

The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population

Nadine Jalkh, Sandra Corbani, Zahraa Haidar, Nadine Hamdan, Elias Farah, Joelle Abou Ghoch, Rouba Ghosn, Nabiha Salem, Ali Fawaz, Claudia Djambas Khayat, et al.

▶ **To cite this version:**

Nadine Jalkh, Sandra Corbani, Zahraa Haidar, Nadine Hamdan, Elias Farah, et al.. The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population. BMC Medical Genomics, 2019, 12 (1), 10.1186/s12920-019-0474-y . hal-02048914

HAL Id: hal-02048914
https://amu.hal.science/hal-02048914

Submitted on 25 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Jalkh et al. BMC Medical Genomics (2019) 12:11
https://doi.org/10.1186/s12920-019-0474-y

BMC Medical Genomics

RESEARCH ARTICLE Open Access

The added value of WES reanalysis in the field of genetic diagnosis: lessons learned from 200 exomes in the Lebanese population

Nadine Jalkh[1], Sandra Corbani[1], Zahraa Haidar[1], Nadine Hamdan[1], Elias Farah[2], Joelle Abou Ghoch[1], Rouba Ghosn[1], Nabiha Salem[1], Ali Fawaz[3], Claudia Djambas Khayat[4,5], Mariam Rajab[6], Chebl Mourani[4,5], Adib Moukarzel[4,5], Simon Rassi[7], Bernard Gerbaka[4,5], Hicham Mansour[8], Malek Baassiri[9], Rawane Dagher[10], David Breich[11], André Mégarbané[12,13], Jean Pierre Desvignes[14], Valérie Delague[14], Cybel Mehawej[1] and Eliane Chouery[1*]

Abstract

Background: The past few decades have witnessed a tremendous development in the field of genetics. The implementation of next generation sequencing (NGS) technologies revolutionized the field of molecular biology and made the genetic information accessible at a large scale. However, connecting a rare genetic variation to a complex phenotype remains challenging. Indeed, identifying the cause of a genetic disease requires a multidisciplinary approach, starting with the establishment of a clear phenotype with a detailed family history and ending, in some cases, with functional assays that are crucial for the validation of the pathogenicity of a mutation.

Methods: Two hundred Lebanese patients, presenting a wide spectrum of genetic disorders (neurodevelopmental, neuromuscular or metabolic disorders, etc.), sporadic or inherited, dominant or recessive, were referred, over the last three and a half years, to the Medical Genetics Unit (UGM) of Saint Joseph University (USJ). In order to identify the genetic basis of these diseases, Whole Exome Sequencing (WES), followed by a targeted analysis, was performed for each case. In order to improve the genetic diagnostic yield, WES data, generated during the first 2 years of this study, were reanalyzed for all patients who were left undiagnosed at the genetic level. Reanalysis was based on updated bioinformatics tools and novel gene discoveries.

Results: Our initial analysis allowed us to identify the specific genetic mutation causing the disease in 49.5% of the cases, in line with other international studies. Repeated WES analysis enabled us to increase the diagnostics yield to 56%.

Conclusion: The present article reports the detailed results of both analysis and pinpoints the contribution of WES data reanalysis to an efficient genetic diagnosis. Lessons learned from WES reanalysis and interpretation are also shared.

Keywords: High throughput sequencing, Exome, NGS, Mutations, Genetic heterogeneity, Genetic diagnostics, Lebanon

* Correspondence: eliane.chouery@usj.edu.lb
[1]Unité de Génétique Médicale, Faculté de Médecine, Campus De l'innovation et du sport, Université Saint-Joseph, rue de Damas, Beirut, Lebanon
Full list of author information is available at the end of the article

© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Background

The identification and characterization of the molecular basis of genetic disorders is crucial for the establishment of a specific diagnosis. This allows the family to benefit from an accurate genetic counseling, the patient to be aware of his disease's prognosis and the physician to implement adequate therapeutic approaches, when possible. The last few years have seen an outstanding improvement in the field of molecular biology [1].

Indeed, the emergence of new high-throughput technologies or Next Generation Sequencing methods (NGS) such as WGS (Whole Genome Sequencing) and WES (Whole Exome Sequencing), has revolutionized the field of genetic diagnosis and made it much more efficient and affordable [2]. While WGS provides a thorough picture of the genome, WES was developed as a practical and cost effective alternative to WGS, targeting only coding regions and canonical splice sites that represent 1–1.5% of the human genome. The rationale behind this approach is that 80–85% of the mutations responsible for Mendelian diseases are located in these regions [3–5].

All these advances are driving NGS from the research field into the clinic. However, data interpretation remains the main challenge especially that thousands of variants of unknown significance are detected in each patient [6, 7].

Here we report the use of WES for genetic diagnostics purposes in 200 Lebanese patients who presented with a wide range of phenotypes, suggesting clinical and genetic heterogeneity. An initial analysis was performed followed, 2 years later, by data reanalysis in order to improve the genetic diagnostic yield. The outcome of WES in the Lebanese population is discussed and lessons learned from this study are shared.

Methods

Clinical samples

From January 2015 to June 2018, 200 patients with genetically heterogeneous disorders were included in our study. They were referred from different Lebanese areas, by private sector and academic physicians. All patients underwent a detailed review of their clinical history and a laboratory evaluation. The most common features of the 200 patients were global developmental delay/intellectual disability, encephalopathy, muscular weakness, failure to thrive and microcephaly.

Approval to conduct the study was obtained from the Ethics Committee of Saint Joseph University, Beirut, Lebanon. All patients, parents or legal guardians signed an informed consent for participation, sample collection and data publication. Peripheral blood was then collected from each individual enrolled in this study and DNA was extracted using the salting out method [8].

WES analysis

Exon capture and sequencing: The exome was captured using the SureSelect Human All Exons, reagents (Agilent Inc.® Santa Clara, CA) according to the manufacturer's standard protocol. The concentration of each library was determined using Agilent's QPCR NGS Library Quantification Kit (G4880A). Samples were pooled prior to sequencing with a final concentration of each sample equal to 10 nM. Sequencing was performed on the Illumina HiSeq2000 platform using TruSeq v3 chemistry.

Mapping and alignment: Reads files (FASTQ) were generated from the sequencing platform via the manufacturer's proprietary software. Reads were aligned to the hg19/b37 reference genome using the Burrows-Wheeler Aligner (BWA) package v0.6.1 [9]. Local realignment of the mapped reads around potential insertion/deletion (Indel) sites was carried out with the Genome Analysis Tool Kit (GATK) v1.6 [10]. Duplicate reads were marked using Picard v1.62. Additional BAM file manipulations were performed with Samtools 0.1.18 [11]. Base quality (Phred scale) scores were recalibrated using GATK's covariance recalibration. SNP and Indel variants called using the GATK Unified Genotyper for each sample [12]. Variants were called using high stringency settings and annotated with VarAFT software 1.61 [13] containing information from dbSNP147 and ExAC (http://exac.broadinstitute.org/). For all patients who were left undiagnosed, reanalysis of vcf files was carried out using VarAFT2.131 and polymorphisms were further filtered using our in-house local database containing more than 300 exomes (Fig. 1).

