

Clinical exome sequencing – Norwegian findings

BACKGROUND New DNA-sequencing technology is revolutionising medical diagnostics. Through the use of exome sequencing, it is now possible to sequence all human genes in parallel. This technology has been widely used in research over the last few years and is now also being applied to diagnostics. The aim of this study was to systematically examine initial experiences with diagnostic exome sequencing in Norway.

MATERIAL AND METHOD This is a retrospective observational study of the results of all exome sequencing performed by the Section of Medical Genetics at Telemark Hospital between December 2012 and October 2014, and includes 125 persons in 46 families. The majority of these families were being investigated for a syndrome ($n = 35$, 76 %) or neurological disease ($n = 9$, 20 %).

RESULTS Exome sequencing detected pathogenic sequence variants in 15 of 46 probands, and variants of unknown significance in 12 probands. Of the 100 patients who stated their wishes regarding feedback of any incidental findings, six indicated that they did not wish to receive such information. There were no incidental findings in this study, but neither were such sequence variants actively looked for.

INTERPRETATION Exome sequencing can enable more patients with syndromes or neurological diseases to receive a causal diagnosis, and to receive this diagnosis at an earlier stage. However, the patients in this study were quite highly selected, and the results must therefore be interpreted with caution.

New technology is causing ground-breaking changes in genetic diagnostics. The new DNA-sequencing technology, referred to variously as next-generation sequencing, high-throughput sequencing and large-scale DNA sequencing, is used to determine the order of the building blocks (nucleotides) in DNA. The technology can also be used to sequence the entire human genome, whole-genome sequencing, but it has not as yet become widespread, for practical reasons.

Most of the known sequence variants responsible for monogenic diseases or conditions are to be found in the coding regions of genes, the exons (1). Consequently, only these regions, spanning 1–2 % of the genome, are usually sequenced. This is accordingly referred to as exome sequencing. Since whole-genome and exome sequencing have many similarities, the term «genome sequencing» is sometimes used as a collective term for both.

Exome sequencing is much more efficient than conventional DNA sequencing (Sanger sequencing). The fact that all of the roughly 20,000 human genes can now be sequenced in parallel is changing the way in which we test for genetic diseases, in what has been described as a diagnostic revolution (2).

Until now, clinical examinations have served as the basis for the selection and subsequent Sanger sequencing of genes, one by one. If nothing was found, it was necessary to go back and review the clinical data again

or perform further tests to identify other relevant genes to sequence. This was both time-consuming and resource intensive. Because exome sequencing enables all genes to be studied in parallel, testing can now be performed without any prior selection of genes.

When a recessive condition or new (*de novo*) mutations are suspected, it is beneficial to study both the affected proband and their healthy parents. This is referred to as trio sequencing and it is often used in the investigation of syndromes, intellectual disability and neurological disease (3, 4).

When all genes are sequenced in parallel, there is a possibility of making incidental findings that are unrelated to the study indication. These may have predictive value – for example, a sequence variant associated with high cancer risk might be detected in a patient who was initially tested for a neurological disease. This is regulated in Norway by the Biotechnology Act, and there is a requirement for written consent and genetic counselling of patients in predictive studies (5). Institutions must also be approved by the Norwegian Directorate of Health to perform such testing.

In 2014, the report «Personalised medicine in the health services» was prepared by the regional health authorities at the request of the Norwegian Ministry of Health and Care Services (6). The report recommends expanding the use of exome and whole-genome sequencing in the diagnosis of pa-

Øystein L. Holla
oholla@sthf.no
Øyvind L. Busk
Kristian Tveten
Hilde T. Hilmarsen
Linda Strand
Helle Høyer
Anette Bakken
Camilla F. Skjelbred
Geir J. Braathen
Section of Medical Genetics
Department of Laboratory Medicine
Telemark Hospital

Appendix at www.tidsskriftet.no/hollaengappendiks

MAIN POINTS

New DNA-sequencing technology makes it possible to study all human genes simultaneously in a single assay

In this observational study, the sequencing technology was applied to patients who had been investigated for a number of years for a suspected hereditary syndrome or neurological disease

