

REVIEW

The promise of whole-exome sequencing in medical genetics

Bahareh Rabbani¹, Mustafa Tekin² and Nejat Mahdieh³

Massively parallel DNA-sequencing systems provide sequence of huge numbers of different DNA strands at once. These technologies are revolutionizing our understanding in medical genetics, accelerating health-improvement projects, and ushering to a fully understood personalized medicine in near future. Whole-exome sequencing (WES) is application of the next-generation technology to determine the variations of all coding regions, or exons, of known genes. WES provides coverage of more than 95% of the exons, which contains 85% of disease-causing mutations in Mendelian disorders and many disease-predisposing SNPs throughout the genome. The role of more than 150 genes has been distinguished by means of WES, and this statistics is quickly growing. In this review, the impacts of WES in medical genetics as well as its consequences leading to improve health care are summarized.

Journal of Human Genetics (2014) 59, 5–15; doi:10.1038/jhg.2013.114; published online 7 November 2013

Keywords: cancer; common disease; medical genomics; Mendelian disorder; whole-exome sequencing

INTRODUCTION

DNA sequencing is one of the main concerns of medical research nowadays. Union of chain termination sequencing by Sanger *et al.*¹ and the polymerase chain reaction (PCR) by Mullis *et al.*² established many marked events such as the completion of the Human Genome Project (HGP), providing a barely sufficient reference to investigate the genetic alterations in the associated phenotypes.^{3–5}

In Sanger sequencing, all four of the standard deoxynucleotides (dNTPs) and the four chain-terminating dideoxynucleotides (ddNTP) with different fluorescent colors attached on each are crucial to perform a sequencing reaction as well as a DNA primer, a single-stranded template, and a DNA polymerase enzyme. Eventually, newly developed technologies replaced the traditional methods for whole-genome sequencing (WGS) and whole-exome sequencing (WES) with a less expensive sequencing cost per genome/exome. These revolutionizing sequencing technologies, so-called next-generation sequencing (NGS), are promising to be used in clinic to improve human health, although their expensive costs, ethical issues related to the produced genetic data and the need for user-friendly software in the analysis of the raw sequence have to be addressed. The current NGS platforms include those produced by 454 Life Sciences, Illumina, Applied BioSystems, Helicos Biosciences, Danaher Motion, Oxford Nanopore Technologies and Pacific Biosciences. NGS has influenced all fields of biological research; here, we provide a review of the recent literature focusing on applications of WES in medical research and how it is imbuing our understanding about genetic mechanisms of the diseases.

VALUE OF WES IN MEDICINE

Human genome comprises $\sim 3 \times 10^9$ bases having coding and noncoding sequences. About 3×10^7 base pairs (1%) (30 Mb) of the genome are the coding sequences. The genome assembly from the Genome Reference Consortium (GRCh37.p10, Feb 2009) in Ensembl (http://asia.ensembl.org/Homo_sapiens/Info/StatsTable?db=core) includes $\sim 20\,000$ protein-coding genes, pseudogenes and noncoding genes. The HUGO Gene Nomenclature Committee approves more than 19 000 protein-coding genes, 8000 pseudogenes, 4000 noncoding RNA genes (http://www.genenames.org/cgi-bin/hgnc_stats) and almost 34 000 gene symbols.⁶ In fact, less than 10% (~ 3 Mb) of the whole-genome sequence is characterized, and especially insufficient clinical knowledge is acquired from genome sequence.⁷ In addition, it is estimated that 85% of the disease-causing mutations are located in coding and functional regions of the genome.^{8,9} For this reason, sequencing of the complete coding regions (exome) has the potential to uncover the causes of large number of rare, mostly monogenic, genetic disorders as well as predisposing variants in common diseases and cancers. Although the identification of the causal variations in coding and noncoding sequences leads to a clinical phenotype, the epigenetic modifications are to an extent other critical determinants of the phenotype.¹⁰

Details of the WES strategy have been recently reviewed.^{9,11} In general, the human genome project encompasses about 3.5 million SNPs. A multistep filtering against public databases such as dbSNP, 1000 Genome Project¹² and HapMap (www.hapmap.org) is used that narrows down the identified variants to a small number. On the basis

¹Department of Medical Genetics, Faculty of Medicine, Qazvin University of Medical Sciences, Qazvin, Iran; ²John P Hussman Institute for Human Genomics, Miller School of Medicine, University of Miami, Miami, FL, USA and ³Department of Medical Genetics, Faculty of Medicine, Ilam University of Medical Sciences, Ilam, Iran
Correspondence: Dr N Mahdieh, Department of Medical Genetics, Faculty of Medicine, Ilam University of Medical Sciences, Ilam 6939177143, Iran.
E-mail: nmahdieh@yahoo.com

Received 11 August 2013; revised 29 September 2013; accepted 11 October 2013; published online 7 November 2013

of the disease being studied and its mode of inheritance, a specific strategy can be applied to reach the possible causing variants; the functional and biological significance of each variant follows a gradient that ranges from normal to pathological effect.^{13,14} Notably, variant calling and gene annotation of WES data of each individual exome contains about 10 000 nonsynonymous variants depending on ethnicity and calling methods. A normal individual has been estimated to have 50–100 mutations in the heterozygous state that can cause a recessive Mendelian disorder when being homozygous.¹²

Identification of the variants causing the disease brings the research into clinical practice. Disease-causing variants with large pathogenic effect (high penetrance) mostly seen in single gene disorders are the first group of classified variants. These variants are mainly rare. Although some variants are in handful of individuals who are associated with rare or uncommon diseases, they are categorized as likely disease-causing variants with less certainty of the variants causing the disease due to incomplete penetrance. NGS approach would verify these variants, which is helpful in management of individuals carrying such variants. Other group is the variants with higher frequency and lower penetrance in cases than controls based on genome-wide association studies. These variants could be detected with DNA chip genotyping and NGS approaches. WES could identify these variant and is used in clinical management of individuals, for example, in familial hypercholesterolemia dietary management in individuals having the causal variants is lifesaving. High penetrance variants detected by WES are important in diagnosis of the patients and healthy carriers.

There are functional variants including insertions, deletions, nonsense variants, splice variants and copy-number variation with no association to disease and are suspected to cause a disease. They may be novel, and WES of different individuals carrying the variants may be a clue for their effect. Finally, a large number of variants has an unknown clinical significance located in introns and intergenic regions.¹³ They may be benign or pathogenic.

WES IN RESEARCH

NGS has improved our understanding of the genetic pathology of the diseases. The first report of selectively sequencing of whole exome was published by Ng *et al.*¹⁵ in 2009. They reported the targeted capture and massively parallel sequencing of the exomes of 12 humans including eight individuals previously characterized by HapMap and Human Genome Structural Variation project and four unrelated individuals affected by a rare dominantly inherited disorder called Freeman-Sheldon syndrome, which is known to be caused by mutations in the *MYH3* gene (FSS; MIM193700). Rare and common variants were identified. After assessment of the quality of the exome data, 13 347 variants were appeared to be novel. Subsequent filtering of these variants against dbSNP (v129) or those found in the HapMap samples defined the *MYH3* gene as a disease-causing gene in FSS patients. Consequently, WES could help to define the causal variants more easily. This kind of study is performed to uncover the genetic basis of genetic disorders, which would lead to treatment and management of patients. A PUBMED search limited to Title/Abstract on 'whole-exome sequencing' revealed that more than 150 articles have been published about gene discovery in various disorders.¹⁴ Current literature of WES is growing exponentially.

WES IS CHARMING ITS NEW WAY IN DIAGNOSTIC AND CLINICAL GENETICS

In the era of genetic medicine, most developed technologies are coping with diagnosis and the development of new therapies. Medical

sequencing projects have challenge of identifying the cause of rare genetic disorders and cancers. Different genes and causal variants have been discovered by WGS and WES. The aim is to provide efficient and effective genetic data for best treatment; albeit, accurate, fast and cost-effective diagnosis of the patients are a major concern. Similarly, the preclinical individuals at risk of having the disease could be identified; however, prognostic evaluation and determining the appropriate therapy is applied for those individuals with defined sequence by the NGS technology.

The first report that pinpointed an exact diagnosis established by WES was published by Choi *et al.*¹⁶ in 2009. WES on a patient referred for Bartter syndrome (MIM241200), a rare inherited disorder characterized by hypokalemia, showed that the patient has a novel homozygous mutation in *SLC26A3* gene; mutations in this gene have been known to cause congenital chloride-losing diarrhea (CLD; MIM214700). Clinical re-evaluation of the patient who was misdiagnosed as Bartter syndrome established the correct diagnosis as CLD. Another study showed that WES helped in the diagnosis of a patient with Leber congenital amaurosis who had mutation in *PEX1* gene associated with peroxisome biogenesis disorders.¹⁷ A patient misdiagnosed as having the Hermansky–Pudlak syndrome type 2 was considered for homozygosity mapping and WES because of other phenotypic effects including oculocutaneous albinism (OCA) and neutropenia. No mutation was found in the related *AP3B1* gene; followed by WES, two disease causative genes were identified, naming *SLC45A2* (related to OCA) and *G6PC3* (related to neutropenia) genes.¹⁸

Indeed, when applying for diagnostic, sequencing of many known disease-associated genes at once via NGS (in case of genetically heterogeneous disorders) is probably enough rather than WES.¹⁹ In a survey by Shen *et al.*,²⁰ 524 candidate nuclear-encoded mitochondrial genes were sequenced for molecular diagnosis. Because of genetic heterogeneity of inherited cardiomyopathies, Meder and colleagues²¹ used targeted exome sequence of about 30 selected genes. Jones and colleagues²² selected 24 genes of congenital disorders of glycosylation by means of targeted NGS. In this case, greater sequencing depth is achieved at a relatively low cost as well as a time-to-diagnosis would be improved using this targeted gene approach.²³

Disorders with phenotypic/genetic heterogeneity or patients with overlapping symptoms are difficult to diagnose. The clinical differential diagnostic tests may be lengthy and costly.²⁴ WES provides insights into genetic diagnosis of the challenging cases. For example, a phenotypically and genetically heterogeneous disorder named neuroacanthocytosis (NA) syndrome includes chorea-acanthocytosis (ChAc), X-linked McLeod syndrome (MLS), Huntington's disease-like 2 (HDL2) and pantothenate kinase-associated neurodegeneration (PKAN). WES of the genes made the diagnosis easier.²⁵

WES has a useful application in disease treatment, screening and prenatal diagnosis. A 15-month-old boy with an immune deficiency was exome sequenced and diagnosed as Crohn disease due to a mutation in the X-linked inhibitor of apoptosis gene. For the management of the patient, hematopoietic progenitor cell transplant with an excellent outcome was performed.²⁶

Neonatal screening of diabetes mellitus (NDM) has been studied by WES; oral treatment with sulfonylurea drugs instead of insulin therapy is used for patient care in patients carrying a mutation in *KCNJ11* or *ABCC8* genes.²⁷

Prenatal diagnosis using fetal DNA in maternal serum represented the applicability of WES in finding aneuploidies with a noninvasive method.^{28–30} In a survey by Bell *et al.*³¹ carrier testing carried out for 448 severe recessive childhood diseases by NGS. The expense of

carrier testing with such a method is lower than the cost of treatment and management of the children born with such disorders.