In terms of functional annotation, we included only protein-altering variants, including truncating variants (stop gain/loss, start loss, or frameshift), missense variants, canonical splice-site variants, inframe indels affecting protein-coding regions, and variants within the intron–exon boundary (ten bases flanking the exonic boundaries). We focused on genotypes absent in our local control data sets. We systematically considered four different genetic models, using stratified European and African Americans in the Exon Variant Server (EVS) for minor allele frequency estimations: (a) germ-line de novo mutations, also absent in the local control populations; (b) recessive homozygous genotypes, which were heterozygous in both parents, never homozygous in controls, with a control allele frequency < 1%; (c) hemizygous X chromosome variants inherited from an unaffected heterozygous mother, with a control allele frequency < 1% and never observed in male controls or homozygous in female controls; and (d) compound heterozygous genotypes in the patient (one variant inherited from each heterozygous parent, with the two variants occurring at different genomic positions within the same gene), for which neither variant was ever homozygous in

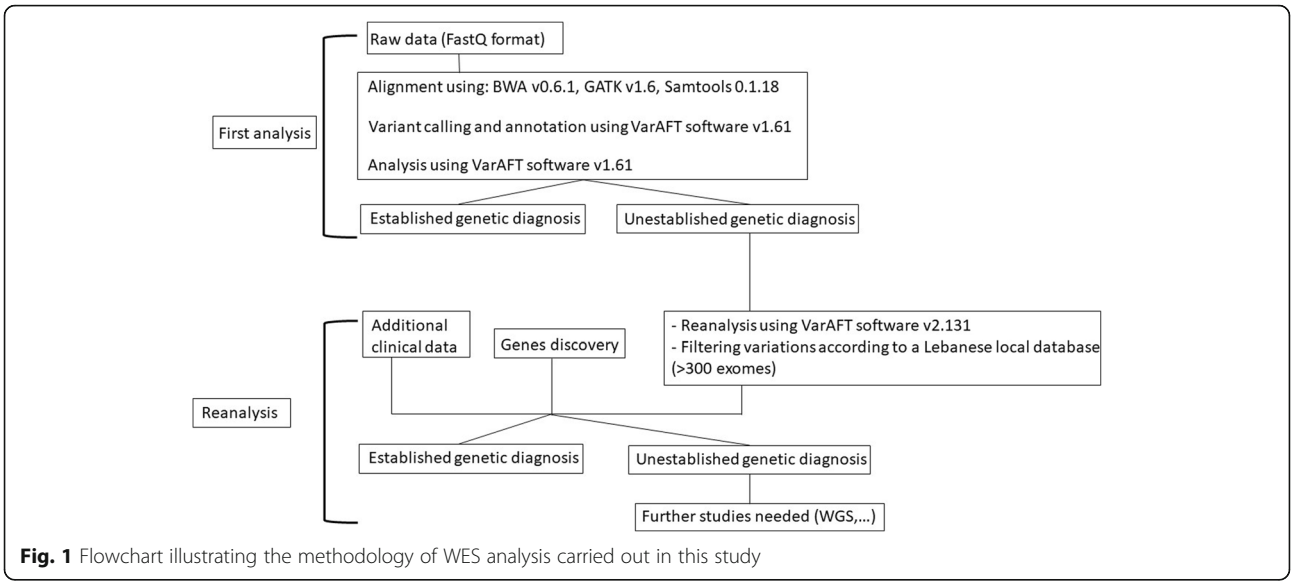


Fig. 1 Flowchart illustrating the methodology of WES analysis carried out in this study

controls, and each had a control allele frequency < 1%. For the compound heterozygous genotypes, we further required study of phasing of the two variants. Genotypes meeting these criteria were referred to as "candidate genotypes," with the genes harboring candidate genotypes referred to as "candidate genes" [14].

Variants confirmation and segregation studies by sanger sequencing

Primers were designed using Primer 3 (http://frodo.wi.mit.edu) and OLIGOS v.9.3, and checked for specificity using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi). DNA sequences were obtained from UCSC and Genbank databases. Standard PCR reactions were performed using Taq DNA polymerase (Invitrogen Life Technologies, Carlsbad, CA, USA) and both strands of the resultant products were sequenced using the BigDye® Terminator v1.1 Cycle Sequencing Kit (ThermoFisher Scientific, Waltham, MA, USA) under standard conditions. The labeled products were subjected to electrophoresis on an ABI 3500 Genetic Analyzer sequencing system (ThermoFisher Scientific, Waltham, MA, USA). Electropherograms were analyzed using Sequence Analysis Software version 5.4 (ThermoFisher Scientific, Waltham, MA, USA) and compared to reference sequences using ChromasPro version 1.7.7 (Technelysium, Queensland, Australia). Nucleotide numbering reflects cDNA numbering with + 1 corresponding to the A of the ATG translation initiation codon in the reference sequences.

Results

Over a 42-month period, 200 cases with heterogeneous genetic disorders were enrolled in this study. Among these patients, around 80% (159 samples) were younger than 18 years of age. From the latter, 150 cases were children, including the case of a fetus whose specimen was collected after a terminated pregnancy.

Patients were referred by physicians for presenting with (Table 1): neurodevelopmental disorders (in 39.5% of the cases), neuromuscular disorders (10%), metabolic and mitochondrial disorders (9.5%), renal disorders (5%), hearing disorders (3.5%), isolated epilepsy (3%), bone diseases (2%), leukodystrophy (2%), visual disorders (1%) and other rare diseases referred as "others" in all tables (24.5%). WES was performed on these patients for diagnosis purposes.

Our first analysis yielded a success rate of 49.5%. In other words, genetic diagnosis was established in 99 patients, after the first analysis (Table 2). In order to improve the diagnosis yield, reanalysis of WES data, generated during the first 2 years of this study, was carried out. Reanalysis was performed using new bioinformatics algorithms and was based on the newly established local database, new genes discoveries and additional clinical information for each of the undiagnosed patients. The second analysis allowed the identification of the pathogenic mutations in 13 additional cases (Table 3) corresponding to 6.5% increase in the genetic diagnosis rate, thus leading to an overall success rate of 56%. The 112 patients with established genetic diagnosis included 54 patients with autosomal dominant diseases, 49 with autosomal recessive diseases, and 9 with X-linked diseases. Different types of mutations were detected including small frameshift, nonsense, splice sites, and missense mutations.

