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

### **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

<sup>&</sup>lt;sup>1</sup>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



<sup>\*</sup> Correspondence: eliane.chouery@usj.edu.lb

Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 2 of 7

### **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.

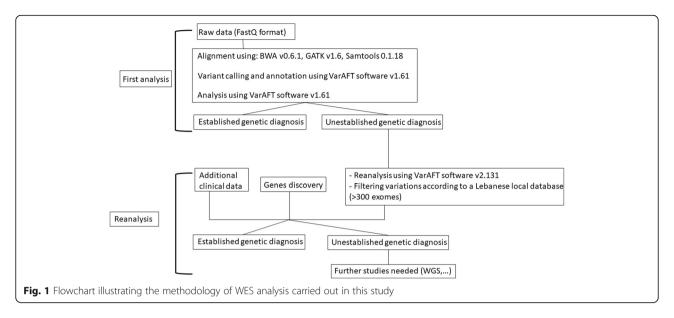
Approval to conduct the study was obtained from the Ethics Committee of Saint Joseph University, Beirut, Lebanon. All patients, parents or legal guardians signed an informed consent for participation, sample collection and data publication. Peripheral blood was then collected from each individual enrolled 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 Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 3 of 7



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

Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 4 of 7

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.

Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 5 of 7

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

Patheol reds	Arter! age in your	Affected	Community	Coefficient of remangainty	Mode of inheritance severding to the moderne	Pathology	Clinical signs	Gow	Manadea	Xypedity
n	20-00			134	AD AR	Noncommunitier dimender Noncommunitier dimender	At the years old, absormed got work street, dietal association brokernal deficiency, produmenter delay, conductor	ANNU PRIN, DESIGN	print(print(s))	18a
	10.00						stola, myschnic stola, corbeller strophy, scione Norrokivckpromal-icky, stola, startic debala, shot			
PIPE				136	14.	Neuromandar dureder	stania, speria diplogia, eksel cercentico Two bids descared from heart	SCHOOL DRIVE (MINER)	\$8170 (-2000-C) \$7700 (-2000-C)	Home
P576	21.90	4		With the purport 121	AR.	Other boar doose and adventages.	false. Our of them processed also subconstitute.	ANCI (SME BIOSE ANTAC (SSE BIOSE)	\$1.00P to SHITH CO \$COMP to 2000 CO p.8290C to 500 CO	Ho He
		٠	•			Observations	Albaton, depigneration of the skin syntagene, phospholis	20K-000(30023)	(135.16-0) •2185.1-280-0	Companyal B
PK	8.10 9.10	٠				Nonemanie dande	Manufer dynagely	CAPST (NOT MINUTE) DOMEST (NOT MINUTE)	NY1-1, 4809CAG 8739CGK 1990-1G	Corposal B
P19	31.40			336	AL	Horig-horie	Mild to server dealbase	ANOTHE SEA ONES	#2000 ( 1900-0) #2000 ( 1900-0) #A500 ( 1900-1) #81000 ( 1900-1)	Corposei E
PII	11-20			1105	AK X Select	Nounder-obspaced deceder	Normalizatiopmental dates, appaints; delectional forchard Normalizatiopmental dates,	ATES (NA_0000)	p.1879(p.2870+C)	Homogete
P12	8.30				denovo	Symdesdepsend donler	čemophi, microsyloly, testralar ocepia, micropeni, cospenial by homple occupyle	CASC PRESENTED	p.2000W (c.6100-x)	Ho
							amblical array, balocyamphy Bildered severe dealers, mand			
P13	6-30		-		AR.	Maubok diseder	they, convolves crisis, secondor, seisser activity and specialty, shutaled lastic acid level	PETITO (SM_00171155)	умпкжча	Ho
F14	6-10		- (same village)		AL.	How desires	Acceptance in included deading, recent relatively process. (My), increasing opposits, and white motor delay, memory delay, metrorist association of the layer delay, metrorist association of the layer delay, memory, and basic appression, and control appression, and memory delays delay del	POZNO PONĮ OKTORNOS	p.1334 (s. 334-4)	Ha
P15	25-30			236	AR.	Horig-director	Dealana Nasrodrodesmostal delan	SCORE IN SHEET	p1.2000 (p. 7900-0) a80079 (c.2000-0)	Hone 180
Pla	8.00	٠			All, de mon	Newsdewingswood doorder	statural delay dynnaryba. plugincaptuly, door sock, beyonde	MECET (MACHINELY)	pt.380(6/444)-1117_1280	19an
P17	21.80		(see vlop)		A.R.	Neurole-ologousid decelor	Macrocopholy, psychomotor delay, hypersignal from the white mater, historial enlargement of the latent randomless, singely of the conduction, singely of conducting fundaments, and conducting conducting fundaments.	352/4 (5M_00000)	p396 (c.2004)	Flore