Preclinical application of WES was recognized by characterization of mutations in genes causing phenotypically similar disorders. Jiao and colleagues³² studied 10 nonfamilial pancreatic neuroendocrine tumors with WES. The frequent gene mutations were used for prognosis.

Our understanding of the genetic and epigenetic profiles of a disease facilitates the identification of mutations for diagnosis, prognosis and new therapeutic approaches for health-care programs.⁷ WES has evolved the biomedical research. The characterization of the disease and genes involved are understood by massively parallel sequencing. WES is less challenging than WGS, although the cost will become less. More or less, all NGS strategies (WES, RNAseq, targeted sequencing, WGS,...) could be used in genetic diagnosis depending on the disease and the available information; for example, targeted NGS works for X-linked mental retardation, but WGS may be appropriate for mental retardation and congenital malformation due to structural variations. The complexity of genome may affect the disease progression as in cancers; therefore, different approaches may be used. As above mentioned, WES has been applied in different areas of research and diagnostics: in brief, including diagnosis (prenatal diagnosis (PND), preimplantation genetic diagnosis (PGD), carrier/mutation detection of the heterogeneous disorders

such as hearing loss with many causal genes), prognosis of preclinical individuals, newborn screening procedures and treatment. Certainly, gene discovery of unknown disorders and a predisposed SNP detection of common disorders clarify the genetic basis and molecular mechanism of disorders. The interaction of molecules in cells (interactome) is revealed, and the genes in biological networks are determined. In particular, WES could improve health care by influencing disease management and drug discovery, and later personalized medicine is achieved for customizing the health of each individual (Figure 1). These are maintained for different disorders ranging from rare Mendelian disorders and heterogeneous disorders to polygenic disorders complex diseases such as common disorders and cancers. Here, we focus on the usage of WES in different category of diseases.

USAGE OF WES IN HUMAN DISORDERS

Monogenic disorders

Currently, over 6000 presumably monogenic disorders have been described, but for nearly two-thirds of these the molecular basis has not been reported (OMIM Statistics). Understanding the pathogenic mechanism of a disease mostly depends on finding the causative gene and variants associated with the phenotype. In case of a newly identified variant in a patient or in a small family, a clearly defined genetic diagnosis is difficult to be plausible only on the basis of

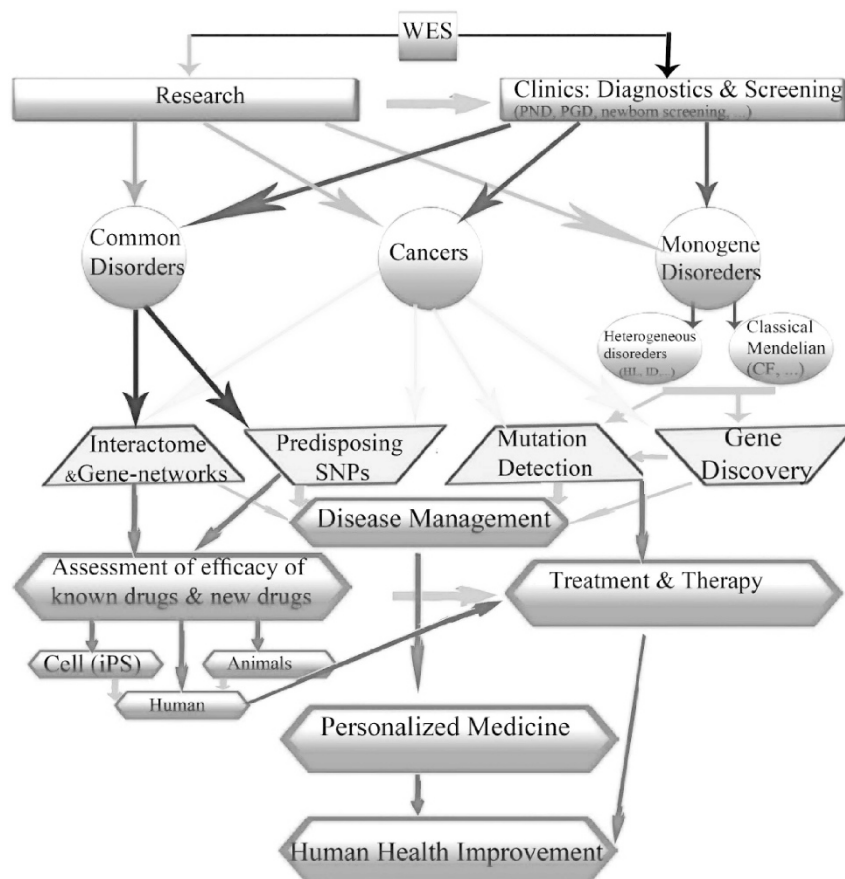


Figure 1 WES and impact of its genetic consequences on human public health. Variant calling of the WES data is applied for various purposes in different disorders: diagnosis (prenatal diagnosis: PND, preimplantation genetic diagnosis: PGD and mutation detection in heterogeneous diseases such as hearing loss using selected genes for faster diagnosis), screening procedures and research. Consequently, WES benefits for treatment and management of patients, gene discovery, SNP detection for drug effects and finding disease mechanism and genes network. All these will hint at health improvement including disease management, personalized medicine. A full color version of this figure is available at the *Journal of Human Genetics* journal online.

variant finding. Whereas its absence in the population controls is verified, the presence of the same and other variants in the same gene in other patients or families with the same disease are usually used to confirm the new pathogenic variant/s. If the disorder is extremely rare, it is hard to find more patients. However, further functional experiments are crucial to validate the pathologic impact of the newly determined variants; if the mutated gene has a defined role in a well-known pathway related to the disease, it is acceptable to perform biochemical confirmatory experiments. Identification of novel genes causing rare monogenic disorders is crucial to comprehend the biological pathways as well as treatment or therapeutic management. Recent publications on WES show that it is a powerful tool in finding the causal genes for Mendelian disorders;¹⁴ here we focus only on application of genetically heterogeneous single-gene disorders.

Heterogeneous monogenic phenotypes

There are many genetically heterogeneous conditions such as hearing loss, intellectual disabilities, autism spectrum disorders and retinitis pigmentosa. WES has been successfully used to identify the causative variants in several heterogeneous conditions. Here, we briefly mention some examples.

Hearing loss. Hearing loss (HL) is the most prevalent sensory defect in humans affecting 1.86 in 1000 newborns worldwide, half of which is due to genetic causes.³³ On the basis of estimation, about 1% of human genes (200 to 250 genes) may be involved in hereditary HL.³⁴ To date, more than 60 genes have been reported to cause nonsyndromic HL.³⁵ This picture would be more complicated if we consider that more than 400 syndromes related to HL have been described in the OMIM. Identification of novel genes using classical gene discovery methods is usually laborious and time consuming. WES is a rapid way to get a full picture of the protein-coding sequence variations. For example, Walsh *et al.*³⁶ described the application of WES in conjunction with homozygosity mapping to define the causative variant for recessive deafness in a consanguineous family. By using the regions of homozygosity as an evidence, they quickly identified the causative mutation, p.R127X, in *GPSM2* that maps to the *DFNB82* locus.^{36,37} Diaz-Horta *et al.*,³⁸ in 2012, after

excluding mutations in the most common autosomal-recessive nonsyndromic HL (*ARNSHL*) gene, *GJB2*, performed WES on samples from 20 unrelated multiplex consanguineous families with *ARNSHL*. They found twelve homozygous mutations in known deafness genes in 12 families. Thereupon, rare causative mutations in known *ARNSHL* genes were reliably determined via WES. More than 15 genes associated with syndromic or nonsyndromic HL have been identified using WES (Table 1).

Intellectual disabilities. Intellectual disability (ID), referred to as cognitive impairment or mental retardation, is usually defined by a considerably below-average score on tests of mental capability or intelligence and limitations in the ability to function in areas of daily life³⁹ ranging from mild to profound. It appears as nonsyndromic or associates with other clinical features in syndromic forms. ID accounts for up to 3% of the population in industrial countries.^{40,41} Cognition is a consequence of a sequential and simultaneous collaboration among a complex network of neuron and brain-expressed genes in the cellular and molecular levels. In addition to non-genetic factors, ID can have genetic bases including point mutations of single genes, large cytogenetic abnormalities and epigenetic alterations. More than 60 X-linked genes and seven autosomal genes for nonsyndromic intellectual disability have been recognized so far.^{42,43} Over thirty studies have used WES to identify the causal variants in various affected families including those with sporadic and familial nonsyndromic and syndromic forms of intellectual disability (Table 2). For instance, Caliskan *et al.*⁴³ identified the causal mutation, p.P182L in *TECR* gene, in a consanguineous family with five affected individuals. *De novo* mutations are involved in many sporadic cases of ID; 45 and 7 out of 74 studied cases are due to *de novo* and dominant mutations, respectively (Table 2). Correspondingly, WES could be an effective and interesting tool for identifying the genetic basis of such heterogeneous disorders such as ID.