There was a trend towards an association between the rate of positive diagnosis and the disease group. Indeed, the highest diagnosis success rate corresponds to the group of hearing and visual disorders (100%), followed by neuromuscular disorders (85%), metabolic and

Table 1 Distribution of cases between disease groups with the corresponding success rate

Pathologies	Number of patients presenting the pathology	Number of patients with established molecular diagnosis	Percentage of patients with established molecular diagnosis (%)
Neurodevelopmental disorders	79	24	30.4
Neuromuscular disorders	20	17	85
Metabolic and mitochondrial disorders	19	16	84.2
Renal disorders	10	6	60
Hearing disorders	7	7	100
Epilepsy	6	4	66.7
Bone diseases	4	3	75
Leukodystrophy	4	3	75
Visual disorders	2	2	100
Others	49	30	61.2
Total	200	112	56

mitochondrial disorders (84.2%), bone diseases and leukodystrophy (75%), epilepsy (66.7%), others (61.2%), renal disorders (60%) and neurodevelopmental disorders (30.4%) (Table 1).

Of the 136 identified mutated alleles, 66.9% were novel variants at the time of diagnosis.

Discussion

WES is a relevant, efficient and cost effective genetic diagnostic tool that allows the detection of a wide range of genetic variants including point mutations (missense and nonsense mutations), small deletions and insertions in addition to canonical splicing mutations [15]. However, linking genetic variations to diseases still represents a big challenge in many cases.

Here we report the first study on WES outcome in the Lebanese population. Exome sequencing performed in 200 cases included in our study first yielded a 49.5% overall success rate, which is concordant with other international studies [16]. This rate was variable depending on the disease group. It barely reached 30.4% in some cases such as in neurodevelopmental diseases including intellectual disabilities (ID) and autism spectrum disorders (ASD) but was extended to 100% in the case of some diseases with well-established molecular mechanisms, such as visual and hearing disorders.

In order to improve the diagnostic yield, WES raw data that was generated over the first 2 years were reanalyzed. Repeated WES analysis enabled us to establish the genetic diagnosis, for 13 additional patients who were left undiagnosed, corresponding to around 7% increase in the diagnosis rate. This positive outcome is due to many factors, among which: the development of new annotation and filtering tools (including new bioinformatics variant effect prediction logarithms), the continuous revelation of new genes in addition to the implementation at our research unit of a Lebanese WES database that allowed us to filter out all "private" Lebanese polymorphisms. Apart from these factors, the 2 years of experience in WES analysis were an asset that helped us improve our performance in data interpretation. Here we report some of the lessons learned from our study:

Two monogenic diseases can masquerade as a single entity: the first molecular investigation of both parents (P5 and P6) of two deceased patients presenting with heart failure and achromatopsia did not lead to any success. However, reanalysis helped to identify, in each parent, two heterozygous mutations in *MYL3* and *PDE6C*, responsible for two different conditions: heart defect (OMIM #608751) and achromatopsia (OMIM #613093), respectively. The observation, in a single patient, of multiple "hits" leading to a unique clinical presentation is nowadays becoming more and more acceptable.

Molecular geneticists need to be aware of filtering out homozygous mutations classified, by bioinformatics logarithms, as "heterozygous uncertain" due to poor read depth: A homozygous mutation in *PET100* in the patient P13, classified as heterozygous uncertain by our logarithm was missed in our first WES analysis. Reanalysis allowed the detection of this candidate mutation and Sanger sequencing enabled its confirmation.

Patients from consanguineous families can still present with compound heterozygous mutations: Patients P10 and P95, issued form consanguineous Lebanese families, were for instance shown to carry respectively two compound heterozygous mutations in *MYO15A* responsible for an autosomal recessive type of deafness (OMIM #600316) and in *SPG11* responsible for spastic paraplegia type 11 (OMIM #604360).

De-novo mutations linked to autosomal dominant conditions can occur in patients from consanguineous families: A de novo mutation in *GJB3* (p.S199R) involved in hearing loss was detected in patient P66 who is issued from a consanguineous Lebanese family.

Table 2 Clinical and genetic characteristics of families with established genetic diagnosis following the first analysis

