

# Genome sequencing as a diagnostic test

Gregory Costain MD PhD, Ronald D. Cohn MD, Stephen W. Scherer PhD DSc, Christian R. Marshall PhD

■ Cite as: *CMAJ* 2021 October 25;193:E1626-9. doi: 10.1503/cmaj.210549

**G**enetic testing of patient constitutional DNA (i.e., their genome) is increasingly performed in medical practice.<sup>1-3</sup> Sequencing an entire human genome (about 3.2 billion nucleotides) is now possible to complete in days to weeks, and at a similar cost to some advanced imaging tests or to a brief admission to hospital.<sup>3,4</sup> Genome sequencing is being integrated into health care systems internationally, most notably in the United Kingdom.<sup>5</sup> Starting in 2021, genome sequencing is being performed as a clinical genetic test in Ontario, Canada.

## What is genome sequencing?

Genome sequencing (or whole genome sequencing) is a comprehensive test capable of detecting nearly all DNA variation in a genome. Sequencing can diagnose most of the > 6000 conditions listed in the Online Mendelian Inheritance in Man database ([www.omim.org](http://www.omim.org)) for which the genetic basis is currently understood. These include cystic fibrosis, Duchenne muscular dystrophy, familial hypercholesterolemia, hemophilia A, Lynch syndrome, Marfan syndrome and multiple endocrine neoplasia. Most diagnoses are individually rare (i.e., < 1 in 2000 live births) or ultrarare (i.e., < 100 people reported in the medical literature), and thus less likely to be suspected a priori (e.g., Kabuki syndrome; Figure 1). Patients may present with unusual constellations of features, or with common diseases like autism spectrum disorder, cardiomyopathy, congenital heart disease, epilepsy, cancer, schizophrenia or dementia, although this list is not comprehensive.

Genome sequencing is broader in scope than other commonly used genetic tests (Box 1), and data can be analyzed in both hypothesis-driven and hypothesis-generating ways. For these reasons, genome sequencing will likely eventually supplant exome sequencing, large next-generation sequencing gene panel tests and chromosomal microarray analysis.

## How is genome sequencing delivered?

Genome sequencing is a 3-stage process (Figure 1). First, a medical geneticist or other health care professional collects the necessary information on phenotype and family history. Second, the genome data are generated and reviewed by a clinical laboratory geneticist. Third, a physician correlates the genetic findings with the clinical phenotype. Primary findings are genetic variants that may explain all or a component of the clinical presentation. Identifying

## Key points

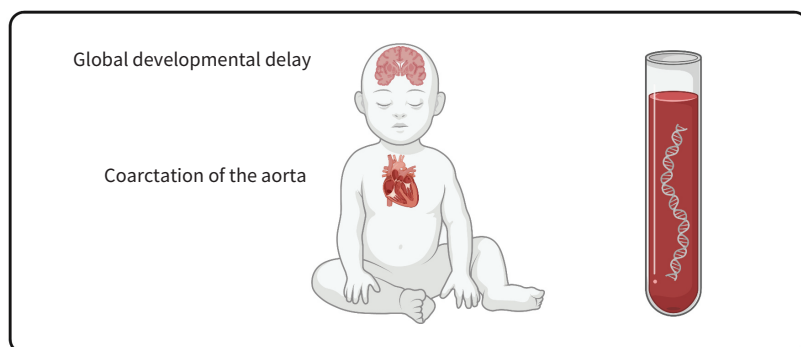
- Genome sequencing is a comprehensive genetic test that is being integrated into health care systems internationally.
- Test indications include suspected genetic disorders in children and adults for whom a targeted genetic testing approach is likely to be low yield or has already failed.
- Analytic validity, diagnostic yield and clinical utility are similar or superior to other clinical genetic tests, such as exome sequencing, chromosomal microarray analysis and next-generation sequencing gene panel tests.
- Appropriate adoption of genome sequencing as a molecular diagnostic test in Canada would be facilitated by a cohesive national strategy for genomic medicine.

such results is the main purpose of genome sequencing as a clinical diagnostic test. In North America, some laboratories will also actively look for secondary findings, namely disease-causing variants in specific genes associated with medically actionable conditions that are unrelated to the reason for testing.<sup>7</sup> Communication of results is embedded within the framework of genetic counselling (Box 2) and is often facilitated by a medical geneticist and genetic counsellor.

