva) marketed and sold directory to consumers without the involvement of a healthcare provider (including on the internet, e.g. 23andMe), often at a low cost (e.g. <US$ 100). While some interpretation is provided by these providers about the significance of an individual's results, such testing raises a number of complex issues about what people, their families and wider society will feel and do once the results are known.

However, despite advances in genetic technologies and exponential drops in costs, inequalities in healthcare systems, deficits in the genetic research workforce and a lack of access to research funding have prevented knowledge produced by genomic research from truly informing and improving the global public good, particularly in Africa and other low resource settings.

Nevertheless, research into African genomic variations is a scientific imperative for all populations because African genomes, more than any other population, harbour millions of uncaptured variants accumulated over 300,000 years of modern humans' evolutionary history.[21] Moreover, investigating all world populations will contribute to making the outcomes of genetic medicine truly equitable.

Notes

1 Claussnitzer M et al. A brief history of human disease genetics. *Nature* 2020;577;179–89.
2 Genomics beyond health – report overview. UK Government Office for Science, Foresight, UK, 2022.
3 Bilkey GA et al. Optimizing precision medicine for public health. *Front Public Health* 2019;7:42.
4 Zheng Y et al. Association of cardiovascular health through young adulthood with genome-wide DNA methylation patterns in midlife: the CARDIA study. *Circulation* 2022;146:94–109.
5 Sun BB et al. Genetic associations of protein-coding variants in human disease. *Nature* 2022;603:95–102.
6 Khera AV et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* 2019;177:587–96.
7 Khera AV et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219–24.
8 Merino J et al. Polygenic scores, diet quality, and type-2 diabetes risk: an observational study among 35,759 adults from 3 US cohorts. *PLoS Med* 2022;19:e1003972.
9 Brigden T et al. *Implementing polygenic scores for cardiovascular disease into NHS health checks*. University of Cambridge, UK: PHG Foundation, 2021.
10 Richardson TG et al. Harnessing whole genome polygenic risk scores to stratify individuals based on cardiometabolic risk factors and biomarkers at age 10 in the lifecourse – brief report. *Arterioscler Thromb Vasc Biol* 2022;42:362–65.
11 Hollands GJ et al. The impact of communicating genetic risks of disease on risk-reducing health behaviour: systematic review with meta-analysis. *BMJ* 2016;352:i1102.
12 Rutten-Jacobs LC et al. Genetic risk, incident stroke, and the benefits of adhering to a healthy lifestyle: cohort study of 306'473 UK Biobank participants. *BMJ* 2018;363:k4168.
13 Widén E et al. How communicating polygenic and clinical risk for atherosclerotic cardiovascular disease impacts health behavior: an observational follow-up study. *Circ Genom Precis Med* 2022;15:e003459.
14 Ludwig KK et al. Risk reduction and survival benefit of prophylactic surgery in BRCA mutation carriers, a systematic review. *Am J Surg* 2016;212:660–69.
15 Zhou Y, Lauschke VL. Population pharmacogenomics: an update on ethnogeographic differences and opportunities for precision public health. *Hum Genet* 2022;141:1113–36.
16 Esrick EB et al. Post-transcriptional genetic silencing of BCL11A to treat sickle cell disease. *NEJM* 2021;384:205–15.
17 Patrinos GP et al. Roadmap for establishing large-scale genomic medicine initiatives in low- and middle-income countries. *Am J Hum Genet* 2020;107:589–95.
18 Wonkam A. Sequence three million genomes across Africa. *Nature* 2021;590:209–11.
19 Bycroft C. The UK Biobank resource with deep phenotyping and genomic data. *Nature* 2018;562:203–09.
20 Institute of Medicine. *Integrating large-scale genomic information into clinical practice*. Washington, DC: National Academies Press, 2012.
21 Wonkam A et al. Five priorities of African genomics research: the next frontier. *Annu Rev Genomics Hum Genet* 2022;23:499–521.

Taylor & Francis
Taylor & Francis Group
http://taylorandfrancis.com