Patient code	Actual age in years	Affected	Consanguinity	Coefficient of consanguinity	Mode of inheritance according to the pedigree	Pathology	Clinical signs	Gene	Mutation	Zygosity
P1	11-16	+	-		AD	Neurovascular disorder	At 16 years old, abnormal gait, mild cognitive deficiency, intellectual deficiency, psychomotor delay, cerebellar ataxia, strabismus ataxia, cerebellar atrophy, seizures	*SGSH* (NM_000199)	p.L110= (c.330C>G)	Htz
P2	11-16	+	+	1/16	AR	Neurovascular disorder	Neurodevelopmental delay, psychomotor delay, cerebellar ataxia, strabismus ataxia, cerebellar atrophy, seizures	*LRP4* (NM_170707)	p.T425A (c.1087A>G)	Homo
P3/P4	11-16	+	+	1/16	AR	Neurovascular disorder	Neurodevelopmental delay, psychomotor delay, cerebellar ataxia, cerebellar atrophy, seizures	*CTTNBP2* (NM_052875)	p.R1797 (c.5392C>T)	Homo
P5/P6	11-16	+	+	1/16	AR	Other: heart disease and achromotopsia	Two kids deceased from heart failure - One of them presented also achromotopsia	*SCN5A* (NM_198056) *ATX2* (NM_002974) *PDYN2* (NM_006820)	p.T1620M (c.4859C>T) p.L124P (c.371T>C) p.C76R (c.226T>C)	Htz Htz Htz
P7	11-16	+	-		AR	Other: Ichtiosis	Ichtiosis, oligomenorrhea at 16 y/o, nystagmus, photophobia	*TGS* (NM_000372)	c.[823C>T];[=] (855G>A); p.[T12A];[=](T4=)	Compound Htz
P8	11-16	+	-		AR	Neuromuscular disorder	Muscular dystrophy	*CAPN3* (NM_000070)	IVS7-1_4delAC/wt	Compound Htz
P9	11-16	+	-		NR, de novo	Hearing disorder	Deafness in the age of 2 years	*SLC4A7*) (NM_003615) *GJB2* (NM_004004)	p.A267Val (c.1784A>G) p.R127H (c.380G>A)	Htz/Homo
P10	11-16	+	+	1/16	AR	Hearing disorder	Mild to severe deafness	*MYO15A* (NM_016239)	p.A2492Val (c.7474G>T) p.C1726= (c.5178C>T)	Compound Htz
P11	11-16	+	+	1/128	AR, X-linked	Neurodevelopmental disorder	Neurodevelopmental delay, psychomotor delay, mental retardation	*ANPS* (NM_000005)	p.E875Q (c.2617G>C)	Homozygote
P12	6-10	+	-		de novo	Neurodevelopmental disorder	Neurodevelopmental delay, dysmorphic, microcephaly, molecular scopio, strabismus, hypotonia, cerebellar atrophy, artificial anus, kidney malformation	*CSNK1* (NM_017788)	p.D350N (c.1051G>A)	Htz
P13	6-10	+	-		AR	Metabolic disorder	Bilateral severe deafness, mitral flux, convulsive crisis, recurrent vomiting by acid aspiration, abnormal hairs and bones	*PETH* (NM_001171535)	p.M11=/C>G	Htz
P14	6-10	+	- (same village)		AR	Bone disease	Acrodysostosis, intellectual disability, neonatal respiratory distress, short stature, short + thin tubular skin, neurocephaly, convulsive crisis, hypotelorism, epicanthic fold, broad base to nose, small and thin lips, micrognathia, high arched palate, universal tooth, mid face hypoplasia, open mouth appearance, small hands, brachydactyly	*PDE4D* (NM_001904811)	p.T23A (c.73A>G)	Htz
P15	21-30	+	-	1/16	AR	Hearing disorder	Deafness	*CDH23* (NM_022124) *SLC26A4* (NM_000441)	p.[D9N];[=]*(T9D=) p.K679= (c.2037A>G)	Homo Htz
P16	6-10	+	-		AD, de novo	Neurodevelopmental disorder	Neurodevelopmental delay, mental delay, dysmorphy, plagiocephaly, short neck, hypotonia	*SHOT9* (NM_004992.3)	p.L3868del (c.11597_11600del)	Htz
P17	21-30	+	- (same village)		AR	Neurodevelopmental disorder	Microcephaly, psychomotor delay, hypotypia from the white matter, bilateral enlargement of the lateral ventricules, atrophy of the cerebellum, delayed myelination, cerebellar vermis is in the frontal and subcortical regions of the lobe.	*SNX14* (NM_020869)	p.D19E (c.55A>C) p.M672= (c.2016delC)	Homo Homo
P18	6-10	+	-		AD, de novo	Hearing disorder	Sensorineural delay, hypoacusia, cerebellar ataxia, psychomotor delay, cognitive delay	*CDH23* (NM_022124) *TJP2* (NM_004817) *ADGV1* (NM_001079)	p.E1644K (c.4930G>A) p.P1041T (c.3122C>T) c.4841T0wt/wt and (c.1348-1367)	Htz Htz Htz
P19	21-30	+	-		AR	Metabolic disease to be deceased child	Multi- organ-like disease from deceased child	*ACADM* (NM_ 000016)	p.K329_G.=(c.986C>T)	Htz
P20	6-10	+	+	1/16	AR, AD	Metabolic disorder	Psychomotor delay, bilateral optic atrophy, choreoathetosis/Dystonia with cerebellar ataxia, generalized sensory polyneuropathy of spinocerebellar degeneration, muscle biopsy: mitochondrial dysfunction	*CLCN1* (NM_000083)	p.T859M (c.2786delT)	Htz
P21	6-10	+	+	1/16	AR, AD	Renal dysfunction	Renal failure	*PKD1* (NM_ 001009944)	p.[1830fs];[=](.5914C>A)	Htz
P22	6-10	+	+	1/8	AR	Leukoencephalopathy	Leukoencephalopathy with vanishing white matter, microcephaly, thrombocytopenia, plagiocephaly, sub-ependymic thick hair, hypertonia, increased appetite, thrombocytopenia, epilepsy, hepatomegaly	*RNASEH2B* (NM_024570)	p.G40R (c.119A>G) p.A177 (c.529G>A)	Homo Homo
P23/P24	21-30	+	-	With the parents 1/16	Other - Epidermal bullosa in their two deceased children	Decreased children Epidermal bullosa around finger nails at the knee leg around joints, - Cessation of blisters in response to minor physical injury	*LAMA3* (NM_000227)	c.1789-3delTTC	Htz
P25	6-10	+	-		de novo	Neurovascular disorder	Psychomotor delay, forefoot varus	*ATN1* (NM_ 001940) *ATN1* (NM_001007)	c.2560del (c.[=]2560del)(=) p.R1100Lfs*3 (c.3300C>T)	Htz Htz
P26	6-10	+	-		AD, de novo	Neurodevelopmental disorder	Psychomotor delay, intellectual disability, hypotonia, cerebellar ataxia	*KCNC1* (NM_004977)	p.A335A (c.4654delG)	Htz
P27	6-10	+	-		NR, NR, de novo, X-linked	Epilepsy	Epilepsy	*PIGN* (NM_176787)	p.E952Q (c.2945C>A)	Homo