Movement disorders

Movement disorders, a group of neurological conditions, cause abnormal voluntary or involuntary movements or slow, reduced

Table 1 The reported new genes in HL identified by WES

Type of HL	Disease (HL type)	Gene	Studied subjects	Reference
Syndromic	Autosomal-recessive retinitis pigmentosa (late-onset)	<i>USH1C</i>	Two Yemenite Jewish families	85
	D-bifunctional protein deficiency type IV (severe sensorineural)	<i>HSD17B4</i>	Two brothers	86
	Brown-Vialetto-Van Laere syndrome (Progressive sensorineural)	<i>SLC52A2</i>	Two extended Lebanese kindred	87
	MEGDEL syndrome	<i>SERAC1</i>	Two individuals	88
	Chudley-McCullough syndrome (Profound sensorineural)	<i>GPSM2</i>	Eight families	89
	CHIME syndrome (Conductive)	<i>PIGL</i>	Five families	90
	Fuchs corneal dystrophy (Progressive)	<i>LOXHD1</i>	Three large families	91
	Autosomal-dominant cerebellar ataxia, deafness and narcolepsy (ensorineural)	<i>DNMT1</i>	Three families	92
	Mandibulofacialdysostosis with microcephaly (Sensorineural)	<i>EFTUD2</i>	12 families	93
	Hereditary sensory and autonomic neuropathy with dementia and HL	<i>DNMT1</i>	Two American kindreds and one Japanese kindred	94
	Autosomal-recessive Carnevale, Malpuech, Michels, and oculo-skeletal-abdominal (OSA) syndromes	<i>MASP1</i>	Two consanguineous Turkish families	37
	Myhre syndrome	<i>SMAD4</i>	11 individuals	95
	Perrault syndrome (Sensorineural)	<i>HSD17B4</i>	One family	96
	Nonsyndromic	ADNSHL (Progressive)	<i>COCH</i>	One Chinese family
ARNSHL		<i>TSPEAR</i>	One consanguineous Iranian family	98
DFNB82		<i>GPSM2</i>	One consanguineous Palestinian family	36

Table 2 The identified new genes in ID by means of WES

Type of ID	Disease (inheritance)	Gene	No. of Studied families	Reference
Syndromic	Blepharophimosis-ptosis-intellectual-disability syndrome (AR)	<i>UBE3B</i>	Three families	99
	Hereditary spastic paraparesis (AR)	<i>TECPR2</i>	Three Jewish Bukharian families	100
	Intellectual-disability syndrome (AD)	<i>PACS1</i>	Two individuals	101
	Shprintzen-Goldberg syndrome (AD)	<i>SKI</i>	One individual	102
	Hyperphosphatasia with mental retardation syndrome (AR)	<i>PIGO</i>	Two individuals	103
	Acrodysostosis (AD)	<i>PDE4D</i>	Two individuals	104
	Wiedemann-Steiner syndrome (<i>de novo</i>)	<i>MLL</i>	Five individuals	105
	Dyggve-Melchior-Clausen syndrome (AR)	<i>RAB33B</i>	One family	106
	Coffin-Siris syndrome (<i>de novo</i>)	<i>ARID1B</i>	Three individuals	107
	Coffin-Siris syndrome (<i>de novo</i>)	<i>SMARCB1</i>	Five individuals	108
	Intellectual disability with neuronal migration defects (<i>de novo</i>)	<i>DYNC1H1</i>	One family	109
	Baraitser-Winter syndrome (<i>de novo</i>)	<i>ACTB and ACTG1</i>	Three families	110
	X-linked Joubert syndrome (X-linked)	<i>OFD1</i>	One family	111
	Genitopatellar syndrome (<i>de novo</i>)	<i>KAT6B</i>	Five individuals; three individuals	112,113
	Floating-Harbor syndrome (<i>de novo</i>)	<i>SRCAP</i>	Five individuals	114
	Myhre syndrome (<i>de novo</i>)	<i>SMAD4</i>	One individual;11 individuals	95,115
	Sjögren-Larsson syndrome (AR)	<i>ELOVL4</i>	Two individuals	83
	Say-Barber-Biesecker-Young-Simpson syndrome (<i>de novo</i>)	<i>KAT6B</i>	Three individuals	116
	KBG syndrome (AD)	<i>ANKRD11</i>	Two families	117
	3-M syndrome (AR)	<i>CCDC8</i>	Three families	118
	Adaptor protein complex 4 deficiency (AR)	<i>AP4S1</i>	Three families	119
	Nonsyndromic	Hyperphosphatasia mental retardation (HPMR) syndrome (AR)	<i>PIGV</i>	One family
Nonsyndromic intellectual disability (X-linked)		<i>HCFC1</i>	One individual	121
Autosomal-recessive primary microcephaly		<i>CEP135</i>	One northern Pakistan family	122
Autosomal-recessive nonsyndromic mental retardation		<i>TECR</i>	one family	43

movements, for example, ataxias, dystonias and so on. Fuchs *et al.*,⁴⁴ using exome sequencing in two families with primary torsion dystonia, identified a nonsense and a missense causative mutation in *GNAL* gene in two families.⁴⁴ In a study on a large essential tremor-affected family using exome sequencing, a *FUS* nonsense mutation was appeared to cause the disease.⁴⁵ Rosewich *et al.*⁴⁶ employed WES on three proband-parent trios and determined three heterozygous *de novo* missense mutations in alternating hemiplegia of childhood (AHC); they concluded mutation analysis of the *ATPIA3* gene in AHC patients could provide an accurate genetic diagnosis.⁴⁶ Similarly, WES of four siblings with pontocerebellar hypoplasia type 1 showed mutations in *EXOSC3* gene.⁴⁷ In addition, congenital mirror movements, involuntary movements of one side of the body that mirror intentional movements on the opposite side, are genetically heterogeneous and frequently inherited as autosomal-dominant mode of inheritance. By combining genome-wide linkage analysis and WES, Depienne *et al.*⁴⁸ identified heterozygous mutations introducing premature termination codons in *RAD51* gene in two families affected with congenital mirror movements. Weissbach *et al.*,⁴⁹ using a combined approach of genome-wide linkage analysis, WES and a systematic validation procedure in a German family with restless legs syndrome (RLS), found variants in four genes (*FAT2*, *ATRN*, *WWC2* and *PCDHA3*) that may cause RLS in this family.⁴⁹ RLS is one of the most common movement disorders in Europe and the United States.

Monogenic types of Diabetes

To identify the variants causing obesity in a patient suffering from permanent neonatal diabetes mellitus (NDM), a rare monogenic form of non-autoimmune diabetes, Bonnefond *et al.*²⁷ applied WES on the patient who had no mutation in *KCNJ11*, *ABCC8* and *INS* genes and other abnormalities in chromosome 6q24 region in their previous

studies. A *de novo* non-synonymous mutation, p.Q485H, was found in the *ABCC8* gene.

Maturity onset of the young (MODY), a heterogeneous type of diabetes showing an autosomal-dominant mode of inheritance, has an onset before the age of 25 years; it is due to a primary deficit in the pancreatic beta-cell function. Genetic causes of about 30% of MODY cases are still unknown (MODY-X). By using WES, a four-generation MODY-X family was studied to find the genetic causes of the condition. After co-segregation analysis within the family, Bonnefond *et al.*²⁷ found that a mutation (p.Glu227Lys) in *KCNJ11* gene is segregated with the disease. In addition to neonatal diabetes mellitus, *KCNJ11* has a role in MODY (MODY13).

Common diseases and complex disorders

Molecular genetics of complex phenotypes including common diseases and multigenic traits has been gradually developed; the common disease-common variant hypothesis (CD-CV) has been the main force for genome-wide association studies (GWAS). However, it turns out that common alleles are responsible for a fraction of complex traits. Thus, a rare variants-common disease (RV-CD) hypothesis is now being emerged.^{50,51} WES will identify rare and novel as well as common genetic variants in coding regions associated within complex and common traits.

Cardiovascular disease. Cardiovascular disease (CVD) is the major cause of death in the United States.⁵² On the basis of the etiology, there are two forms of CVDs: the rare form that is monogenic (Mendelian: includes structural cardiomyopathies, channelopathies and familial dyslipidemias) and the common one having polygenic/multifactorial causes. Recent advances have evolved our understanding of many cardiovascular diseases. Defining the role of genes associated with Mendelian forms of cardiovascular diseases

provides insights to other types of heart diseases and its managements (prevention/treatment) of the disorder. WES has been used to reveal genetic variants in affected individuals with rare forms of heart disease.

Cardiomyopathies may be presented as hypertrophic cardiomyopathy (HCM) or dilated cardiomyopathy (DCM). A family history of the disease could be traced in HCM form,⁵³ while about 20–35% of DCM may show a familial history of the disease.⁵⁴ HCM is mainly due to mutations of genes encoding sarcomere proteins; in turn, more heterogeneous etiology of DCM was shown.^{55,56} Meder *et al.*²¹ established a comprehensive genetic screening in patients with hereditary DCM or HCM in a cost-efficient manner using an array-based subgenomic enrichment systems. They found two microdeletions and four point mutations in six patients. WES provides a new way for the diagnosis of cardiomyopathies. Several new genes have been discovered in cardiomyopathies (Table 3). WES has characterized the causal gene of familial combined hypolipidemia (*ANGPTL3*),⁵⁷ severe hypercholesterolemia (*ABCG*)⁵⁸ and familial dilated cardiomyopathy (*BAG3*).⁵⁹ WES promises to revolutionize the genetic research about CVDs in the near future. Using this technology, identification of rare SNPs susceptible to CVDs is also possible.

Hypertension. One billion people suffer from hypertension worldwide. To date, a few genes causing Mendelian forms of hypertension have been described. More recently, Boyden *et al.*⁶⁰ studied 52 unrelated patients suffering from pseudohypoaldosteronism type II (PHAII), a rare Mendelian syndrome characterized by hypertension, hyperkalemia and metabolic acidosis; mutations in kelch-like 3 (*KLHL3*) or cullin3 (*CUL3*) were identified using WES in PHAII patients. Austin *et al.*,⁶¹ using WES, studied a three-generation family with multiple affected family members with pulmonary arterial hypertension and found a frameshift mutation in caveolin-1 (*CAV1*) gene. Familial hyperkalemic hypertension (FHHT), a Mendelian form of arterial hypertension, is partly defined by *WNK1* and *WNK4* gene mutations. By means of WES, *KLHL3*

mutations were determined as a third gene responsible for FHHT;⁶² elucidation of the role of genes involved in the signaling pathway regulating ion homeostasis, one of the impacts of WES, could provide molecular insights into understanding mechanisms controlling blood pressure.

Obesity and diabetes. Prevalence of the obesity is rapidly growing; 11% of adults aged 20 and over were obese in 2008 according to WHO. Obesity is a main potential risk factor for diabetes type 2 (T2D) and has usually a high predisposition for cardiovascular disorder and hypertension.⁶³ Molecular investigation of obesity is crucial for understanding the underlying mechanisms of regulation of adiposity and the pathophysiology of obesity and unraveling the molecular regulation of appetite. Finding and designing the strategies for obesity management (prevention and therapy) demand to define molecular factors causing the disease. Mendelian obesity is involved at least in 5% of the severely obese individuals.⁶³ WES facilitates identifying new highly penetrant SNPs. Until now, there is no a huge literature about using WES for detecting the underlying or predisposing SNPs among obese population. As the cost of WES is rapidly decreasing, sequencing of 1000 obese people is proposed to perform like 1000 Genomes Project; then, management of the obesity and associated disorders will be expected to be done in a well-understood manner.