DNA sequencing is becoming less expensive and more accurate because of technological advances. However, the cost of offering genome sequencing as a clinical test is still in the range of several thousand Canadian dollars per family-based investigation, not including up-front investments in personnel, equipment and infrastructure.

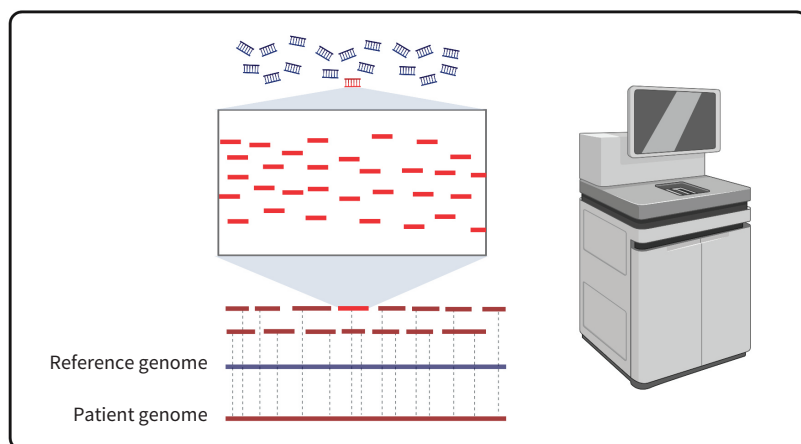
## Who is eligible for genome sequencing?

Genome sequencing is a consideration for children and adults with suspected genetic disorders for whom a targeted genetic testing approach is unlikely to succeed or has already failed. In Ontario, genome sequencing is undergoing a 2-year pilot assessment as part of the Genome-wide Sequencing Ontario project ([www.gsontario.ca](http://www.gsontario.ca)), with publication of interim results pending. Testing can be ordered only by a medical geneticist or other physician with comparable expertise. The approval criteria currently favour individuals with multiple unexplained congenital anomalies or a moderate-to-severe developmental disability, and affect scenarios where a unifying diagnosis is expected to impact management.



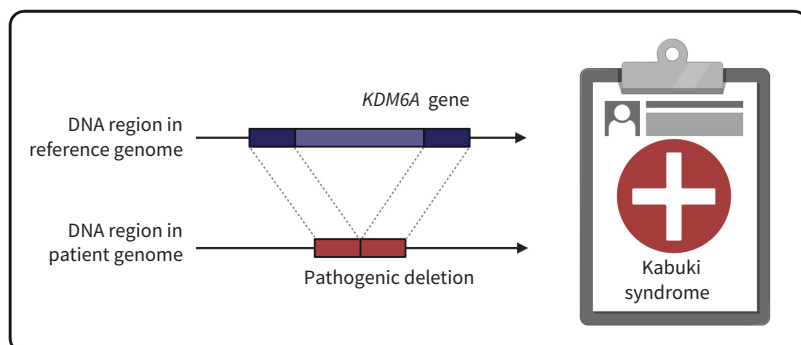
### Presequencing

- 1 Gather necessary medical and family history
- 2 Provide pretest counselling
- 3 Collect sample(s) for DNA extraction



### Sequencing

- 4 Break DNA into small fragments
- 5 Generate hundreds of millions of sequence reads
- 6 Align sequence reads to a reference genome



### Postsequencing

- 7 Identify variant(s) potentially associated with disease
- 8 Correlate with clinical phenotype
- 9 Return result and provide genetic counselling

**Figure 1:** Genome sequencing process. Middle panel was adapted from “Whole Genome Sequencing” by BioRender.com (2021); retrieved from <https://app.biorender.com/biorender-templates>. Note: Variants in the *KDM6A* gene are linked to Kabuki syndrome. After appropriate pretest counselling, including review of possible results, a DNA sample is obtained from the patient, usually via a blood draw. Samples from biological parents are often sequenced at the same time to give additional context regarding inheritance. DNA is fragmented into small pieces and run on a high-throughput sequencing machine to generate millions of “nucleotide reads.” Reads are aligned to a reference genome, like puzzle pieces being assembled by looking at the image pictured on the box. The average number of reads covering a given genomic position (the “read depth”) is typically 30–40. Variants (differences) from that reference are annotated using sophisticated bioinformatics tools and large-scale databases of genomic variation. These variants are filtered and reviewed by a genome analyst to identify 1 or more that might be causal for the observed clinical presentation.