P28	11-16	+	Parents from the same village		AD, de novo, AR	Neurodevelopmental disorder	Bone marrow signal abnormalities, bone pain, psychomotor disorder, spreading diffu ulbus, hypotelos	*PGAP1* (NM_ 024644)	p.R218C (c.652C>T)	Htz
P29	6-10	+	-	1/128	AR	Epilepsy	Pyridoxine dependent epilepsy, psychomotor delay	*ALDH7A1* (NM_001182)	p.V464A (c.2393T>A)	Htz
P30	11-16	+	+		AD	Renal dysfunction	Polycystic kidney disease	*PKHD1* (NM_ 138694)	(c.79+3T>G)	Htz
P31	11-16	+	+	1/16	AR	Neurovascular disorder	Psychomotor delay, mental retardation, cerebellar ataxia, paravertical diplegia	*COL6A3* (NM_004369)	p.H871Y (c.2611C>A)	Htz
P32	deceased	+	Parents from the same village		AR	Metabolic disease - Niemann- Pick disease	Hepatosplenomegaly, recurrent respiratory infections, abnormal white matter signal, psychomotor delay, cerebellar disorder	*SMPD1* (NM_000543)	p.H419R (c.1256A>H)	Homo
P33	6-10	+	-		NR, de novo, Slinked	Lymphopathy	Floppy baby, hypotonia, microcyst, white matter injury	*PGN* (NM_000311)	p.G2R (c.5G>C)	Homozygote
P34	6-10	+	-		AR	Renal dysfunction	Renal dysfunction	*PKHD1* (NM_ 138694.3)	p.[4157=];(.4615C>A)	Htz
P35	6-10	+	-		AD, de novo	Neurodevelopmental disorder	Psychomotor delay, dysmorphic protruding eyes, triangular face, low set ears	*SNX9* (NM_ 001165)	p.R623Q (c.1862G>A)	Htz
P36	6-10	+	-		AD, de novo	Neurodevelopmental disorder	Psychomotor delay, anemia, nystagmus, macrocytic anemia	*SLC12A1* (NM_ 000338.2)	(c. 3279-2A>C)	Htz
P37	6-10	+	+	1/16	AR	Hearing disorder	Deafness	*TECTA* (NM_005422) *CDH23* (NM_022124)	p.T1486=(c.4458C>T) p.E1644K (c.4930G>A)	Homo Htz
P38	11-16	+	Parents from the same village		AR	Neurodevelopmental disorder	Cardiac malformation, epilepsy, microcephaly, hypotonia	*SLC13A5* (NM_ 001times)	p.G399=(c.1197C>G)	Htz
P39	11-16	+	+		AD	Neurodevelopmental disorder	Intellectual disability, hypotonia, cerebellar ataxia	*ATP1A3* (NM_ 152296)	p.A452= (c.1356C>T)	Htz
P40	6-10	+	-		AD, Slinked	Neurodevelopmental disorder	Psychomotor delay, cerebellar delay, hypotonicity, intellectual disability	*ANPS* (NM_000005)	p.R139Q (c.416G>C) p.R139T (c.416G>T)	Homozygote
P41	11-16	+	+	1/16	AR	Neuromuscular disorder	Muscular weakness, Gowers sign, psychomotor delay, cerebellar ataxia, difficulties	*TUBA3E*(NM_ 020675)	p.T307=(c.909C>A)	Htz
P42	6-10	+	+	1/16	AR	Other: Usher syndrome	Visual and hearing disorders, choreoathetosis	*ANAG1* (NM_ 022124)	p.G1041fs (c.3641dupA)	Homo
P43	11-16	+	+		AD, de novo, X-Mosh	Neurodevelopmental disorder	Psychomotor delay, speaking delay, mild encephalopathy, epilepsy, hypotonia, cerebellar ataxia	*COL1A1* (NM_ 000083)	p.H85N (c.265T>A)	Htz
P44	11-16	+	Parents from the same village		NR, AD, de novo	Other - Microphthalmia	Visual acuity, in both eyes, cataract	*RPL13* (NM_ 176875)	p.G102E (c.304G>C)	Htz
P45	6-10	+	+		AD, de novo, Slinked	Neurodevelopmental disorder	Neurodevelopmental delay, psychomotor delay, hypotonia, cerebellar ataxia	*HUS1* (NM_ 004507)	p.R26C (c.76C>T)	Homozygote
P46	11-16	+	+		AD, de novo, Slinked	Neurodevelopmental disorder	Hypotonicity, short neck, growth delay, psychomotor delay, mild dysmorphic feature	*HUS1* (NM_ 025205)	p.V30L (c.233C>T)	Htz
P47	16-20	+	+		AR	Neurovascular disorder	CTG AWR - Limb girdle muscular dystrophy	*ANPS* (NM_ 005004)	p.L412F (c.1235C>C)	Homo
P48	6-10	+	+	1/8	AR	Neurovascular disorder	Neurodevelopmental disorder	*SLC12A2*(NM_ 001346)	p.T1236= (c.3708C>T)	Homo
P49	6-10	+	+	X-Mosh, de novo	Neurovascular disorder	Psychomotor delay, genius mental retard, weakness	*GJB1* (NM_000166)	p.E47K (c.139G>A)	Homozygote
P50	6-10	+	+	de novo	Bone disease	Osteoporosis	*COL1A1* (NM_ 001454)	p.G1106_G1107ins(.)	Htz
P51	6-10	+	+	de novo	Neurodevelopmental disorder	Psychomotor delay	*ARG1D1* (NM_ 020805)	p.D149=(c.447C>T)	Htz
P52	6-10	+	+	1/16	AR	Metabolic disease	Mean cardiac joint disease	*RUNX1T1* (NM_ 004349)	p.D371G (c.1109A>G)	Htz
P54	11-16	+	+		AR	Other: Immunological chronic	Autoimmune polyendocrinopathy syndrome	*AIRE* (NM_ 007358)	p.L86Q (c.256T>A)	Homo
P55	6-10	+	+		AR, D Slinked	Neurodevelopmental disorder	Abdominal malformations, reflux (Cardiac malformation), agenesis of one of the renal cortex, bladder hypertrophy	*SMN1* (NM_ 000344)	p.Q87R (c.259C>G) p.N370S (c.1109A>G)	Homo Htz
P56	16-20	+	+		AR	Renal disease	Diabetes	*ABCA4* (NM_ 000350)	p.G1228A (c.3683G>A)	Homo
P57	11-16	+	+		de novo	Other: ADHD	Attention deficit hyperactivity disorder	*DRD4* (NM_ 000797)	p.G124Afs (c.370dupA)	Htz
P58 Mother	11-16	+	+	With her partner 1/16	AR	Metabolic diseases	Her sons present intellectual deficiency, hypotonia, convulsive seizures/epilepsy, deficiency. One of her sons epilepsy at the age of 4 months	*GLDC* (NM_ 000170)	p.G19C (c.52G>T)	Htz
P59	6-10	+	+		D Slinked	Metabolic disease, epilepsy	Epilepsy, muscle biopsy showing mitochondrial dysfunction	*ALG13* (NM_ 001099922)	p.G1110T (c.3230A>T)	Htz
P60	11-16	+	+		de novo	Renal dysfunction	Nephritic syndrome, developmental delay	*NPHS2* (NM_ 014625)	p.V180M (c.539G>A)	Htz
P61	11-16	+	+	1/16	AR, AD	Other: recurrent fever	Microcytic anemia, Fever, thrombotic thrombocytopenic deficiency, deficiencies disease, Spina bifida	*IL36RA* (NM_ 003558)	p.E98K (c.1726=A)	Htz
P62	6-10	+	+	1/16	AR	Metabolic disease / Other: Leigh syndrome	Leigh syndrome, bilateral optic atrophy, low severe myotonia and anemia	*MECR* (NM_ 016011)	p.E153Q (c.517A>G)	Htz