Lehne *et al.*⁶⁴ investigated the distribution of association signal with respect to protein-coding genes for seven diseases including Crohn's disease, type 1 and type 2 diabetes, rheumatoid arthritis, hypertension, coronary artery disease and bipolar disorder from the Wellcome Trust Case Control Consortium. Their study showed that there is a consistently stronger association signal in coding than in noncoding regions for these disorders. Albrechtsen *et al.*⁶⁵ applied WES to identify novel associations of coding polymorphisms at minor allele frequencies (MAFs) >1% with common metabolic phenotypes. In stage 1 of their study, WES in 1000 cases with type 2 diabetes, BMI >27.5 kg m⁻² and hypertension and in 1000 controls was performed and they selected 16 192 SNPs associated with case-control status, from four selected annotation categories or from loci reported to associate with metabolic traits. In the next step (stage 2), these researchers genotyped the identified variants in 15 989 Danes to search for association with 12 metabolic phenotypes. Finally (stage 3), SNPs with potential associations were genotyped in a further 63 896 Europeans. Eventually, robust associations for coding SNPs in three genes were identified: *CD300LG* (fasting HDL cholesterol: MAF 3.5%), *COBLL1* (type 2 diabetes: MAF 12.5%) and *MACF1* (type 2 diabetes: MAF 23.4%). On the basis of these results, they concluded that coding SNPs with MAF above 1% might have no high effect size on the metabolic phenotypes. WES, therefore, is an appropriate approach to find causal and predisposing SNPs (Figure 1).

Cancer

Accumulation of genetic alterations during the life can lead to a malignant neoplasm or cancer. Because of heterogeneous entity of human cancers, multiple gene tests should be applied. This potentiality is allowed by WES at once. WES could highlight novel insights into cancer mechanisms;⁶⁶ identification of the DNA sequence of cancer cells in comparison with that of normal cells could help to reach an in-depth understanding of cancer. Sequence variations may influence the predisposition for cancer development. Using WES, it is feasible to check germline and somatic mutations in cancer (Table 4). Table 4 shows some of the identified genes involved in cancer development. Chang *et al.*⁶⁷ applied an exome-sequencing

Table 3 Reported WES approaches in cardiovascular diseases

Category	Disease (Inheritance)	Identified gene	Studied subjects	References
Cardiomyopathy	Recessive HCM	<i>MRPL44</i>	Two siblings	123
	Recessive HCM	<i>MTO1</i>	Two unrelated families	124
	Dominant ARVC	<i>DES</i>	One family	125
	Recessive HCM (Sengers syndrome)	<i>AGK</i>	One individual	126
	Recessive DCM	<i>GATAD1</i>	One family	127
	Recessive HCM	<i>MRPL3</i>	One family	128
	Recessive HCM	<i>AARS2</i>	One family	129
	Dominant DCM	<i>BAG3</i>	One	59
Heterotaxy cardiovascular	Recessive	<i>SHROOM3</i>	One	130
Familial combined hypolipidemia	Recessive	<i>ANGPTL3</i>	Two families	57
Channelopathy	X-linked intellectual disability, atrial fibrillation, cardiomegaly, congestive heart failure	<i>CLIC2</i>	Two males	131

Abbreviation: ARVC: arrhythmogenic right ventricular cardiomyopathy.

Table 4 Some of cancers studied by WES

Cancer	Description	Samples	Findings	Authors' conclusion	References
Gastric cancer	Multiple environmental etiologies and alternative pathways of carcinogenesis	22	Inactivating mutations of <i>ARID1A</i> and genes involved in chromatin remodeling	Importance of chromatin remodeling, in the molecular taxonomy	74
Gastric cancer		15	Mutations in chromatin remodeling genes (<i>ARID1A</i> , <i>MLL3</i> and <i>MLL</i>)	Somatic mutations of <i>FAT4</i> and <i>ARID1A</i> are probably key tumorigenic events	132
Prostate cancer	A clinically and pathologically heterogeneous disease	4	High level of intraprostatic heterogeneity	Personalized cancer treatment and biomarker discovery	133
Prostate cancer		112 tumor and normal tissue pairs	Mutations of <i>SPOP</i> , <i>MED12</i> and <i>FOXA1</i>	Role of <i>SPOP</i> mutations in a new molecular subtype of prostate cancer	134
Familial breast cancer	Malignancy of inner lining of milk ducts or the lobules	50	Median number of mutations was 10 in individuals	Intelligent sample and variant prioritization strategies to deliver robust genetic evidence of disease association needed in WES	135
Lung cancer	Non-small cell lung carcinoma	14	Mutations of <i>MRA5</i> as a second most frequently mutated gene	Altered extracellular matrix-remodeling as a possible cause of NSCLC	136
Lung cancer	Non-small cell lung cancer	31	Many newly genes including <i>CSMD3</i>	Mutational processes, cellular pathways and gene networks	76
Endometrial cancer	The most common gynecological malignancy	13	12 potential driver cancer genes	First unbiased view of somatic coding mutations	137
Endometrial cancer		13	High frequencies of somatic mutations in 8 genes	Frequent mutational disruption in the molecular pathogenesis of this cancer	138
Breast cancer	Malignancy of inner lining of milk ducts or the lobules	33	Mutations in DNA repair genes <i>FANCC</i> and <i>BLM</i>	Role of intra-family WES approaches	139
Melanoma	Malignant tumor of melanocytes	147	Somatic mutations in <i>RAC1</i> , <i>BRAF</i> and <i>NRAS</i>	Pharmacological inhibition of <i>RAC1</i> signaling effectors for therapeutic benefit	75
Melanoma		14	68 genes	The role of glutamate signaling pathway	140
Opisthorchisviverrini-related cholangiocarcinoma	Fatal bile duct cancer	8	10 newly implicated genes	Mutational landscape contributing to this cancer	141
Head and neck squamous cell carcinoma	The sixth most common cancer	32	Mutations in newly genes: <i>FBXW7</i> and <i>NOTCH1</i>	<i>NOTCH1</i> is likely a tumor suppressor gene but an oncogene in this tumor type	142
Renal carcinoma	Clear cell carcinoma	7	Several new cancer genes including <i>UTX</i> , <i>JARID1C</i> and <i>SETD2</i>	The somatic genetic architecture of ccRCC and the marked contribution of aberrant chromatin biology	69

technology using Roche Nimblegen capture paired with 454 sequencing to determine variations and mutations in eight commonly used cancer cell lines; they showed that this technology is able to identify sequence variations so that it can provide 95% concordance with an Affymetrix SNP Array 6.0 (Affymetrix, Santa Clara, CA, USA) performed on the same cell lines. These researchers mentioned that WES can be a reliable and cost-effective way to identify variations in cancer genomes.

In 2010, Summerer *et al.*⁶⁸ reported an automated exome-capture method for a subset of 115 cancer-related genes using microfluidic DNA arrays. WES was applied on a series of primary clear cell renal cell carcinoma (ccRCC).⁶⁹ A gene, *PBRM1*, encoding for SWI/SNF chromatin remodeling complex was found in this study.

Recently, it has been reported that mutations in *SF3B1* gene encoding a spliceosomal protein were identified in a distinct form of myelodysplastic syndromes (MDS) with ring sideroblasts.

Mutations of this gene are associated with a good prognosis.⁷⁰ In addition to complexity of variations in different types of cancers, mutations in oncogenes and tumor suppressor genes that initiate tumorigenesis should be considered. To confirm variations in tumor suppressor genes leading to loss of heterozygosity, comparison of the identified variants in diseased tissues with normal tissues from the same individual is required.

Finding the driver mutations is one of the major concerns in genomic analysis of cancer. Thus, several studies have attempted to identify driver mutations within the exome in various types of cancer including leukemias, myelomas and solid tumors.^{71–73}

WES is providing a detailed understanding of cancer pathways and unraveling the molecular mechanisms of cancer. Wang *et al.*⁷⁴ applied WES on 22 gastric cancer individuals and identified unreported mutated genes and pathway alterations; they particularly found genes involving in chromatin remodeling. Besides, WES is also being

applied for therapeutic aims, for example, identifying relevant pharmacogenetic variants and targeted gene–disease–drug interactions.^{75,76}

In brief, applications of WES in cancers are (1) somatic mutation detection, (2) driver mutation detection, (3) mutation network reconstruction and (4) identification of predisposing variants; even intra-family WES approaches to uncover cancer predisposition genes need to be considered for some cancers.⁷⁷

Human variome and ClinSeq projects

The human genome has about four million DNA sequence variability. The important fact is to characterize the clinical importance of these differences. In coming years, a huge amount of gene variations will be produced with application of NGS technology. The challenging issue is how to use these data into clinical practice. This explosively growing data should be managed in an organized manner so that clinicians and researchers could easily find the required information of previously characterized variations. In the era of the personalized medicine, collection of the variants based on the ethnicity and origin of patients would provide a valuable solution for management of the disease in near future.^{78–80} The Human Variome Project (<http://www.humanvariomeproject.org>) will collect systematically mutations that cause human disease on the basis of population. Gathering and characterizing all mutations and their effects would shed light on the road of management of the genetic disorders.

The NGS technology would help to identify not only the causal variants in genotype–phenotype correlation studies but also the modifier genes with the comparison of the alleles in the family members. The clinical phenotype is the result of several alleles with different pathogenic effect, which makes the interpretation of variations difficult.

Most studies are based on small number of samples for identifying the causal variants. To rely on the genetic information in clinic, much more information is needed. In practice, ClinSeq project aims to provide enough amount of genome for study, number of subjects and the associated clinical data. In a single-gene study, only a single gene with clinical phenotype is evaluated, although the 1000 Genome Project aims to evaluate large number of samples with no phenotype evaluation.

The ClinSeq project was organized to use the WGS data of massively parallel sequencing of different patients suffering from diseases to determine the genotype–phenotype association of variants. Identification of variants and knowledge on genotype–phenotype correlation could be used to change the clinical management of patients.⁸¹ For example, pharmacological treatment is useful for patients affected with familial hypercholesterolemia that are dietary dependent.⁸² The pilot study started with large-scale medical sequencing of 1000 participants for atherosclerotic heart disease and makes a model relating the genotype to the phenotype.