One third of the patients received a causal diagnosis

Table 1 Summary of study participants

Study participants	125
Probands (pedigrees)	46
Pedigrees with 1 study participant	6
Pedigrees with 2 study participants	5
Pedigrees with 3 study participants	31
Pedigrees with 4 study participants	4
Probands with likely and clearly pathogenic variants	15
<i>De novo</i>	7
Inherited	8
Probands with variants of unknown significance	12
Only <i>de novo</i>	5
<i>De novo</i> and inherited	2
Only inherited	5
Consent form with stated wishes regarding feedback of incidental findings	100
Only on study indication	6
Also incidental findings regarding treatable diseases	38
All incidental findings	56

tients with rare hereditary single gene disorders, and evaluating the usefulness of such data for the patients concerned. There are also draft Norwegian guidelines on the use of genome sequencing, intended for clinicians, researchers and regional ethics committees (7). Further guidelines are now being prepared by interdisciplinary working groups initiated by the Norwegian Medical Association, through the Norwegian Society of Medical Genetics and the Norwegian Society of Human Genetics.

Exome sequencing has been used in research on Norwegian datasets for purposes including the detection of a new disease gene for chronic diarrhoea (8) and the mapping of genetic causes of maturity-onset diabetes of the young (9). However, there are no Norwegian reports on the use of this technology in diagnostics. The purpose of this study was therefore to systematically review initial experiences of the use of exome sequencing in genetic diagnostics in Norway and to determine whether exome sequencing has worked out as intended.

Material and method

This retrospective observational study includes all diagnostic exome sequencing

analyses in which gene lists for specific diseases were not used to limit the genes studied. The results were obtained from the Section of Medical Genetics at Telemark Hospital from the start of the analysis in December 2012 to October 2014. A specialist in medical genetics considered it likely that the patients had a rare monogenic condition, but the underlying cause had not been determined by other tests.

Trio sequencing was performed, which also entails the study of unaffected individuals. The dataset does not include prenatal diagnostics or studies of deceased persons. The study was approved by the Norwegian Social Science Data Services/Data Protection Officer for quality assurance purposes, and hence unnecessary to present to the regional ethics committee for medical research.

All participants received genetic counselling before and after exome sequencing and consented to the study. Telemark Hospital uses an informed consent form with three options for feedback regarding any incidental findings (see www.tidsskriftet.no/hollaengappendiks).

DNA was extracted from blood, and samples were worked up using the TruSeq

Exome Enrichment Kit or Nextera Rapid Capture Exome Kit (Illumina, San Diego, CA) before sequencing on a HiScanSQ (Illumina). For each sample, an average of 57,000 sequence variants were detected, of which one or two were assumed to be responsible for the patient's disease/condition. Sequence variants with high frequencies were filtered out since the diseases being tested for are rare.

An alternative allele frequency of 0.01 was initially used for recessive inheritance and a frequency of 0.001 for dominant inheritance. International databases of sequence variants were consulted, but given that the Norwegian population has distinctive normal variants, an in-house database of Norwegian variation was used in addition. Sequence variants that were synonymous (no change in the amino acids), intronic (outside the exons and splice sites) or in untranslated regions (UTR) were discarded. There were also requirements with respect to technical quality. Genes and variants that remained after filtering were then manually reviewed in the light of available clinical data to determine causality. The key databases used for this were Online Mendelian Inheritance in Man (OMIM), the Human Gene Mutation Database (HGMD) and Medline/PubMed.

Sequence variants were categorised into five classes: class 5: Clearly pathogenic, class 4: Likely to be pathogenic, class 3: Unknown significance, class 2: Unlikely to be pathogenic, class 1: Clearly not pathogenic. Any remaining sequence variants in genes not associated with the patient's phenotype were not considered further. Key aspects to consider when classifying sequence variants can be found in guidelines from the Association for Clinical Genetic Science (10) and the American College of Medical Genetics and Genomics (11).

Only variants in class 4 and class 5 were reported to the patient, and all of these were verified with Sanger sequencing of all available family members. Laboratory methods and the interpretation of sequence variants have been described in more detail previously (12, 13). The exome sequencing methodology along with accompanying bioinformatics and interpretation of sequence variants has been validated and accredited in accordance with ISO15189.