							eyes in the Frental and subvertical regions of the lefts Speaking-kiley, hyposocouls, wild to review secondarised deathers, recognist ratio	CHOLOU WHILE	AND STREET	. The
PES	8.90	٠			AD, de seuo	Hookg-backe	mild to severe seasons and deathers, terranest onto	MINYA POM (SATE) MOGLA POM (SATE)	p.E4179/P.45(1130-11) p.E4179/P.45(1130-130)	190 190 19a
P20	8-10			18	ALAD	Mitthek desse is her doorsel dild.  Mitthek desse	donard child Psychosom desp, blaved QK 2006; chomples with a server georable with a server georable states princespate of placespate of placespate desperading much begre chickendrial desperading	CLCV col many	p2006 (c.2006/E)	No.
P00	8.10			2	44	Lesindystephy	Leukomopleápalty váh vanishty vátr mator, márosopleár, kneodystesky, plajinopleár, stropadána, doch lair. Ingozneir, bennod apprin, ferodocynymia.	EVERYOR OWN ROLLS	p-Q408 p-1594-43	Hone
P20/P34	31-40		4	PUEL de putour	AR.	Oher: Spidematindona in their tere decreased (MRV)	appleps, inconsequences, appleps; logatiophenosysty.  Documel children Deliva beams around larger rash at the lower log around joints. Exemples of blaces in congress to miner physical égary.	LAMAS (POR 188027)	same same	No
173	b 30				de mon	Noncommunity decodes	to mine physical injury Psychomogy delay, Souther South	STATE DAY 16560 STATE DAY 00001	ARREST CORROLA PRINTED COMPONIO	Ho He
PS	8.30		٠		AD, de sevo	Numbridgecest diorder	Psychococor delay , starlice test della tency, marsia, dyserbeis, conductor disorder	ACAC PROJECTS	pX155-y-45600	Ho
P27	11-20		Para-to from the same village		AD, do novo, AR	System Security Secur	Epitopy  Bose moreor signal shoreralities, bose pais , proclaim distrains, spending deliculation, legenquista	2557 (NM_15000)	pR90(s), 2940-A) pR20C (s, 850-7)	Hone
129	6.30			1128	AI.	tykoy	difficulties, legacyticia Pyrikkeite-dependent spikpey, psychomotor delay	ACARTAL DNI, MILITON	p.100G/g.29T-G	Hone
P30	30-80	b	-		AD	Kend dydanian	Prijeryste labory dramas, sond custo Psychomory drky, muck	PERSONAL DRIVEN	6.9000.0	160
PN	11.30	٠		196	AL.	Neuromaniar dander	mindness, obverted CPK, general host position Elephosphinosopphy, to concer- requisitely articizes, alternated	COLLEG POLICEONS	pERFEC(s20000A)	Home
P10	devenni	٠	Parada from the same village			Modele donor Niesen Phi-hone	requisitely will clima, almost and white matter signal, aphenically, charry and spen-	AMPLY (NA_00001)	\$1809C(s.1256C4)	Home
PSI PSI meter	8-90 45-50		-		AES, de nove, Xilmberk AR	Louiselysingley Rotal dysfunction	mocifs who many since Boal dynamics	PERSONAL CONTROL (CONTROL)	p(045%(c4470°T)	Heriotype Hz
P56	6-10				AD, de novo	Numbrulopsuntil disolar	Psychonia ddy dywedda proudig eyo, tiagdir ber,	2004 (MI) MI 200	\$8021Q161800F4Q	Bo
179										
P36 P32	6-30 8-30			19	AD, de nevo	Novembro deprecial decedor Horizo director	Producesor Arlay, stanta, impaccia, mociotic oplicare Deplicar	SECZAY PRIN (888/94.2) FECZA PAN (884/25)	9-179-10-C) a T1986-0-30C-Ti	16a Hem
P36 P36 P38	8-90 8-90 8-90	-	Para-to them the same village	396	AD, de sevo AB, AB, AB,	Numbridgewid doder Horkelooke Horkglooke Symbologewid doder	Pryderwar delay, sterio, Impeterio, especiario episper Destinos Cardio, maternativos, opisper, microsoptuly, leguento	ACCLAY PORT (MATELL) FECTAL PORT (MATELL) EVELY - CORT (MATELL) SECTAL (SMC, MODER)	6.279.20-C) g.T1986-6.39C-T) g.Y1987-6.489T-A) p.G089-6.287-G)	Hora Hora Hora Hora
P50 P56	8-90 8-90		Para to these the same village	19	4	Neurolandepsentid disorder Herrige disorder Herrige disorder Stande departed disorder Versolandering disorder Versolandering disorder Versolandering disorder	Psychoware deleg viente, impossio, resolvent carleger Dealleres Cardes, milliorantous, oplopy, minocopholy, legotimis Associated information or Psychomotor class, speaking delay ingressione, millional	SECTAL DOS (MEDIC)  FECTA DAM (MEDIC)  FECTA DAM (MEDIC)  SECTA (MA) (MEDIC)  META DAM (MEDIC)  ACTO (MA) (MEDIC)	pTHMM (2000-T) pYHMM is shift DA) pCHMM (2000-T)	Flore Flore