## What are the harms?

The procedure of sequencing is safe; however, possible negative consequences are tied to how results are interpreted and disclosed. First, genome sequencing can be misconstrued as a diagnostic panacea. Accurate clinical information and family history remain important for interpreting results. A positive result may not explain all of the

patient’s features, and a negative result does not rule out a genetic contribution or invalidate an unambiguous clinical diagnosis.<sup>11</sup> Second, the classification of a genetic variant can change over time as new information becomes available.<sup>12</sup> Certain ethnic groups remain under-represented in the large-scale reference databases of genomic variation that guide interpretation, and therefore misdiagnosis is a possibility for these groups.<sup>12</sup> Third, genetic test results can

### Box 1: Overview of selected clinical genetic test modalities<sup>6</sup>

For patients with phenotypes with known genetic heterogeneity, the following tests are commonly employed in clinical practice:

- Chromosomal microarray analysis: a genome-wide test that typically detects only copy number variations (i.e., chromosome imbalances).
- Next-generation sequencing gene panel test: a targeted test focusing on a predefined list of genes that typically detects only exonic sequence-level variants and exon-level deletions or duplications in those genes.
- Exome sequencing: a genome-wide test that typically detects only exonic sequence-level variants, a subset of exon-level deletions or duplications, and a subset of copy number variations.
- Genome sequencing: this approach offers myriad advantages as a single comprehensive test. Current short-read genome sequencing can reliably detect sequence, structural and copy number variations, both within and outside of exons, as well as clinically relevant short tandem repeats, pseudogenes and mitochondrial DNA variation.

### Box 2: Overview of genetic counselling

Genetic counsellors are health professionals with specialized training and experience in providing genetic counselling. In Canada, they may be members of the Canadian Association of Genetic Counsellors ([www.cagc-accg.ca/](http://www.cagc-accg.ca/)). The National Society of Genetic Counselors defines genetic counselling as “the process of helping people understand and adapt to the medical, psychological and familial implications of genetic contributions to disease. This process integrates the following:

- Interpretation of family and medical histories to assess the chance of disease occurrence or recurrence.
- Education about inheritance, testing, management, prevention, resources and research.
- Counselling to promote informed choices and adaptation to the risk or condition.”<sup>8</sup>

Professional societies like the Canadian College of Medical Geneticists<sup>9</sup> and the American College of Medical Genetics and Genomics<sup>10</sup> recognize the importance of pre- and post-test genetic counselling in the context of clinical genome-wide (i.e., exome or genome) sequencing.

reveal unexpected information about the individual, family members or their relationships to one another. These considerations underscore the importance of detailed pre- and post-test counselling and the need for trained genetics professionals (Box 2). An additional safeguard against harms is Canada’s *Genetic Non-Discrimination Act*, which became law in 2017 and is intended to protect individuals from genetic discrimination on the basis of their test results.<sup>13</sup>

## What is the evidence so far?

Most data are from prospective observational trials of testing in clinically heterogeneous populations with suspected rare genetic diseases.<sup>1,2,4,5,14,15</sup> Primary outcome measures are usually diagnostic yield or time to diagnosis. Clinical utility and cost-effectiveness are secondary outcomes of interest. Genome sequencing has a higher diagnostic yield than exome sequencing and chromosomal microarray analysis,<sup>1,11,14,15</sup> at an increasingly competitive cost.<sup>2</sup> Different phenotype categories are associated with different diagnostic yields.<sup>2,3,5</sup> For example, yield may be more than 50% in individuals with severe-to-profound intellectual disability, but less than 10% in most cohorts with single, major congenital anomalies. One comprehensive review of the literature identified 36 studies addressing key laboratory and clinical issues that collectively support a role for genome sequencing in place of other current standard genetic tests (Box 1).<sup>2</sup>