Results

The patient sample comprises 46 distinct pedigrees (probands) with testing of 125 persons (Table 1). Most pedigrees ($n = 31$) were trios with healthy biological parents and their affected offspring. All were resident in Southern or Eastern Norway, with the exception of three persons. Half the pro-

Table 2 Sequence variants judged to be pathogenic

Category	Mutated gene	<i>De novo</i> sequence variant	Genetic diagnosis
Neurological disease	<i>ACTG2</i>	–	Visceral myopathy
	<i>DARS2</i>	–	Leukoencephalopathy
	<i>RYR1</i>	–	Myopathy/«central core disease», malignant hyperthermia
	<i>SBF1</i>	–	Charcot-Marie-Tooth disease
	<i>TTPA</i>	–	Ataxia with vitamin E deficiency
	<i>ACTB</i>	Yes	Baraitser-Winter syndrome
	<i>ARID1B</i>	Yes	Coffin-Siris syndrome
Syndrome	<i>EPG5</i>	–	Vici syndrome
	<i>GRIN2A</i>	Yes	Epileptic encephalopathy
	<i>MED13L</i>	Yes	Intellectual disability
	<i>NSD1</i>	Yes	Sotos syndrome
	<i>NSD1</i>	Yes	Sotos syndrome
	<i>PTEN</i>	Yes	Cowden syndrome/ <i>PTEN</i> hamartoma tumour syndrome
	<i>RECQL4</i>	–	Rothmund-Thomson syndrome
	<i>SPG11</i>	–	Spastic paraplegia, complex

bands (n = 23) were resident in Telemark. The average age of the probands when genetic diagnosis was requested was 16 years, and 16 subjects were less than five years old.

Median coverage of sequenced exons \pm 2 base pairs was 79x. In all, 88 % of this area was sequenced > 20x and 83 % was sequenced > 30x. In the 31 trios studied, 0–8 (mean 2.1) rare *de novo* coding variants were detected in the probands. On average, sequence variants in 60 genes were consistent with dominant inheritance, and variants in 31 genes with recessive inheritance.

The patients were being evaluated for a syndrome (n = 35), neurological disease (n = 9), haematological/immunological disorder (n = 1) or endocrine disorder (n = 1). Typical symptoms and signs for patients with a suspected syndrome were intellectual disability, dysmorphic features and delayed psychomotor development. Patients with suspected neurological disease also had a broad spectrum of symptoms and signs, including ataxia, speech impediments, intestinal pseudo-obstruction and paresis.

The probands had previously been thoroughly assessed with X-ray imaging, biochemical analyses and clinical examinations. They had also undergone molecular genetic testing, such as comparative genomic hybridisation (aCHG, n = 45), G-band karyotyping (n = 41), multiplex ligation-

dependent probe amplification (MLPA, n = 32) and Sanger sequencing of one or more single genes (n = 18), without any clearly pathogenic variants being found. These are minimum numbers, since some patients may have been studied at other hospitals without our knowledge.

Exome sequencing revealed putative causal variants in 15 of 46 pedigrees (33 %) (Table 2). Inherited autosomal dominant diseases were found in two pedigrees, malignant hyperthermia and visceral myopathy (12). Both can have serious consequences which it is possible to prevent. Six recessive conditions were found, in which the affected individual had inherited a mutated allele from each parent. Seven of the variants detected were *de novo* heterozygous mutations. All of these were found in patients with a suspected syndrome. At the time of the genetic diagnosis, half of the patients had been under evaluation for more than five years; the average was ten years.

Variants of unknown significance (VUS) were detected in 12 probands without a definitive genetic diagnosis. Seven of these individuals had *de novo* heterozygous variants, and two of the seven had inherited variants too. Five probands had inherited VUS only.

There were no incidental findings in any of the 125 study participants. Of those examined, 100 had stated in writing whether

they wished to be informed of any incidental findings. More than half (n = 56) asked to receive information about all incidental findings that could have implications for their health. A total of 38 persons wished to be informed only about incidental findings regarding conditions that can be treated or prevented, while six wished to receive only the results the study was intended to yield.

Discussion

This is the first report on the use of exome sequencing in medical diagnostics in Norway. The majority of the patients in the study were assumed to have a rare syndrome, and all had previously undergone extensive workup without receiving a precise diagnosis. The patient sample consists of individuals who were referred for genetic testing; it is thus quite highly selected and includes a broad spectrum of phenotypes. The results must therefore be interpreted with caution. Exome sequencing revealed a putative pathogenic sequence variant in 15 of 46 probands, a proportion consistent with international studies (14). Five variants were detected in nine patients with neurological disease. Higher diagnostic yield in patients investigated for neurological symptoms has been reported by others (15, 16).