P30 P38 P30	8.30 8-30 10-30 20-40		Para-to free the same village	3 M 3 M	AL AL	Numbrahmenti diseler	microsophuly, leptomia Americania bitanti adamon	2022/ (994,000162) 2022/ (994,0000) 2022/ (994,0000) 2022/ (994,0000) 2022/ (994,0000) 2022/ (994,0000)	#TI9864/380/TI #YURT (CHRITIA)	Hone Hone Hone
20 20 20 20 20 20 20 20 20 20 20 20 20 2	640 640 840 840 640 640		same village	336 336 336	AR AR AR AD AP, Misted AR	Numberlement froder Standerlement froder Neuroneite froder Oter Uder yalone	minocopholy, lypoinnis Americado Islancia derech Psychosote Girs, speaking delay Ingermeiren, mediannal Geolobie Marculat medianne, Govern total positive, progressive scaline difficultie Y and not horring Georder, ghospholia	WORLD DAY SELLS THE SELLS OF TH	p T19866 (200C-T) p Y1987 (200C-C) p X1987 (200C-C)	Hone Hone Horizoper Hone Hone Hone Hone Hone
S 2 2 2 2 2 2	8-10 1-30 1-30 1-30 1-30 1-30		uma vilage		AR AR AD	Seminorium und dende Seminorium dende Normaniste dende Sommodie dende Obet Untergalone Seminorium dende	microsophuly, leptomia Americania bitanti adamon	COSTAN GOVERNOON  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON  VERS	ETTIMO (1907)  ETTIMO	Hone Hone Hone Hone Hone Hone Hone Hone
20 20 20 20 20 20 20 20 20 20 20 20 20 2	640 640 840 840 640 640		same village		AR AR AR AD AP, Misted AR	Numberlement froder Standerlement froder Neuroneite froder Oter Uder yalone	minocopholy, lypoinnis Americado Islancia derech Psychosote Girs, speaking delay Ingermeiren, mediannal Geolobie Marculat medianne, Govern total positive, progressive scaline difficultie Y and not horring Georder, ghospholia	WORLD DAY SELLS THE SELLS OF TH	p T19866 (200C-T) p Y1987 (200C-C) p X1987 (200C-C)	Hone Hone Horizoper Hone Hone Hone Hone Hone
S 2 2 2 2 2 2	8-10 1-30 1-30 1-30 1-30 1-30		uma vilage		AR AR AD	Seminorium und dende Seminorium dende Normaniste dende Sommodie dende Obet Untergalone Seminorium dende	microspholy, legisteins Amountain fastari durante. Psychotoste Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, microspholys Vander vasilature, Gostott teat pendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, vander meily gelieby, bytaklig, midden Microspholostein, Koronierig E- vande meily in Service and Service vande of Service and S	COSTAN GOVERNOON  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON  VERSON CONTROL  VERSON  VERS	ETTIMO (1907)  ETTIMO	Hone Hone Hone Hone Hone Hone Hone Hone
8 2 2 2 2 2 2 2 2	1.30 1.30 1.30 1.30 1.30 1.30 1.30 1.30		Parvis that the same village		AR A	Verander-Americal Bonder  Verander-Amproisi Bonder  Neuromensche dier der  Oden Chilor erstellene  Verander-Amproisi der der  Oden Mongelichen  Verander-Amproisi dereiche  Verander-Amproisi dereiche  Verander-Amproisi dereiche  Neuromensche dereiche  Neuromensche dereiche	minocopholy, lypoinnis Americado Islancia derech Psychosote Girs, speaking delay Ingermeiren, mediannal Geolobie Marculat medianne, Govern total positive, progressive scaline difficultie Y and not horring Georder, ghospholia	2007/2008 (2009)  FELEORS ON (2009)	2 THE CHIEF OF THE	Home Home Home Home Home Home Home Home
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	1.00 10.00 1		Parvio that the same village		AB.	Standardsmooth fluids Standardsmooth fluids Neumanish fluids Neumanish fluids Neumanish fluids Other Uniterative Standardsmooth fluids Standardspooth flui	microspholy, legisteins Amountain fastari durante. Psychotoste Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, microspholys Vander vasilature, Gostott teat pendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, vander meily gelieby, bytaklig, midden Microspholostein, Koronierig E- vande meily in Service and Service vande of Service and S	2007 (2001 (2000) 2007 (2001 (2001) 2007 (2001)	#THM (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007)	Home Home Home Home Home Home Home Home