The goal is to use genetic data in clinical practice improving our understanding of genetic basis of diseases in health care, diagnosis and therapy. Preliminary data from ClinSeq study showed that results should be interpreted and returned to the individuals. Depending on the disease, variants should be reported. High penetrance variants with high reliability of finding the disease are important to be noticed. They may be used in management and treatment of the disease.

CONCERNS IN WES

Massively parallel DNA sequencing systems have increased the throughput of sequencing compared with the classical sequencing method. Reduced cost of WES rather than WGS led to its applications

in the clinical genetics. As WES focuses on only ~1% of the genome and is limited to the coding and splice-site variants in annotated genes, it is suitable for gene discovery in highly penetrant Mendelian diseases. Also, exon capture step may potentiate technical biases along the procedure and limits its usage in detecting copy-number variants as well as in genomic regions where capture is less efficient.

During search for causative variants, availability of the data throughout the genome makes the clinical phenotype less important, when big families with multiple affected individuals or multiple pedigrees with the same genetically homogeneous disorder are available; if only simplex cases or small families are studied, then detailed clinical evaluations and laboratory tests have important roles to narrow down the candidate causative variants. Aldahmesh *et al.*,⁸³ for example, reported two individuals showing more severe neurologic phenotype of Sjögren-Larsson syndrome (SLS; MIM 270200), characterized by ichthyosis, seizures, intellectual disability and spasticity. However, fibroblast fatty aldehyde dehydrogenase (FALDH) deficiency, the hallmark of SLS, was assessed to be normal in their patients. Biallelic *ELOVL4* mutations were found in these pseudo-SLS patients. Thus, differential diagnosis of the condition may accelerate and be advantageous to find the causal variants. However, as we discussed before, WES is an excellent tool to help finding the correct diagnosis even if the initial diagnosis is incorrect.

One of the most important facts in medical genetics is the ethical issue. Informed consent, which is a guiding principle of medical research, may not be so meaningful because information derived from DNA does not only belong to participant but also reveals properties of the relatives and sometimes original population. As WES provides the massive scale of DNA variants, using it in a clinical setting needs to meet more detailed legal and ethical standards. How many relatives of the participant should be informed? How much information is enough to say that a participant is informed? Who should and how to ensure the patient and relatives that their information will not be misused in their employments and insurance companies? Besides, who does guarantee that their information will be secured in future research for their offspring?

The ethical issues about the privacy, confidentiality and return of results are major concerns and need more evaluation by professionals. However, employment and insurance issues should be solved; a comprehensive agenda considering all possible consequences of WES need to be established to use WES for improving human health.

Briefly, according to Tabor *et al.*,⁸⁴ the informed consent for NGS analysis is provided for the participants giving description of the overall experiment, explanation of the risk of sharing the data, privacy of the data and explication to return the results. The consent explains the return of the results in groups including the related causal variant, overall variants and variants associated with the common diseases. It is optional for participants whether or not to receive the research results. It is suggested that a framework for the release of information for the family and participants should be dependent on their preferences and could be changed over time.

CONCLUSION

Understanding the genetic variants provides appealing insights into the human disease for prevention strategies, diagnostic applications and therapeutic methods. Systematic resequencing of human DNA (large-scale medical sequencing) is revolutionizing medicine.⁸² As whole-genome sequencing is still costly, WES is a temporary alternative to get a picture of the coding genome. Sequence data production, which is exponentially growing, should be managed and

organized so that the information to be available for researchers and clinicians in a systematic framework. Further, more user-friendly bioinformatics' tools are demanded to be established for improvement of current knowledge about genetic disease management including counseling, prevention and treatment/therapies and drug responses. Considering this impact of DNA sequencing, a well-understood personalized medicine is expected in coming years.

WES is charming its way in research, diagnostic and clinical setting; it has been used for variant detection among both common and rare diseases as well as SNPs associations and pharmacogenetics. It is recommended that for clinical and diagnostic applications of the NGS, it is better to focus on specific human genes within the pathways for faster diagnosis. This could be based on our knowledge of different developmental biology pathways (interactome) relating to the disease. In addition, targeted capture is suitable in a diagnostic setting, especially for those diseases caused by numerous genes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank Dr Bahman Akbari for critical reading of the manuscript.

- Sanger, F., Nicklen, S. & Coulson, A. R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl Acad. Sci. USA* **74**, 5463–5467 (1977).
- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp. Quant. Biol.* **51** (Pt 1), 263–273 (1986).
- Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G. *et al*. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).
- Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G. *et al*. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature* **409**, 928–933 (2001).
- Venter, J. C. A part of the human genome sequence. *Science* **299**, 1183–1184 (2003).
- Seal, R. L., Gordon, S. M., Lush, M. J., Wright, M. W. & Bruford, E. A. genenames.org: the HGNC resources in 2011. *Nucleic Acids Res.* **39**, D514–D519 (2011).
- Rizzo, J. M. & Buck, M. J. Key principles and clinical applications of "next-generation" DNA sequencing. *Cancer Prev. Res.* **5**, 887–900 (2012).
- Botstein, D. & Risch, N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat. Genet.* **33** (Suppl), 228–237 (2003).
- Majewski, J., Schwartzentruber, J., Lalonde, E., Montpetit, A. & Jabado, N. What can exome sequencing do for you? *J. Med. Genet.* **48**, 580–589 (2011).
- Bonasio, R., Tu, S. & Reinberg, D. Molecular signals of epigenetic states. *Science* **330**, 612–616 (2010).
- Singleton, A. B. Exome sequencing: a transformative technology. *Lancet Neurol.* **10**, 942–946 (2011).
- Consortium, G. P. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