Exome sequencing can help to provide patients with genetic diseases with a specific

diagnosis. This can reveal risk factors that they should avoid, or even make it possible for them to receive treatment (17). The prognosis can be assessed, and the risk of recurrence in any future pregnancies estimated. If the identity of the pathogenic sequence variant for serious disorders is known, this may open up the possibility of prenatal genetic diagnostics.

A specific genetic diagnosis can be used as a basis for further follow up. This also applies to symptoms and sequelae that the patient does not yet have, but which they are at increased risk of developing in the future. Many of these sequelae are not obvious. For example, patients with *ARID1B* mutations should be monitored by the dental services from the age of three years due to difficulties with eating combined with tooth development disorders (18).

Most patients had been thoroughly assessed over a number of years prior to the sequencing and had been evaluated by specialists both at home and abroad. As stated above, half had been investigated for more than five years at the time of genetic diagnosis; the average was a whole ten years. The possibility of using exome sequencing earlier in the diagnostic process should therefore be considered for certain patient groups. As well as receiving a diagnosis more quickly, these individuals could potentially be spared invasive tests, incorrect treatment and anxiety.

To reduce the number of false positive results with exome sequencing, all pathogenic variants (class 4 and class 5) were verified by Sanger sequencing. False negatives can also occur with exome sequencing, as parts of the sequence may be of poor technical quality (coverage), enabling variants to escape detection. If there is a strong suspicion that the patient's condition is caused by a pathogenic variant in a very limited region of the genome, such as a single gene, Sanger sequencing or targeted deep sequencing may be more suitable than exome sequencing. However, calculations show that the probability of detecting pathogenic point mutations or small insertions/deletions with exome sequencing is 93% if the average coverage is over 100x (15). In common with Sanger sequencing, exome sequencing is not well-suited to detecting larger insertions/deletions or expansions.

Knowledge of the functions of genes and their role in disease pathogenesis is increasing steadily. As many as five of the 14 genes shown to contain pathogenic variants in this study were first linked to the disease/condition in question in 2012. This means that if the analysis had been performed before then, these variants would probably not have been interpreted as pathogenic. There is thus

reason to believe that variants of unknown significance today could be viewed as pathogenic in the future.

Interpretation of sequencing data requires knowledge of wet lab techniques, bioinformatics, molecular genetics and clinical genetics – not to mention knowledge of the patient's symptoms and the family's medical history. Most sequence variants can be assumed to be benign because they occur at high frequencies in normal populations or do not change the protein's amino acid sequence. However, some variants require a thorough manual assessment. Interdisciplinary collaboration is essential to determine how genetic data are related to clinical symptoms.

Clinical information is a key factor in achieving good results. If such information is lacking, the results of exome sequencing can nevertheless be used to guide further clinical tests. Even in cases where clinical information subsequently proves to be highly consistent with the final diagnosis, the correct diagnosis is often not made until the results of exome sequencing are available. A new clinical assessment of the patient may allow the diagnosis to be confirmed or ruled out, since it is easier to demonstrate or exclude symptoms when one knows what to look for.

There were no incidental findings in this study. However, such sequence variants were not actively looked for either, since this would require additional resources and is not recommended by current Norwegian guidelines. In studies that do actively search for such variants, secondary findings are made in up to about 5% of patients, depending on how many genes are examined (16). Only six participants in this study (6%) did not wish to be informed of any incidental findings. A further 38% did not wish to be informed of incidental findings associated with untreatable disorders.

Exome sequencing is effective – the cause of disease was identified in one third of these patients with syndromes or neurological diseases who had already undergone other forms of genetic testing without receiving a diagnosis.

Øystein Lunde Holla (born 1978)

PhD, chief biomedical laboratory scientist and head of deep sequencing. The author has completed the ICMJE form and reports no conflicts of interest.

Øyvind Løvold Busk (born 1982)

PhD, bioinformatician and head of bioinformatics. The author has completed the ICMJE form and reports no conflicts of interest.

Kristian Tveten (born 1977)

PhD, chief biomedical laboratory scientist. The author has completed the ICMJE form and reports no conflicts of interest.