8 2 2 2 2 2 2 2 2	100 100 100 100 100 100 100 100 100 100		Parvio that the same village		AR.	Street of control famile Street of control famile Street of control famile Street of control famile Obt Chin problem Street of control Str	microspholy, legisteins Amountain fastari durante. Psychotoste Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, microspholys Vander vasilature, Gostott teat pendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, vander meily gelieby, bytaklig, midden Microspholostein, Koronierig E- vande meily in Service and Service vande of Service and S	APPL AND SERVED APPL STORM STORM APPL STORM STORM APPL STORM STORM APPL STORM STORM APPL	#THM (-1907) #THM	Home Home Home Home Home Home Home Home
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	840 840 840 840 840 840 840 840 840 840		Pare Cost be		AB, AB, ABA, ABA, ABA, ABA, ABA, ABA, A	Vermite-objected desired funder.  Neutronic-objected desired in Neutronico desired desired in Neutronico desired desired in Neutronico desired desired in Neutronico de de desired in Neutronico de desired in Neutronico de de desired in Neutronico de de de de de desired in Neutronico de	microspholy, legisteins Amountain fastari durante. Psychotoste Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, legistein Gilys, spaklig dely, microspholys Vander vasilature, Gostott teat pendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, psychotostein Gilys, spaklig dely, mid et respendirer, vander meily gelieby, bytaklig, midden Microspholostein, Koronierig E- vande meily in Service and Service vande of Service and S	2007 (2001 (2000) 2007 (2001 (2001) 2007 (2001)	#THM (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007) #100 (#007)	Home Home Home Home Home Home Home Home
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	100 100 100 100 100 100 100 100 100 100		Parvio that the same village		AR.	Street of control famile Street of control famile Street of control famile Street of control famile Obt Chin problem Street of control Str	microphylic, legromia microphylic, legromia daly ilgromi sity, microbia daly ilgromi sity, microbia daly ilgromi sity, microbia daly ilgromi sity, microbia	2017 (PM 00000) PELOPS 704 (0000) PELOPS 705 (0000)	ATMINI (1997)  PORT (1997)	Home Home Home Home Home Home Home Home
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	840 840 840 840 840 840 840 840 840 840		Pare Cost be		AB, AB, ABA, ABA, ABA, ABA, ABA, ABA, A	Vermite-objected desired funder.  Neutronic-objected desired in Neutronico desired desired in Neutronico desired desired in Neutronico desired desired in Neutronico de de desired in Neutronico de desired in Neutronico de de desired in Neutronico de de de de de desired in Neutronico de	no organic hypomes memorine described by the process of a figure of the process of a figure of the process to the process of the process or the process of the process or the process of the process or the process of the process process process process of the process	ASSESSMENT OF THE STATE OF THE	ATTIME (1997)	Home Thom Thom Thom Thom Thom Thom Thom Thom
SE S E Z Z Z Z Z Z Z Z Z Z E Z E Z E E E E	1.00 10.00 1		Pare Cost be		AB A	Seeming designed from the Seeming designed f	microphylic, legromia microphylic, legromia daly ilgromi sity, microbia daly ilgromi sity, microbia daly ilgromi sity, microbia daly ilgromi sity, microbia	ADVENDE EXPERIMENTAL PROPERTY OF A CONTROL O	ATTIME (1997)  ATTIME	Home Theory Theo
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	8.00 917 10.00 10.		Para o front the same village	19	AB. Miled AB. Mi	Name of the Control o	no organic hypomes memorine described by the process of a figure of the process of a figure of the process to the process of the process or the process of the process or the process of the process or the process of the process process process process of the process	SECTION CONTROL  SECTIO	### (### (### (### (### (### (### (###	Hone Hone Hone Hone Hone Hone Hone Hone