P63	6-10	+	+		AD, de novo	Other: craniosynostosis	Craniosynostosis, frontal bossing, late tooth and anterior wide spacing, hypertelorism, arched palate, retrusive face, long face	*TCF12* (NM_ 002327)	p.Q622* (c.1864C>T)	Htz
P64	6-10	+	+		de novo, ? Slinked	Neurodevelopmental disorder	Hypotonia, microgyria/dysplasia, short stature, developmental delay, cerebellar ataxia, arched palate, retrusive face	*CPS1* (NM_ 017980)	p.Q1050E (c.1759C>G)	Htz
P65	6-10	+	+		AR	Visual disorder	Cone and dystrophy, macular disorders	*ABCA4* (NM_ 000350)	p.L2278P (c.6832T>C) p.V4600I (c.13807G>A)	Compound Htz
P66	6-10	+	+	1/16	AR	Hearing disorder	Mild/Moderate sensorineural deafness	*GJB2* (NM_ 003320)	p.G59R (c.175G>A)	Htz
P67	11-16	+	+	1/16	AR	Other: Ichtiosis vulgaris	Ichtiosis vulgaris	*FLG2* (NM_ 004613)	Ichtiosis vulgaris	Htz
P68	11-16	+	+		AR	Neuromuscular disorder	Unsteady gait, muscle, cerebellar syndrome	*KAT6* (NM_ 012330.3)	p.L1147= (c.3441C>T)	Homo
P69	6-10	+	+			Metabolic disorder	Congenital and speech delay with bone scoliosis, convulsed hypotonia	*SMPD1* (NM_ 001942251)	p.S189=(c.567C>T)	Homozygote
P70	6-10	+	+		AD	Metabolic disease	Glucosylation disorder, psychomotor/Dystonia, muscle rigidity, poor facial, failure to thrive, hypotonia, cerebellar cerebral atrophy	*PMM2* (NM_ 000303)	p.V129M (c.385G>A) p.R141H (c.422G>A)	Compound Htz
P71	6-10	+	+		de novo	Epilepsy	Epilepsy, West syndr	*SCN8A* (NM_ 001077)	p.V429M (c.1285G>A)	Htz
P72	6-10	+	+		AD	Neuromuscular disorder	Cerebellar ataxia, psychomotor delay, hypotonia, cerebellar ataxia, cerebellar atrophy	*CACNA1A* (NM_ 001127357)	p.E1864Q (c.5590G>C)	Htz
P73	6-10	+	+	1/64	AR	Other: Glaucoma/Keratoconus	Glaucoma/Keratoconus	*ZNF469* (NM_ 001079)	p.Q4536=(c.13608C>T)	Htz
P74	6-10	+	+	1/16	AR	Neurodevelopmental disorder	Bone cyst/ribs hypoplasia / severe dysmotile ataxia	*TBX6S* (NM_ 052836)	p.V445I (c.1333G>A)	Homo
P75	6-10	+	+		AR	Other: Periodic fever	? Pyrexia, fever, edema, urticaria, ataxia every 2 to 3 weeks	*MEFV* (NM_ 000243)	p.V726A (c.2177T>C)	Homo
P76	6-10	+	+		X-linked	Other: Hemophagocyte lymphohistiocytosis	Hemophagocyte	*XIAP* NM_001167	p.N164 (c.102_102delA)	Homozygote
P77	6-10	+	+		de novo	Other: Cerebral ataxia	Cerebral ataxia, severe motor dysfunction, muscle tone, epilepsy	*CACNA1A*(NM_001127357)	p.R1864Q (c.5591G>A)	Htz
P78	6-10	+	+		de novo	Metabolic disease / Epilepsy / Encephalopathy	Metabolic ataxia, seizure (Epilepsy)	*SCN8A* (NM_ 014191)	p.R1872Q (c.5615G>A)	Htz
P79	6-10	+	+	1/32	AR	Other - Hemophagocytic lymphohistiocytosis / Deficiency	Severe hemophagocytosis before the abdominal crisis	*PRF1* (NM_ 001083116)	p.T61M (c.182C>T)	Homo
P80	11-16	+	+		AR	Metabolic disorder	Hypomyelinopathy, severe ataxia	*GJA1* (NM_ 000165.5)	p.S83C (c.248C>T)	Htz
P82	6-10	+	+	1/16	AR, AD	Other: Neurofibromatosis	Neurofibromatosis, thalassemia	*NF1* (NM_ 000267)	(c.2430G>C);(=2754dupA)	Htz
P83	6-10	+	+	1/16	AR	Other: Thrombocythemia	Bleeding, low blood platelet count	*ITGA2B* (NM_ 000419) *ITGA2B* (NM_ 000419)	p.T162=(c.1134C>T) p.N841Ifs (c.2520C>A)	Htz Htz
P84	6-10	+	+	1/16	AR	Other: Cholestasis	Hypotonia, hepatomegaly, cholestasis, cirrhosis, abnormal	*ITGA2B* (NM_ 000419)	p.L169fs (c.499delA)	Homo
P85	Deceased	+	+		AR	Other: Pyruvate/pyruvatemia	Pyruvatemia, cirrhosis, cholestasis, abnormal	*ALDH3A* (NM_000736)	(c.1421G>A)	Htz
P86	6-10	+	+	1/2	AR	Metabolic disorder	Ceramine deficiency, speaking delay, elevated pyruvic acid and lactate	*TNXB* (NM_ 019980)	p.E6O3* (c.1808A>G)	Homo
P87	6-10	+	+		AD, XLD	Metabolic disorder	Favism, hemolytic anemia	*G6PD* (NM_ 001042351)	p.S188F (c.563C>T)	Htz
P88	11-16	+	+		AD	Other: CADASIL	Cerebral arteriopathy with subcortical infarcts and leukoencephalopathy	*NOTCH3* (NM_ 000435)	p.C455Y (c.1365T>C)	Htz
P89	6-10	+	+	1/16	AR, AD, de novo	Other: Achondroplasia	Achondroplasia, bone fusion, abnormal liver function tests, increased serum bile acids	*FGFR3* (NM_ 000142)	p.N858=(c.2574C>G)	Htz
P90	6-10	+	+	1/16	AR	Other: Cholestasis	Progressive familial intrahepatic cholestasis	*ATP8B1* (NM_ 005603)	p.R352Q (c.1055G>A)	Htz
P91	11-16	+	(same village)		AR, AD	Other: Cholestasis	Progressive familial intrahepatic cholestasis	*ATP8B1* (NM_ 005603)	p.H285Y (c.853C>T)	Htz
P92	6-10	+	+		AR, AD	Other: Periodic fever	Abdominal pain, Fever episodes	*MEFV* (NM_ 000243)	p.V726A (c.2177T>C)	Htz
P93	6-10	+	+	1/8	AR	Renal dysfunction	Hypomagnesia	*CLDN16* (NM_ 006580)	p.T233= (c.699C>T)	Htz
P94	11-16	+	+		AD, Slinked	Neurodevelopmental disorder	Psychomotor delay, speaking, intellectual disability	*SMN2*(NM_ 017411)	p.L196fs (c.586_587insA)	Homozygote
P95	11-16	+	+	1/16	AR	Neurovascular disorder	Psychomotor delay, speaking, intellectual disability	*SLC12A2* (NM_ 001346)	p.T1236= (c.3708C>T)	Compound Htz
P96	6-10	+	+		AD	Other: Periodic fever	Recurrent fever, elevated CPK, LDH, CRP and fibrinogen	*TNFRSF1A* (NM_ 001065)	p.C70Y (c.211T>A)	Htz
P97	6-10	+	+		AD, de novo	Metabolic disorder	Psychomotor delay, delayed fine, late, late, late	*SRD5A3* (NM_ 024592) *SRD5A3* (NM_ 024592)	p.A132V (c.395C>T) p.T240A (c.718A>G)	Htz Htz
P98	11-16	+	+		AD, de novo	Other: Noonan syndrome	Psychomotor delay, delayed speaking, intellectual disability, hypotonia, cerebellar ataxia, cryptorchidism, not prominently stained eyes	*PTPN11* (NM_ 002834)	p.M504V (c.1510A>G)	Htz
P99	6-10	+	+		AD, de novo	Mitochondrial disorder	Psychomotor delay, strabismus, ophthalmoplegia, facies retinal	*SCAP1* (NM_ 003172)	p.F296L (c.876delT)	Htz