- Marian, A. J. Medical DNA sequencing. *Curr. Opin. Cardiol.* **26**, 175–180 (2011).
- Rabbani, B., Mahdieh, N., Hosomichi, K., Nakaoka, H. & Inoue, I. Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. *J. Hum. Genet.* **57**, 621–632 (2012).
- Ng, S. B., Turner, E. H., Robertson, P. D., Flygare, S. D., Bigham, A. W., Lee, C. *et al*. Targeted capture and massively parallel sequencing of 12 human exomes. *Nature* **461**, 272–276 (2009).
- Choi, M., Scholl, U. I., Ji, W., Liu, T., Tikhonova, I. R., Zumbo, P. *et al*. Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl Acad. Sci. USA* **106**, 19096–19101 (2009).
- Majewski, J., Wang, Z., Lopez, I., Al Humaid, S., Ren, H., Racine, J. *et al*. A new ocular phenotype associated with an unexpected but known systemic disorder and mutation: novel use of genomic diagnostics and exome sequencing. *J. Med. Genet.* **48**, 593–596 (2011).
- Cullinane, A. R., Vilboux, T., O'Brien, K., Curry, J. A., Maynard, D. M., Carlson-Donohoe, H. *et al*. Homozygosity mapping and whole-exome sequencing to detect SLC45A2 and G6PC3 mutations in a single patient with oculocutaneous albinism and neutropenia. *J. Invest. Dermatol.* **131**, 2017–2025 (2011).
- Hoischen, A., Gilissen, C., Arts, P., Wieskamp, N., van der Vliet, W., Vermeer, S. *et al*. Massively parallel sequencing of ataxia genes after array-based enrichment. *Hum. Mutat.* **31**, 494–499 (2010).
- Shen, P., Wang, W., Krishnakumar, S., Palm, C., Chi, A. K., Enns, G. M. *et al*. High-quality DNA sequence capture of 524 disease candidate genes. *Proc. Natl Acad. Sci. USA* **108**, 6549–6554 (2011).
- Meder, B., Haas, J., Keller, A., Heid, C., Just, S., Borries, A. *et al*. Targeted next-generation sequencing for the molecular genetic diagnostics of cardiomyopathies. *Circ. Cardiovasc. Genet.* **4**, 110–122 (2011).
- Jones, M. A., Bhide, S., Chin, E., Ng, B. G., Rhodenizer, D., Zhang, V. W. *et al*. Targeted polymerase chain reaction-based enrichment and next generation sequencing for diagnostic testing of congenital disorders of glycosylation. *Genet. Med.* **13**, 921–932 (2011).
- Ku, C. S., Naidoo, N. & Pawitan, Y. Revisiting Mendelian disorders through exome sequencing. *Hum. Genet.* **129**, 351–370 (2011).
- Kingsmore, S. F. & Saunders, C. J. Deep sequencing of patient genomes for disease diagnosis: when will it become routine? *Sci. Transl. Med.* **3**, 87ps23 (2011).
- Walker, R. H., Schulz, V. P., Tikhonova, I. R., Mahajan, M. C., Mane, S., Arroyo Muniz, M. *et al*. Genetic diagnosis of neuroacanthocytosis disorders using exome sequencing. *Mov. Disord.* **27**, 539–543 (2011).
- Worthey, E. A., Mayer, A. N., Syverson, G. D., Helbling, D., Bonacci, B. B., Decker, B. *et al*. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet. Med.* **13**, 255–262 (2011).
- Bonnefond, A., Durand, E., Sand, O., De Graeve, F., Gallina, S., Busiah, K. *et al*. Molecular diagnosis of neonatal diabetes mellitus using next-generation sequencing of the whole exome. *PLoS One* **5**, e13630 (2010).
- Chiu, R. W., Chan, K. C., Gao, Y., Lau, Y. Y., Zheng, W., Leung, T. Y. *et al*. Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma. *Proc. Natl Acad. Sci. USA* **105**, 20458–20463 (2008).
- Chen, E. Z., Chiu, R. W., Sun, H., Akolekar, R., Chan, K. C., Leung, T. Y. *et al*. Noninvasive prenatal diagnosis of fetal trisomy 18 and trisomy 13 by maternal plasma DNA sequencing. *PLoS One* **6**, e21791 (2011).
- Sehnert, A. J., Rhees, B., Comstock, D., de Feo, E., Heilek, G., Burke, J. *et al*. Optimal detection of fetal chromosomal abnormalities by massively parallel DNA sequencing of cell-free fetal DNA from maternal blood. *Clin. Chem.* **57**, 1042–1049 (2011).
- Bell, C. J., Dinwiddie, D. L., Miller, N. A., Hateley, S. L., Ganusova, E. E., Mudge, J. *et al*. Carrier testing for severe childhood recessive diseases by next-generation sequencing. *Sci. Transl. Med.* **3**, 65ra4 (2011).
- Jiao, Y., Shi, C., Edil, B. H., de Wilde, R. F., Klimstra, D. S., Maitra, A. *et al*. DAXX/ATRAX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors. *Science* **331**, 1199–1203 (2011).
- Morton, C. C. & Nance, W. E. Newborn hearing screening—a silent revolution. *N. Engl. J. Med.* **354**, 2151–2164 (2006).
- Finsterer, J. & Fellinger, J. Nuclear and mitochondrial genes mutated in nonsyndromic impaired hearing. *Int. J. Pediatr. Otorhinolaryngol.* **69**, 621–647 (2005).
- Mahdieh, N., Rabbani, B., Wiley, S., Akbari, M. T. & Zeinali, S. Genetic causes of nonsyndromic hearing loss in Iran in comparison with other populations. *J. Hum. Genet.* **55**, 639–648 (2010).
- Walsh, T., Shahin, H., Elkan-Miller, T., Lee, M. K., Thornton, A. M., Roebe, W. *et al*. Whole exome sequencing and homozygosity mapping identify mutation in the cell polarity protein GSPM2 as the cause of nonsyndromic hearing loss DFNB82. *Am. J. Hum. Genet.* **87**, 90–94 (2010).
- Sirmaci, A., Walsh, T., Akay, H., Spiliopoulos, M., Sakalar, Y. B., Hasanefendioglu-Bayrak, A. *et al*. MASP1 mutations in patients with facial, umbilical, coccygeal, and auditory findings of Carnevale, Malpuech, OSA, and Michels syndromes. *Am. J. Hum. Genet.* **87**, 679–686 (2010).
- Diaz-Horta, O., Duman, D., Foster, J., Sirmaci, A., Gonzalez, M., Mahdieh, N. *et al*. Whole-exome sequencing efficiently detects rare mutations in autosomal recessive nonsyndromic hearing loss. *PLoS One* **7**, e50628 (2012).
- Moeschler, J. B. Genetic evaluation of intellectual disabilities. *Semin. Pediatr. Neurol.* **15**, 2–9 (2008).
- Curry, C. J., Stevenson, R. E., Aughton, D., Byrne, J., Carey, J. C., Cassidy, S. *et al*. Evaluation of mental retardation: recommendations of a Consensus Conference: American College of Medical Genetics. *Am. J. Med. Genet.* **72**, 468–477 (1997).
- Chelly, J., Khelifaoui, M., Francis, F., Cherif, B. & Bienvenu, T. Genetics and pathophysiology of mental retardation. *Eur. J. Hum. Genet.* **14**, 701–713 (2006).
- Molinari, F., Rio, M., Meskenaitė, V., Encha-Razavi, F., Auge, J., Bacq, D. *et al*. Truncating neurotrophin mutation in autosomal recessive nonsyndromic mental retardation. *Science* **298**, 1779–1781 (2002).
- Caliskan, M., Chong, J. X., Uricchio, L., Anderson, R., Chen, P., Sougnez, C. *et al*. Exome sequencing reveals a novel mutation for autosomal recessive non-syndromic mental retardation in the TECR gene on chromosome 19p13. *Hum. Mol. Genet.* **20**, 1285–1289 (2011).
- Fuchs, T., Saunders-Pullman, R., Masuho, I., Luciano, M. S., Raymond, D., Factor, S. *et al*. Mutations in GNAL cause primary torsion dystonia. *Nat. Genet.* **45**, 88–92 (2012).
- Merner, N. D., Girard, S. L., Catoire, H., Bourassa, C. V., Belzil, V. V., Riviere, J. B. *et al*. Exome sequencing identifies FUS mutations as a cause of essential tremor. *Am. J. Hum. Genet.* **91**, 313–319 (2012).
- Rosewich, H., Thiele, H., Ohlenbusch, A., Maschke, U., Altmüller, J., Frommolt, P. *et al*. Heterozygous de-novo mutations in ATP1A3 in patients with alternating hemiplegia of childhood: a whole-exome sequencing gene-identification study. *Lancet Neurol.* **11**, 764–773 (2012).