The use of genome sequencing with rapid turnaround time is most well established in neonatal and pediatric intensive care settings, including 2 randomized controlled trials showing increased timely diagnostic yield and clinical utility relative to routine testing.<sup>1,4</sup>

## What can be expected in the future?

Genome sequencing is anticipated to become a first-tier investigation in children and adults with suspected genetic conditions that have high genetic heterogeneity (i.e., a broad genetic differential diagnosis with many candidate genes or loci).<sup>2,3,5,11,14,15</sup> This will curtail

the longstanding practice of ordering multiple genetic tests sequentially and shorten many diagnostic odysseys. Results returned to patients may also include pharmacogenetic profiles, reproductive carrier status information and genetic risk profiles for later-onset conditions.<sup>16</sup> The role for genome sequencing as a preventive health tool in ostensibly healthy individuals remains unclear.<sup>16</sup>

Further evidence is needed regarding the clinical utility, cost-effectiveness and possible unintended downstream consequences of genome sequencing within our health care system. The added yield of genome sequencing over exome sequencing remains modest in some cohorts; however, this gap will widen with advances in data analysis and larger data sets to compare against.<sup>2,11,15</sup> The anticipation of additional clinically relevant information arising from as-yet-unexplored areas of the genome is also driving investments in genome sequencing technology.<sup>2,5,11</sup>

Appropriate adoption of genome sequencing as a diagnostic test in Canada would be facilitated by a coordinated national strategy for genomic medicine that couples basic and clinical research, as exists in the United Kingdom and Australia. Translational genomics projects, such as Care4Rare Canada ([www.care4rare.ca](http://www.care4rare.ca)), the CAUSES Clinic at BC Children’s Hospital, the Integrated Centre for Pediatric Clinical Genomics in Montréal, the Silent Genomes Project ([www.bcchr.ca/silent-genomes-project](http://www.bcchr.ca/silent-genomes-project)), Genome Canada’s All-for-One initiative ([www.genomecanada.ca](http://www.genomecanada.ca)) and the efforts from individual tertiary hospitals (e.g., The Hospital for Sick Children) continue to evolve. These disease-focused initiatives would benefit from a parallel genome sequencing effort of the national population to study the genetic determinants of health. Delivering diagnostic genome sequencing in Canada will also require that more staff be trained, including medical geneticists, clinical laboratory geneticists and genetic counsellors, to enhance the genetic literacy across the diverse population of Canada. Ensuring equitable access to care informed by the DNA code, irrespective of postal code, is a challenge in countries that span a large geography, such as Canada and the United States, but we suggest that this be a priority for policy-makers.