AD Autosomal dominant, *AR* Autosomal recessive, *Htz* Heterozygous, *Homoz* Homozygous

Variable Expressivity and incomplete penetrance of some diseases might mislead the choice of the filtering strategy: Patient P94 who was expected to carry a de novo mutation, was found to have a hemizygous mutation (p.E1204G) in *SHROOM4*, inherited from his mother who shares with him milder clinical signs.

Pathogenic mutations with a high frequency in a population exist! Filtering out common variants can be risky in the case of genetic disorders that are frequent in a population: In patient P87 presenting with G6PD deficiency, a disease frequently seen in the Lebanese population, WES reanalysis led to the identification of a pathogenic mutation in *G6PD* (p.S188F) that is present in high frequency (greater than 1%) in several arab populations [17].

WES has technical limitations: Large deletions encompassing candidate genes were detected by CGH arrays in two patients (not included in the study), who were left undiagnosed by WES.

WES coverage data might, in some cases, enable the detection of gene deletion: NPHP1 gene coverage was for instance equal to zero in patient P67 who presented with Joubert syndrome. This gene was fully covered in other patients run on the same sequencing chip. Owing to the potential involvement of the gene in the disease, CGH array was performed and a homozygous deletion encompassing the entire *NPHP1* sequence was confirmed, thus explaining the clinical presentation.

Last but not least,

Communication between clinicians and geneticists is the ultimate key for genetic diagnostics success: In our first WES analysis, genetic diagnostic was not established for the patient P106 who was referred to our center as presenting with Bartter syndrome. While following up on all undiagnosed cases during our WES reanalysis, a pathogenic mutation in *SLC9A2* (p.F82 L) was highlighted and communicated to the physician who re-evaluated the patient and modified the initial diagnosis to congenital chloride diarrhea, known to result from mutations in *SLC9A2* [18].

On the other hand, failure of genetic diagnostics in the remaining patients can be due to several factors: i) the clinical and genetic heterogeneity of some entities: More than 1000 genes are for instance linked to ASD, thus making genetic diagnostics very challenging (https://gene.sfari.org/database/human-gene/); ii) the detection in isolated cases of private mutations whose pathogenicity is hard to be demonstrated in addition to iii) purely technical limitations associated with WES: Some bioinformatics pipelines are for example unable to detect CNVs (Copy Number Variation), rearrangements or triplet expansions. Homopolymers can for instance be miscalled, thus generating false positive variants. Furthermore, lack of coverage of coding regions

Table 3 Clinical and genetic characteristics of the 13 additional cases with established genetic diagnosis after reanalysis

Patient code	Actuel age in years	Affected	Consanguinity	Coefficient of consanguinity	Mode of inheritance according to the pedigree	Pathology	Clinical signs	Gene	Mutation	Zygosity
P100	11-20	+	+	1/32	AR	Neurodevelopmental disorder	Mild psychomotor delay, vitiligo, essential tremor	*IVD* (NM_002225)	p.S281R (c.841A>C)	Homo
								FASTKD2 (NM_014929)	p.R358* (c.1072C>T)	Homo
P101	0-10	+	-		AD, de novo, Xlinked	Neurodevelopmental disorder	Psychomotor delay, developmental delay, microcephaly, retrognathism, testicular ectopia, dysmorphia	*ALDH18A1* (NM_002860.3)	c.2207-3C>T	Htz
								ITPR1 (NM_002222.5)	(c.5164+4C>T	Htz
P102	51-60	+	+		AD	Neuromuscular disorder	Ataxic walk, dystonia, tremor, dysgraphia	*ATXN3* (NM_004993)	p.G306fs (c.915_916ins(CAG) $_{13}$)	Htz
P103	11-20	+	+	1/32	AR	Neuromuscular disorder	Bilateral piramydal syndrome, muscular weakness	*TGM6* (NM_198994)	p.R342W (c.1024C>T)	Htz
P104	0-10	+	-		AR	Neurodevelopmental disorder	Psychomotor delay, convulsive crisis, hypotonia, excessive weight, tongue thrust	*NARS2* (NM_024678)	p.M1fs (c.1_2insG)	Compound Htz
									p.P20H (c.59C>A)	
P105	0-10	+	-		de novo	Leukodystrophy / Other: ceroid lipofuscinosis	Mild psychomotor delay, speaking delay, convulsive crisis, ataxia, leukodystrophy, hypotonia, talking regression	*SLC2A1* (NM_006516)	p.S82Y (c.245C>A)	Htz
								DYNC1H1 (NM_001376)	p.N4579I (c.1373A>T)	Htz
P106	0-10	+	+	1/128	AR	Renal disease (tubulopathy). Other : congenital osmotic diarrhea	Watery diarrhea, reccurent thrombocytopenia, anemia, neutropenia	*SLC9A2* (NM_003048)	p.F82L (c.244T>C)	Htz
P107	0-10	+	+	1/32	AR	Bone disease	Hypertelorism, short nose, short neck, bowed femur, bowing bones, wide metaphysis	*NALCN* (NM_052867)	p.I523V (c.1567A>G)	Homo
P108	0-10	+	+	1/16	AR	Other : ophtalmological disorders	Leber, bilateral blindness	*FZD4* (NM_012193)	p.P254fs (c.762delC)	Htz
P109	31-40	+	+	1/16	AR	Neuromuscular disorder	Dystonia, scoliosis	*SPR* (NM_003124)	p.P87H (c.260C>A)	Homo
P110	0-10	+	- (same village)		AR, de novo, X linked	Neurodevelopmental disorder, other : optical neuropathy	Encephalitis, optique neuropathy, epilepsy, hypotonia, convulsive crisis	*SLC9A6* (NM_006359)	p.V419F (c.1255G>T)	Hemizygote
P111	0-10	+	-		X linked, de novo	Neuromuscular disorder	Round face, forehead hump, psychomotor delay	*COL1A2* (NM_000089)	p.D933N (c.2797G>A)	Htz
P112	0-10	+	-		AD, de novo, X linked	Other: connective tissu disease	Hypotonia, aplastic nose, ocular edema, progeroid aspect, narrow thorax, microretrognathism, thick hair, small mouth, cutis laxa (abdominal), dystrophy, curved oesophagus, hyperplasia of the toe nails, pes cavus	*WRN* (NM_000553)	p.K32R (c.95A>G)	Htz

AD Autosomal dominant, *AR* Autosomal recessive, *Htz* Heterozygous, *Homoz* Homozygous.

(inefficient capture of regions rich in GC or bad read depth) and misalignment of reads to the reference sequence are all factors that can occur and lead to false negatives [19, 20].

Conclusion

In conclusion, WES allowed a rapid and cost effective identification of the molecular bases of heterogeneous genetic disorders in the Lebanese population. Our study yielded an overall success rate of 56%, of which around 7% is due to data reanalysis, thus pinpointing the utility of WES reanalysis that should take into consideration the updates in bioinformatics logarithms/annotation tools, novel genetic findings in addition to the occurrence in patients of new clinical manifestations.