- 47 Wan, J., Yourshaw, M., Mamsa, H., Rudnik-Schoneborn, S., Menezes, M. P., Hong, J. E. *et al.* Mutations in the RNA exosome component gene EXOSC3 cause pontocerebellar hypoplasia and spinal motor neuron degeneration. *Nat. Genet.* **44**, 704–708 (2012).
- 48 Depienne, C., Bouteiller, D., Meneret, A., Billot, S., Groppa, S., Klebe, S. *et al.* RAD51 Haploinsufficiency Causes Congenital Mirror Movements in Humans. *Am. J. Hum. Genet.* **90**, 301–307 (2012).
- 49 Weissbach, A., Siegesmund, K., Bruggemann, N., Schmidt, A., Kasten, M., Pichler, I. *et al.* Exome sequencing in a family with restless legs syndrome. *Mov. Disord.* **27**, 1686–1689 (2012).
- 50 Marian, A. J. Molecular genetic studies of complex phenotypes. *Transl. Res.* **159**, 64–79 (2012).
- 51 Schork, N. J., Murray, S. S., Frazer, K. A. & Topol, E. J. Common vs rare allele hypotheses for complex diseases. *Curr. Opin. Genet. Dev.* **19**, 212–219 (2009).
- 52 O'Donnell, C. J. & Nabel, E. G. Genomics of cardiovascular disease. *N. Engl. J. Med.* **365**, 2098–2109 (2011).
- 53 Seidman, J. G. & Seidman, C. The genetic basis for cardiomyopathy: from mutation identification to mechanistic paradigms. *Cell* **104**, 557–567 (2001).
- 54 Colombo, M. G., Botto, N., Vittorini, S., Paradossi, U. & Andreassi, M. G. Clinical utility of genetic tests for inherited hypertrophic and dilated cardiomyopathies. *Cardiovasc. Ultrasound* **6**, 62 (2008).
- 55 Voelkerding, K. V., Dames, S. & Durtschi, J. D. Next generation sequencing for clinical diagnostics—principles and application to targeted resequencing for hypertrophic cardiomyopathy: a paper from the 2009 William Beaumont Hospital Symposium on Molecular Pathology. *J. Mol. Diagn.* **12**, 539–551 (2010).
- 56 Dellefave, L. & McNally, E. M. The genetics of dilated cardiomyopathy. *Curr. Opin. Cardiol.* **25**, 198–204 (2010).
- 57 Musunuru, K., Pirruccello, J. P., Do, R., Peloso, G. M., Guiducci, C., Sougnez, C. *et al.* Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N. Engl. J. Med.* **363**, 2220–2227 (2010).
- 58 Rios, J., Stein, E., Shendure, J., Hobbs, H. H. & Cohen, J. C. Identification by whole-genome resequencing of gene defect responsible for severe hypercholesterolemia. *Hum. Mol. Genet.* **19**, 4313–4318 (2011).
- 59 Norton, N., Li, D., Rieder, M. J., Siegfried, J. D., Rampersaud, E., Zuchner, S. *et al.* Genome-wide studies of copy number variation and exome sequencing identify rare variants in BAG3 as a cause of dilated cardiomyopathy. *Am. J. Hum. Genet.* **88**, 273–282 (2011).
- 60 Boyden, L. M., Choi, M., Choate, K. A., Nelson-Williams, C. J., Farhi, A., Toka, H. R. *et al.* Mutations in kelch-like 3 and cullin 3 cause hypertension and electrolyte abnormalities. *Nature* **482**, 98–102 (2012).
- 61 Austin, E. D., Ma, L., Leduc, C., Berman Rosenzweig, E., Borczuk, A., Phillips, J. A. *et al.* Whole Exome Sequencing to Identify a Novel Gene (Caveolin-1) Associated with Human Pulmonary Arterial Hypertension. *Circ. Cardiovasc. Genet.* **5**, 336–343 (2012).
- 62 Louis-Dit-Picard, H., Barc, J., Trujillano, D., Miserey-Lenkei, S., Bouatia-Naji, N., Pylipenko, O. *et al.* KLHL3 mutations cause familial hyperkalemic hypertension by impairing ion transport in the distal nephron. *Nat. Genet.* **44**, 456–460, S1–3 (2012).
- 63 Blakemore, A. I. & Froguel, P. Investigation of Mendelian forms of obesity holds out the prospect of personalized medicine. *Ann. N. Y. Acad. Sci.* **1214**, 180–189 (2010).
- 64 Lehne, B., Lewis, C. M. & Schlitt, T. Exome localization of complex disease association signals. *BMC Genomics* **12**, 92 (2011).
- 65 Albrechtsen, A., Grarup, N., Li, Y., Sparso, T., Tian, G., Cao, H. *et al.* Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia* **56**, 298–310 (2012).
- 66 Yan, X. J., Xu, J., Gu, Z. H., Pan, C. M., Lu, G., Shen, Y. *et al.* Exome sequencing identifies somatic mutations of DNA methyltransferase gene DNMT3A in acute monocytic leukemia. *Nat. Genet.* **43**, 309–315 (2011).
- 67 Chang, H., Jackson, D. G., Kayne, P. S., Ross-Macdonald, P. B., Ryseck, R. P. & Siemers, N. O. Exome sequencing reveals comprehensive genomic alterations across eight cancer cell lines. *PLoS One* **6**, e21097 (2011).
- 68 Summerer, D., Hevroni, D., Jain, A., Oldenburger, O., Parker, J., Caruso, A. *et al.* A flexible and fully integrated system for amplification, detection and genotyping of genomic DNA targets based on microfluidic oligonucleotide arrays. *N. Biotechnol.* **27**, 149–155 (2010).
- 69 Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C. K., Stephens, P. *et al.* Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature* **469**, 539–542 (2011).
- 70 Makishima, H., Jankowska, A. M., McDevitt, M. A., O'Keefe, C., Dujardin, S., Cazzolli, H. *et al.* CBL, CBLB, TET2, ASXL1, and IDH1/2 mutations and additional chromosomal aberrations constitute molecular events in chronic myelogenous leukemia. *Blood* **117**, e198–e206 (2011).
- 71 Bignell, G. R., Greenman, C. D., Davies, H., Butler, A. P., Edkins, S., Andrews, J. M. *et al.* Signatures of mutation and selection in the cancer genome. *Nature* **463**, 893–898 (2010).
- 72 Dalglish, G. L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A. *et al.* Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. *Nature* **463**, 360–363 (2010).
- 73 Morin, R. D., Johnson, N. A., Severson, T. M., Mungall, A. J., An, J., Goya, R. *et al.* Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. *Nat. Genet.* **42**, 181–185 (2010).
- 74 Wang, K., Kan, J., Yuen, S. T., Shi, S. T., Chu, K. M., Law, S. *et al.* Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* **43**, 1219–1223 (2011).
- 75 Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchicocchi, A., McCusker, J. P. *et al.* Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nat. Genet.* **44**, 1006–1014 (2012).
- 76 Liu, P., Morrison, C., Wang, L., Xiong, D., Vedell, P., Cui, P. *et al.* Identification of somatic mutations in non-small cell lung carcinomas using whole-exome sequencing. *Carcinogenesis* **33**, 1270–1276 (2012).
- 77 Nikolaev, S. I., Sotiriou, S. K., Pateras, I. S., Santoni, F., Sougioultzis, S., Edgren, H. *et al.* A single-nucleotide substitution mutator phenotype revealed by exome sequencing of human colon adenomas. *Cancer Res.* **72**, 6279–6289 (2012).
- 78 Ring, H. Z., Kwok, P. Y. & Cotton, R. G. Human Variome Project: an international collaboration to catalogue human genetic variation. *Pharmacogenomics* **7**, 969–972 (2006).
- 79 Cotton, R. G., Appelbe, W., Auerbach, A. D., Becker, K., Bodmer, W., Boone, D. J. *et al.* Recommendations of the 2006 Human Variome Project meeting. *Nat. Genet.* **39**, 433–436 (2007).
- 80 Mahdieh, N., Bagherian, H., Shirkavand, A., Sharafi, M. & Zeinali, S. High level of intrafamilial phenotypic variability of non-syndromic hearing loss in a Lur family due to delE120 mutation in GJB2 gene. *Int. J. Pediatr. Otorhinolaryngol.* **74**, 1089–1091 (2010).
- 81 Biesecker, L. G. Opportunities and challenges for the integration of massively parallel genomic sequencing into clinical practice: lessons from the ClinSeq project. *Genet. Med.* **14**, 393–398 (2012).
- 82 Biesecker, L. G., Mullikin, J. C., Facio, F. M., Turner, C., Cherukuri, P. F., Blakesley, R. W. *et al.* The ClinSeq Project: piloting large-scale genome sequencing for research in genomic medicine. *Genome Res.* **19**, 1665–1674 (2009).
- 83 Aldahmesh, M. A., Mohamed, J. Y., Alkuraya, H. S., Verma, I. C., Puri, R. D., Alaiya, A. A. *et al.* Recessive mutations in ELOVL4 cause ichthyosis, intellectual disability, and spastic quadriplegia. *Am. J. Hum. Genet.* **89**, 745–750 (2011).
- 84 Tabor, H. K., Stock, J., Brazg, T., McMillin, M. J., Dent, K. M., Yu, J. H. *et al.* Informed consent for whole genome sequencing: A qualitative analysis of participant expectations and perceptions of risks, benefits, and harms. *Am. J. Med. Genet. A.* **158**, 1310–1319 (2012).
- 85 Khateb, S., Zelinger, L., Ben-Yosef, T., Merin, S., Crystal-Shalit, O., Gross, M. *et al.* Exome sequencing identifies a founder frameshift mutation in an alternative exon of USH1C as the cause of autosomal recessive retinitis pigmentosa with late-onset hearing loss. *PLoS One* **7**, e51566 (2012).
- 86 McMillan, H. J., Worthylake, T., Schwartzentruber, J., Gottlieb, C. C., Lawrence, S. E., Mackenzie, A. *et al.* Specific combination of compound heterozygous mutations in 17beta-hydroxysteroid dehydrogenase type 4 (HSD17B4) defines a new subtype of D-bifunctional protein deficiency. *Orphanet J. Rare Dis.* **7**, 90 (2012).
- 87 Johnson, J. O., Gibbs, J. R., Megarbane, A., Urtizberea, J. A., Hernandez, D. G., Foley, A. R. *et al.* Exome sequencing reveals riboflavin transporter mutations as a cause of motor neuron disease. *Brain* **135**, 2875–2882 (2012).
- 88 Wortmann, S. B., Vaz, F. M., Gardeitchik, T., Vissers, L. E., Renkema, G. H., Schuurts-Hoeijmakers, J. H. *et al.* Mutations in the phospholipid remodeling gene SERAC1 impair mitochondrial function and intracellular cholesterol trafficking and cause dystonia and deafness. *Nat. Genet.* **44**, 797–802 (2012).
- 89 Doherty, D., Chudley, A. E., Coghlan, G., Ishak, G. E., Innes, A. M., Lemire, E. G. *et al.* GPSM2 mutations cause the brain malformations and hearing loss in Chudley-McCullough syndrome. *Am. J. Hum. Genet.* **90**, 1088–1093 (2012).
- 90 Ng, B. G., Hackmann, K., Jones, M. A., Eroshkin, A. M., He, P., Williams, R. *et al.* Mutations in the glycosylphosphatidylinositol gene PIGL cause CHIME syndrome. *Am. J. Hum. Genet.* **90**, 685–688 (2012).
- 91 Riazuddin, S. A., Parker, D. S., McGlumphy, E. J., Oh, E. C., Iliff, B. W., Schmedt, T. *et al.* Mutations in LOXHD1, a recessive-deafness locus, cause dominant late-onset Fuchs corneal dystrophy. *Am. J. Hum. Genet.* **90**, 533–539 (2012).
- 92 Winkelmann, J., Lin, L., Schormair, B., Kornum, B. R., Faraco, J., Plazzi, G. *et al.* Mutations in DNMT1 cause autosomal dominant cerebellar ataxia, deafness and narcolepsy. *Hum. Mol. Genet.* **21**, 2205–2210 (2012).
- 93 Lines, M. A., Huang, L., Schwartzentruber, J., Douglas, S. L., Lynch, D. C., Beaulieu, C. *et al.* Haploinsufficiency of a Spliceosomal GTPase Encoded by EFTUD2 Causes Mandibulofacial Dysostosis with Microcephaly. *Am. J. Hum. Genet.* **90**, 369–377 (2012).
- 94 Klein, C. J., Botuyan, M. V., Wu, Y., Ward, C. J., Nicholson, G. A., Hammans, S. *et al.* Mutations in DNMT1 cause hereditary sensory neuropathy with dementia and hearing loss. *Nat. Genet.* **43**, 595–600 (2011).
- 95 Le Goff, C., Mahaut, C., Abhyankar, A., Le Goff, W., Serre, V., Afenjar, A. *et al.* Mutations at a single codon in Mad homology 2 domain of SMAD4 cause Myhre syndrome. *Nat. Genet.* **44**, 85–88 (2011).
- 96 Pierce, S. B., Walsh, T., Chisholm, K. M., Lee, M. K., Thornton, A. M., Fiumara, A. *et al.* Mutations in the DBP-deficiency protein HSD17B4 cause ovarian dysgenesis, hearing loss, and ataxia of Perrault Syndrome. *Am. J. Hum. Genet.* **87**, 282–288 (2010).
- 97 Gao, J., Xue, J., Chen, L., Ke, X., Qi, Y. & Liu, Y. Whole exome sequencing identifies a novel DFNA9 mutation, C162Y. *Clin. Genet.* **83**, 477–481 (2013).
- 98 Delmaghani, S., Aghaie, A., Michalski, N., Bonnet, C., Weil, D. & Petit, C. Defect in the gene encoding the EAR/EPTP domain-containing protein TSPEAR causes DFNB98 profound deafness. *Hum. Mol. Genet.* **21**, 3835–3844 (2012).
- 99 Basel-Vanagaitte, L., Dallapiccola, B., Ramirez-Solis, R., Segref, A., Thiele, H., Edwards, A. *et al.* Deficiency for the ubiquitin ligase UBE3B in a