## References

1. Clark MM, Stark Z, Farnaes L, et al. Meta-analysis of the diagnostic and clinical utility of genome and exome sequencing and chromosomal microarray in children with suspected genetic diseases. *NPJ Genom Med* 2018;3:16.
2. Bick D, Jones M, Taylor SL, et al. Case for genome sequencing in infants and children with rare, undiagnosed or genetic diseases. *J Med Genet* 2019;56:783-91.
3. Manickam K, McClain MR, Demmer LA, et al.; ACMG Board of Directors. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 2021 July 1 [Epub ahead of print]. doi: 10.1038/s41436-021-01242-6.
4. Dimmock DP, Clark MM, Gaughran M, et al.; RIGM Investigators. An RCT of rapid genomic sequencing among seriously ill infants results in high clinical utility, changes in management, and low perceived harm. *Am J Hum Genet* 2020;107:942-52.
5. Turro E, Astle WJ, Megy K, et al. Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* 2020;583:96-102.
6. Wallace SE, Bean LJH. Educational materials: genetic testing — current approaches. In: Adam MP, Ardinger HH, Pagon RA, et al., editors. *GeneReviews*. Seattle (WA): University of Washington; 2017, revised 2020 June 18.
7. Miller DT, Lee K, Chung WK, et al. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG) [published erratum in *Genet Med* 2021;23:1582-4]. *Genet Med* 2021;23:1381-90.
8. National Society of Genetic Counselors' Definition Task Force, Resta R, Biesecker BB, et al. A new definition of genetic counseling: national society of genetic counselors' task force report. *J Genet Couns* 2006;15:77-83.
9. Boycott K, Hartley T, Adam S, et al. The clinical application of genome-wide sequencing for monogenic diseases in Canada: position statement of the Canadian College of Medical Geneticists. *J Med Genet* 2015;52:431-7.
10. Manickam K, McClain MR, Demmer LA, et al. Exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). *Genet Med* 2021 July 1 [Epub ahead of print] doi: 10.1038/s41436-021-01242-6.
11. Marshall CR, Chowdhury S, Taft RJ, et al. Best practices for the analytical validation of clinical whole-genome sequencing intended for the diagnosis of germline disease. *NPJ Genom Med* 2020;5:47.
12. Manrai AK, Funke BH, Rehm HL, et al. Genetic misdiagnoses and the potential for health disparities. *N Engl J Med* 2016;375:655-65.
13. *Genetic Non-discrimination Act (GNA)* (formerly Bill S-201) [fact sheet]. Oakville (ON): Canadian Association of Genetic Counsellors; revised 2018 June 28. Available: <https://www.cagc-accg.ca/doc/revised%20GNA%20fact%20sheet%20-%20Jun%2028%202018.pdf> (accessed 2021 July 20).
14. Lionel AC, Costain G, Monfared N, et al. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet Med* 2018;20:435-43.
15. Costain G, Walker S, Marano M, et al. Genome sequencing as a diagnostic test in children with unexplained medical complexity. *JAMA Netw Open* 2020;3:e2018109.
16. Reuter MS, Walker S, Thiruvahindrapuram B, et al. The Personal Genome Project Canada: findings from whole genome sequences of the inaugural 56 participants. *CMAJ* 2018;190:E126-36.

**Competing interests:** Stephen Scherer reports royalties from Athena Diagnostics, Lineagen and Population Bio, and consulting fees from Population Bio and Deep Genomics. No other competing interests were declared.

This article has been peer reviewed.

**Affiliations:** Division of Clinical and Metabolic Genetics (Costain, Cohn), The Hospital for Sick Children; Department of Paediatrics (Costain, Cohn), University of Toronto; Genetics and Genome Biology (Costain, Cohn, Scherer), SickKids Research Institute; Department of Molecular Genetics (Costain, Cohn, Scherer), University of Toronto; The Centre for Applied Genomics (Scherer, Marshall), The Hospital for Sick Children; Department of Laboratory Medicine and Pathobiology (Marshall), University of Toronto; Genome Diagnostics (Marshall), Department of Paediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ont.

**Contributors:** Gregory Costain drafted the manuscript and Ronald Cohn, Stephen Scherer and Christian Marshall revised it critically for important intellectual content. All authors gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

**Content licence:** This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

**Funding:** This group's translational genomics research studies have been funded in part by Genome Canada, the Canadian Institutes of Health Research, SickKids Foundation, SickKids Centre for Genetic Medicine and the University of Toronto McLaughlin Centre. The Genome-wide Sequencing Ontario pilot project is being conducted collaboratively at the Genetics Diagnostic Laboratory at Children's Hospital of Eastern Ontario and the Genome Diagnostics Laboratory at The Hospital for Sick Children.

**Correspondence to:** Gregory Costain, [gregory.costain@sickkids.ca](mailto:gregory.costain@sickkids.ca)

*CMAJ* invites contributions to Innovations, which highlights recent diagnostic and therapeutic advances. Novel uses of older treatments will also be considered. For publication, the benefits of the innovation, its availability and its limitations must be highlighted clearly, but briefly. Visual elements (images) are essential. Submit brief evidence-based articles (maximum 1000 words and five references) to <http://mc.manuscriptcentral.com/cmaj> or email [kirsten.patrick@cmaj.ca](mailto:kirsten.patrick@cmaj.ca) to discuss ideas.