Undiagnosed patients need to undergo further molecular investigation. For these cases, Whole Genome Sequencing (WGS) can be performed in order to enhance coverage performance and to detect non-coding mutations (variants modifying gene expression or affecting cryptic splice sites) and structural variations including big insertions and deletions. However, this approach remains laborious and very costly.

On the other hand, current genetic diagnosis studies focuses on the protein-coding regions and ignores the vast majority of non-coding regulating elements such us non-coding RNAs. We believe that the recent advances in the fields of computational biology and experimental technology will allow a better characterization of these elements, which might enable their integration in similar studies in the future.

Altogether, the outstanding improvement in high-throughput sequencing techniques will enable the establishment of low-cost genetic diagnostics, the identification of novel genes and the elucidation of physiological mechanisms, all driving towards personalized medicine.

Abbreviations

AD: Autosomal dominant; AR: Autosomal recessive; CGH: Comparative Genome Hybridization; CNVs: Copy number variations; NGS: Next-generation sequencing; WES: Whole-exome sequencing; WGS: Whole-genome sequencing

Acknowledgements

We express our deepest gratitude and sympathy to the families for their full cooperation throughout the study.

Funding

This work (samples collection, analysis, and interpretation of data) was supported by grants from Saint Joseph University, Beirut, Lebanon.

Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author upon a reasonable request.

Authors' contributions

NJ, CyM and EC conceived, designed the study, performed data interpretation and wrote the manuscript. SC, ZH, NH, JAG, RG and NS conducted the experiments and performed data analysis. EF, JPD and VD conducted acquisition and bioinformatics analysis of the raw data. AF, CDK,

MR, ChM, AdM, SR, BG, HM, MB, RD, DB and AnM performed the recruitment and referral of patients, their clinical investigation and the correlation between genotype – phenotype. All authors have read and approved the manuscript.

Ethics approval and consent to participate

Approval to conduct the study was obtained from the Ethics Committee of Saint Joseph University, Beirut, Lebanon (Reference: FM327). All patients, parents or legal guardians signed an informed consent for participation and sample collection.

Consent for publication

All patients, parents or legal guardians signed an informed consent for data publication.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

[1]Unité de Génétique Médicale, Faculté de Médecine, Campus De l'innovation et du sport, Université Saint-Joseph, rue de Damas, Beirut, Lebanon. [2]Service de technologie de l'information, Saint Joseph University, Beirut, Lebanon. [3]Neuropediatrics Department, Lebanese University, Beirut, Lebanon. [4]Division of Pediatrics, Hotel Dieu de France Hospital, Beirut, Lebanon. [5]Department of Pediatrics Faculty of Medicine, Saint Joseph University, Beirut, Lebanon. [6]Department of Pediatrics, Makassed General Hospital, Beirut, Lebanon. [7]Department of Otolaryngology-Head and Neck Surgery, Hotel Dieu de France Hospital, Faculty of Medicine, Saint Joseph University, Beirut, Lebanon. [8]Pediatric Neurometabolic Unit, Saint George University Medical Center, Beyrouth, Lebanon. [9]Department of Oncology, Hammoud Hospital University Medical Center, Saida, Lebanon. [10]Department of Pediatrics, Notre Dame De Secours University Hospital, Byblos, Lebanon. [11]Department of Pediatrics, Chtoura Hospital, Chtoura, Lebanon. [12]Unité de Génétique Médicale, Faculty of Medicine, Saint Joseph University, Beirut, Lebanon. [13]Institut Jérôme Lejeune, Paris, France. [14]Aix Marseille Univ, Inserm, MMG, U 1251 Marseille, France.

Received: 27 July 2018 Accepted: 11 January 2019
Published online: 21 January 2019

References

1. Green ED, Guyer MS. National Human Genome Research Institute. Charting a course for genomic medicine from base pairs to bedside. Nature. 2011; 470:204–13.
2. Soden SE, Saunders CJ, Willig LK, Farrow EG, Smith LD, Petrikin JE, et al. Effectiveness of exome and genome sequencing guided by acuity of illness for diagnosis of neurodevelopmental disorders. Sci Transl Med. 2014;6:265ra168.
3. 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. 2010;5:e13630.
4. Dixon-Salazar TJ, Silhavy JL, Udpa N, Schroth J, Bielas S, Schaffer AE, et al. Exome sequencing can improve diagnosis and alter patient management. Sci Transl Med. 2012;4:138ra78.
5. Ng SB, Turner EH, Robertson PD, Flygare SD, Bigham AW, Lee C, et al. Targeted capture and massively parallel sequencing of 12 human exomes. Nature. 2009;461:272–6.
6. Maddalena A, Bale S, Das S, Grody W, Richards S, ACMG Laboratory Quality Assurance Committee. Technical standards and guidelines: molecular genetic testing for ultra-rare disorders. Genet Med. 2005;7:571–83.
7. Richards CS, Bale S, Bellissimo DB, Das S, Grody WW, Hegde MR, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. Genet Med. 2008;10:294–300.
8. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16:1215.
9. Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, et al. Automated inference of molecular mechanisms of disease from amino acid substitutions. Bioinforma Oxf Engl. 2009;25:2744–50.
10. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res. 2010;20:1297–303.
11. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. Bioinforma Oxf Engl. 2009;25:1754–60.
12. DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 2011;43:491–8.
13. Desvignes J-P, Bartoli M, Delague V, Krahn M, Miltgen M, Béroud C, et al. VarAFT: a variant annotation and filtration system for human next generation sequencing data. Nucleic Acids Res. 2018;46(W1):W545-W553.
14. Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu Y-F, McSweeney KM, et al. Whole-exome sequencing in undiagnosed genetic diseases: interpreting 119 trios. Genet Med. 2015;17:774–81.
15. Kryukov GV, Shpunt A, Stamatoyannopoulos JA, Sunyaev SR. Power of deep, all-exon resequencing for discovery of human trait genes. Proc Natl Acad Sci U S A. 2009;106:3871–6.
16. Yavarna T, Al-Dewik N, Al-Mureikhi M, Ali R, Al-Mesaifri F, Mahmoud L, et al. High diagnostic yield of clinical exome sequencing in middle eastern patients with Mendelian disorders. Hum Genet. 2015;134:967–80.
17. Doss CGP, Alasmar DR, Bux RI, Sneha P, Bakhsh FD, Al-Azwani I, et al. Genetic epidemiology of Glucose-6-phosphate dehydrogenase deficiency in the Arab world. Sci Rep. 2016;6:37284.
18. Thiagarajah JR, Donowitz M, Verkman AS. Secretory diarrhoea: mechanisms and emerging therapies. Nat Rev Gastroenterol Hepatol. 2015;12:446–57.
19. Ku C-S, Cooper DN, Polychronakos C, Naidoo N, Wu M, Soong R. Exome sequencing: dual role as a discovery and diagnostic tool. Ann Neurol. 2012; 71:5–14.
20. Singleton AB. Exome sequencing: a transformative technology. Lancet Neurol. 2011;10:942–6.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

BMC