- blepharophimosis-ptosis-intellectual-disability syndrome. *Am. J. Hum. Genet.* **91**, 998–1010 (2012).
- 100 Oz-Levi, D., Ben-Zeev, B., Ruzzo, E. K., Hitomi, Y., Gelman, A., Pelak, K. *et al.* Mutation in *TECP2* reveals a role for autophagy in hereditary spastic paraparesis. *Am. J. Hum. Genet.* **91**, 1065–1072 (2012).
- 101 Schuurs-Hoeijmakers, J. H., Oh, E. C., Vissers, L. E., Swinkels, M. E., Gilissen, C., Willemsen, M. A. *et al.* Recurrent de novo mutations in *PACS1* cause defective cranial-neural-crest migration and define a recognizable intellectual-disability syndrome. *Am. J. Hum. Genet.* **91**, 1122–1127 (2012).
- 102 Carmignac, V., Thevenon, J., Ades, L., Callewaert, B., Julia, S., Thauvin-Robinet, C. *et al.* In-frame mutations in exon 1 of *SKL* cause dominant Shprintzen-Goldberg syndrome. *Am. J. Hum. Genet.* **91**, 950–957 (2012).
- 103 Krawitz, P. M., Murakami, Y., Hecht, J., Kruger, U., Holder, S. E., Mortier, G. R. *et al.* Mutations in *PIGO*, a member of the GPI-anchor-synthesis pathway, cause hyperphosphatasia with mental retardation. *Am. J. Hum. Genet.* **91**, 146–151 (2012).
- 104 Michot, C., Le Goff, C., Goldenberg, A., Abhyankar, A., Klein, C., Kinning, E. *et al.* Exome sequencing identifies *PDE4D* mutations as another cause of acrodysostosis. *Am. J. Hum. Genet.* **90**, 740–745 (2012).
- 105 Jones, W. D., Dafou, D., McEntagart, M., Woollard, W. J., Elmslie, F. V., Holder-Espinasse, M. *et al.* De novo mutations in *MLL* cause Wiedemann-Steiner syndrome. *Am. J. Hum. Genet.* **91**, 358–364 (2012).
- 106 Alshammari, M. J., Al-Otaibi, L. & Alkuraya, F. S. Mutation in *RAB33B*, which encodes a regulator of retrograde Golgi transport, defines a second Dyggve–Melchior-Clausen locus. *J. Med. Genet.* **49**, 455–461 (2012).
- 107 Santen, G. W., Aten, E., Sun, Y., Almomani, R., Gilissen, C., Nielsen, M. *et al.* Mutations in *SWI/SNF* chromatin remodeling complex gene *ARID1B* cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 379–380 (2012).
- 108 Tsurusaki, Y., Okamoto, N., Ohashi, H., Kosho, T., Imai, Y., Hibi-Ko, Y. *et al.* Mutations affecting components of the *SWI/SNF* complex cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 376–378 (2012).
- 109 Willemsen, M. H., Vissers, L. E., Willemsen, M. A., van Bon, B. W., Kroes, T., de Ligt, J. *et al.* Mutations in *DYNC1H1* cause severe intellectual disability with neuronal migration defects. *J. Med. Genet.* **49**, 179–183 (2012).
- 110 Riviere, J. B., van Bon, B. W., Hoischen, A., Kholmanskikh, S. S., O’Roak, B. J., Gilissen, C. *et al.* De novo mutations in the actin genes *ACTB* and *ACTG1* cause Baraitser-Winter syndrome. *Nat. Genet.* **44**, 440–444, S1–2 (2012).
- 111 Field, M., Scheffer, I. E., Gill, D., Wilson, M., Christie, L., Shaw, M. *et al.* Expanding the molecular basis and phenotypic spectrum of X-linked Joubert syndrome associated with *OFD1* mutations. *Eur. J. Hum. Genet.* **20**, 806–809 (2012).
- 112 Simpson, M. A., Deshpande, C., Dafou, D., Vissers, L. E., Woollard, W. J., Holder, S. E. *et al.* De Novo Mutations of the Gene Encoding the Histone Acetyltransferase *KAT6B* Cause Genitopatellar Syndrome. *Am. J. Hum. Genet.* **90**, 290–294 (2012).
- 113 Campeau, P. M., Kim, J. C., Lu, J. T., Schwartzentruber, J. A., Abdul-Rahman, O. A., Schlaubitz, S. *et al.* Mutations in *KAT6B*, Encoding a Histone Acetyltransferase, Cause Genitopatellar Syndrome. *Am. J. Hum. Genet.* **90**, 282–289 (2012).
- 114 Hood, R. L., Lines, M. A., Nikkel, S. M., Schwartzentruber, J., Beaulieu, C., Nowaczyk, M. J. *et al.* Mutations in *SRCAP*, Encoding *SNF2*-Related *CREBBP* Activator Protein, Cause Floating-Harbor Syndrome. *Am. J. Hum. Genet.* **90**, 308–313 (2012).
- 115 Caputo, V., Cianetti, L., Niceta, M., Carta, C., Ciolfi, A., Bocchinfuso, G. *et al.* A restricted spectrum of mutations in the *SMAD4* tumor-suppressor gene underlies myhre syndrome. *Am. J. Hum. Genet.* **90**, 161–169 (2012).
- 116 Clayton-Smith, J., O’Sullivan, J., Daly, S., Bhaskar, S., Day, R., Anderson, B. *et al.* Whole-exome-sequencing identifies mutations in histone acetyltransferase gene *KAT6B* in individuals with the Say-Barber-Biesecker variant of Ohdo syndrome. *Am. J. Hum. Genet.* **89**, 675–681 (2011).
- 117 Sirmaci, A., Spiliopoulos, M., Brancati, F., Powell, E., Duman, D., Abrams, A. *et al.* Mutations in *ANKRD11* cause KBG syndrome, characterized by intellectual disability, skeletal malformations, and macrodontia. *Am. J. Hum. Genet.* **89**, 289–294 (2011).
- 118 Hanson, D., Murray, P. G., O’Sullivan, J., Urquhart, J., Daly, S., Bhaskar, S. S. *et al.* Exome sequencing identifies *CCDC8* mutations in 3-M syndrome, suggesting that *CCDC8* contributes in a pathway with *CUL7* and *OBSL1* to control human growth. *Am. J. Hum. Genet.* **89**, 148–153 (2011).
- 119 Abou Jamra, R., Philippe, O., Raas-Rothschild, A., Eck, S. H., Graf, E., Buchert, R. *et al.* Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature. *Am. J. Hum. Genet.* **88**, 788–795 (2011).
- 120 Krawitz, P. M., Schweiger, M. R., Rodelsperger, C., Marcelis, C., Kolsch, U., Meisel, C. *et al.* Identity-by-descent filtering of exome sequence data identifies *PIGV* mutations in hyperphosphatasia mental retardation syndrome. *Nat. Genet.* **42**, 827–829 (2010).
- 121 Huang, L., Jolly, L. A., Willis-Owen, S., Gardner, A., Kumar, R., Douglas, E. *et al.* A noncoding, regulatory mutation implicates *HCFC1* in nonsyndromic intellectual disability. *Am. J. Hum. Genet.* **91**, 694–702 (2012).
- 122 Hussain, M. S., Baig, S. M., Neumann, S., Nurnberg, G., Farooq, M., Ahmad, I. *et al.* A truncating mutation of *CEP135* causes primary microcephaly and disturbed centrosomal function. *Am. J. Hum. Genet.* **90**, 871–878 (2012).
- 123 Carroll, C. J., Isohanni, P., Poyhonen, R., Euro, L., Richter, U., Brillhante, V. *et al.* Whole-exome sequencing identifies a mutation in the mitochondrial ribosome protein *MRPL44* to underlie mitochondrial infantile cardiomyopathy. *J. Med. Genet.* **50**, 151–159 (2013).
- 124 Ghezzi, D., Baruffini, E., Haack, T. B., Invernizzi, F., Melchionda, L., Dallabona, C. *et al.* Mutations of the mitochondrial-tRNA modifier *MTO1* cause hypertrophic cardiomyopathy and lactic acidosis. *Am. J. Hum. Genet.* **90**, 1079–1087 (2012).
- 125 Hedberg, C., Melberg, A., Kuhl, A., Jenne, D. & Oldfors, A. Autosomal dominant myofibrillar myopathy with arrhythmogenic right ventricular cardiomyopathy 7 is caused by a *DES* mutation. *Eur. J. Hum. Genet.* **20**, 984–985 (2012).
- 126 Mayr, J. A., Haack, T. B., Graf, E., Zimmermann, F. A., Wieland, T., Haberberger, B. *et al.* Lack of the mitochondrial protein Acylglycerol Kinase causes Sengers Syndrome. *Am. J. Hum. Genet.* **90**, 314–320 (2012).
- 127 Theis, J. L., Sharpe, K. M., Matsumoto, M. E., Chai, H. S., Nair, A. A., Theis, J. D. *et al.* Homozygosity mapping and exome sequencing reveal *GATAD1* mutation in autosomal recessive dilated cardiomyopathy. *Circ. Cardiovasc. Genet.* **4**, 585–594 (2011).
- 128 Galmiche, L., Serre, V., Beinat, M., Assouline, Z., Lebre, A. S., Chretien, D. *et al.* Exome sequencing identifies *MRPL3* mutation in mitochondrial cardiomyopathy. *Hum. Mutat.* **32**, 1225–1231 (2011).
- 129 Gotz, A., Tynjismaa, H., Euro, L., Ellonen, P., Hyotylainen, T., Ojala, T. *et al.* Exome sequencing identifies mitochondrial alanyl-tRNA synthetase mutations in infantile mitochondrial cardiomyopathy. *Am. J. Hum. Genet.* **88**, 635–642 (2011).
- 130 Tariq, M., Belmont, J. W., Lalani, S., Smolarek, T. & Ware, S. M. *SHROOM3* is a novel candidate for heterotaxy identified by whole exome sequencing. *Genome Biol.* **12**, R91 (2011).
- 131 Takano, K., Liu, D., Tarpey, P., Gallant, E., Lam, A., Witham, S. *et al.* An X-linked channelopathy with cardiomegaly due to a *CLIC2* mutation enhancing ryanodine receptor channel activity. *Hum. Mol. Genet.* **21**, 4497–4507 (2012).
- 132 Zang, Z. J., Cutcutache, I., Poon, S. L., Zhang, S. L., McPherson, J. R., Tao, J. *et al.* Exome sequencing of gastric adenocarcinoma identifies recurrent somatic mutations in cell adhesion and chromatin remodeling genes. *Nat. Genet.* **44**, 570–574 (2012).
- 133 Lindberg, J., Klevebring, D., Liu, W., Neiman, M., Xu, J., Wiklund, P. *et al.* Exome sequencing of prostate cancer supports the hypothesis of independent tumour origins. *Eur. Urol.* **63**, 347–353 (2013).
- 134 Barbieri, C. E., Baca, S. C., Lawrence, M. S., Demichelis, F., Blattner, M., Theurillat, J. P. *et al.* Exome sequencing identifies recurrent *SPOP*, *FOXO1* and *MED12* mutations in prostate cancer. *Nat. Genet.* **44**, 685–689 (2012).
- 135 Snape, K., Ruark, E., Tarpey, P., Renwick, A., Turnbull, C., Seal, S. *et al.* Predisposition gene identification in common cancers by exome sequencing: insights from familial breast cancer. *Breast. Cancer. Res. Treat.* **134**, 429–433 (2012).
- 136 Xiong, D., Li, G., Li, K., Xu, Q., Pan, Z., Ding, F. *et al.* Exome sequencing identifies *MXRA5* as a novel cancer gene frequently mutated in non-small cell lung carcinoma from Chinese patients. *Carcinogenesis* **33**, 1797–1805 (2012).
- 137 Liang, H., Cheung, L. W., Li, J., Ju, Z., Yu, S., Stemke-Hale, K. *et al.* Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome Res.* **22**, 2120–2129 (2012).
- 138 Le Gallo, M., O’Hara, A. J., Rudd, M. L., Urlick, M. E., Hansen, N. F., O’Neil, N. J. *et al.* Exome sequencing of serous endometrial tumors identifies recurrent somatic mutations in chromatin-remodeling and ubiquitin ligase complex genes. *Nat. Genet.* **44**, 1310–1315 (2012).
- 139 Thompson, E. R., Doyle, M. A., Ryland, G. L., Rowley, S. M., Choong, D. Y., Tothill, R. W. *et al.* Exome sequencing identifies rare deleterious mutations in DNA repair genes *FANCC* and *BLM* as potential breast cancer susceptibility alleles. *PLoS Genet.* **8**, e1002894 (2012).
- 140 Wei, X., Walia, V., Lin, J. C., Teer, J. K., Prickett, T. D., Gartner, J. *et al.* Exome sequencing identifies *GRIN2A* as frequently mutated in melanoma. *Nat. Genet.* **43**, 442–446 (2011).
- 141 Ong, C. K., Subimerb, C., Pairojkul, C., Wongkham, S., Cutcutache, I., Yu, W. *et al.* Exome sequencing of liver fluke-associated cholangiocarcinoma. *Nat. Genet.* **44**, 690–693 (2012).
- 142 Agrawal, N., Frederick, M. J., Pickering, C. R., Bettgowda, C., Chang, K., Li, R. J. *et al.* Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in *NOTCH1*. *Science* **333**, 1154–1157 (2011).