Hilde Tveitan Hilmarsen (born 1975)

MSc, head of Sanger sequencing. The author has completed the ICMJE form and reports no conflicts of interest.

Helle Høyser (born 1981)

MSc, PhD candidate. She conducts research on genetics, deep sequencing and hereditary peripheral neuropathy. The author has completed the ICMJE form and reports no conflicts of interest.

Linda Strand (born 1968)

MSc, with responsibility for development and quality control. She is an inspector for Norwegian Accreditation. The author has completed the ICMJE form and reports no conflicts of interest.

Anette Bakken (born 1960)

specialist in medical genetics. The author has completed the ICMJE form and reports no conflicts of interest.

Camilla Furu Skjelbred (born 1967)

PhD and head of section. The author has completed the ICMJE form and reports no conflicts of interest.

Geir Julius Braathen (born 1958)

PhD, specialist in neurology and medical genetics. The author has completed the ICMJE form and reports no conflicts of interest.

References

- Human Gene Mutation Database HGMD. www.biobase-international.com/product/hgmd [4.7.2015].
- Prescott T. En diagnostisk revolusjon. *Tidsskr Nor Legeforen* 2013; 133: 1550.
- de Ligt J, Willemsen MH, van Bon BW et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med* 2012; 367: 1921–9.
- Ku CS, Tan EK, Cooper DN. From the periphery to centre stage: de novo single nucleotide variants play a key role in human genetic disease. *J Med Genet* 2013; 50: 203–11.
- Lov om humanmedisinsk bruk av bioteknologi m.m. (bioteknologiloven). Kap. V. <https://lovdata.no/dokument/NL/lov/2003-12-05-100> [4.7.2015].
- Nasjonal utredning av persontilpasset medisin i helsetjenesten. www.helse-sorost.no/fagfolk/_forskning/_persontilpasset-medisin/_Sider/_persontilpasset-medisin-helsetjenesten.aspx [4.7.2015].
- Bioteknologirådet. Forslag til veileder og retningslinjer for bruk av genomsekvensering og genomdata i klinikk og forskning. www.bioteknologiradet.no/filarkiv/2012/11/Veileder_genomsekvensering_091112_til_net.pdf [4.7.2015].

>>>

8. Fiskerstrand T, Arshad N, Haukanes BI et al. Familial diarrhea syndrome caused by an activating GUCY2C mutation. *N Engl J Med* 2012; 366: 1586–95.
9. Johansson S, Irgens H, Chudasama KK et al. Exome sequencing and genetic testing for MODY. *PLoS ONE* 2012; 7: e38050.
10. Wallis Y, Payne S, McAnulty C et al. Practice Guidelines for the Evaluation of Pathogenicity and the Reporting of Sequence Variants in Clinical Molecular Genetics. www.acgs.uk.com/media/774853/evaluation_and_reporting_of_sequence_variants_bpgs_june_2013_-_finalpdf.pdf [4.7.2015].
11. Richards S, Aziz N, Bale S et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 2015; 17: 405–24.
12. Holla OL, Bock G, Busk OL et al. Familial visceral myopathy diagnosed by exome sequencing of a patient with chronic intestinal pseudo-obstruction. *Endoscopy* 2014; 46: 533–7.
13. Hoyer H, Braathen GJ, Busk OL et al. Genetic diagnosis of Charcot-Marie-Tooth disease in a population by next-generation sequencing. *Biomed Res Int* 2014; 2014: 210401.
14. Yang Y, Muzny DM, Reid JG et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013; 369: 1502–11.
15. Lee H, Deignan JL, Dorrani N et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. *JAMA* 2014; 312: 1880–7.
16. Yang Y, Muzny DM, Xia F et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 2014; 312: 1870–9.
17. van Karnebeek CD, Shevell M, Zschocke J et al. The metabolic evaluation of the child with an intellectual developmental disorder: diagnostic algorithm for identification of treatable causes and new digital resource. *Mol Genet Metab* 2014; 111: 428–38.
18. Lovisenberg Diakonale Sykehus. www.lds.no/modules/module_123/proxy.asp?C=272&I=4914&D=2 [4.7.2015].

Received 20 November 2014, first revision submitted 19 March 2014, accepted 4 July 2015. Editor: Are Brean.