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	8.20 9.10 10.3		Para-charde	136 138 336 With her person	AB. AB. AB. AD. Shided AB. AD. AD. AD. AD. AD. AD. AD. AD. AD. AD	Mean Annual Balley  Neman Annu	www.mpho.typem.  See See See See See See See See See Se	ASTON ON STORM  (FLEXION ON STORM ON STORM  (FLEXION ON STORM ON STORM  (FLEXION ON STORM	ATTENDED TO THE STATE OF THE ST	House
\$E & E Z Z Z Z Z Z E E E E E E E E E E E	8.00 9.00 10.0		Pare that by	136 139 136 136 136 136 136	AB.	Mean information design of the Version objection object of the Version object of the Version objection object of the Version object of	security of the property of th	AND THE STATE OF T	4.000 (1.	Home Home Home Home Home Home Home Home
2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	8.20 9.10 10.3		Para-charde	136 138 336 With her person	AB. AB. AB. AD. Shided AB. AD. AD. AD. AD. AD. AD. AD. AD. AD. AD	Mean Annual Balley  Neman Annu	was enabled hypothesis and the control of the contr	ASTON ON STORM  (FLEXION ON STORM ON STORM  (FLEXION ON STORM ON STORM  (FLEXION ON STORM	ATTENDED TO THE STATE OF THE ST	House
84 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	8.20 9.10 10.3		Period Date (Control of Control o	136 139 136 136 136 136 136	AR.  AR.  AR.  AR.  AR.  AR.  AR.  AR.	Security of the Control of the Contr	security of the property of th	AND COLUMN TO THE COLUMN TO TH	4.70% (1970) 4.70%	Hone There T
84 8 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	810 910 10-30 840 10-30		Period Date (Control of Control o	136 139 136 136 136 136 136	AD, March Ma	Mention designed designed for the New Action of Section	The control of the co	AND COLUMN AND ADDRESS OF A STATE	4.70% (1975) 4.70%	Home Them Them Them Them Them Them Them Th
84 8 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	8.10 8.10 10.3		Period Date (Control of Control o	136 139 136 136 136 136 136	34 AS 300 AS 400	Secretary and se	The control of the co	AND COLUMN AND ADDRESS OF A STATE	4.70% (1970) 4.70%	Home Home Home Home Home Home Home Home
84 8 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	100   100		Period Date (Control of Control o	136 139 136 136 136 136 136	34 AS 300 AS 400	Secretary and se	The control of the co	AND COLUMN AND ADDRESS OF A STATE	4.70% (1970) 4.70%	Home Plant P
SE 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	1.00   1.00		Period Date (Control of Control o	13 A SA S	24 A 3 Sheet A 3 A 3 Sheet A 3 A 3 A 3 A 3 A 3 A 3 A 3 A 3 A 3 A	Security of the Control of the Contr	Secretary of the control of the cont	ANTO COLUMNO SERVICE DE LA COLUMNO SERVICE D	ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ACTION AND ACTION	Home Hard Hard Hard Hard Hard Hard Hard Hard
SE 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	100   100		Period Date (Control of Control o	13 A SA S	24 A 3 Sheet A 3 A 3 Sheet A 3 A 3 A 3 A 3 A 3 A 3 A 3 A 3 A 3 A	Security of the Control of the Contr	The control of the co	ANTO COLUMNO SERVICE DE LA COLUMNO SERVICE D	ACTION AC	Home Hard Hard Hard Hard Hard Hard Hard Hard
SE 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	8.20 8.20		per digr	13 A SA S	24 AS Model  43 AS Model  44 AS Model  45 AS Model  45 AS Model  46 AS Model  46 AS Model  47 AS Model  48 AS	Section designed designed for the New Association of the New Associa	Secretary of the control of the cont	AND COMPANY OF THE PARTY OF THE	ACTION CONTROL OF THE	Home Hard Hard Hard Hard Hard Hard Hard Hard
SET 2 12 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	1.20 		are object.	13 A SA S	24 A Shibard A S	Secretary and se	Section of the control of the contro	ANTO COLOR DESCRIPTION OF THE PROPERTY OF THE	ACTION AC	James House
et e e e e e e e e e e e e e e e e e e	100   100		Tree facts	13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	34 At 3 Mark 1 Mark 1 At 3 Mark 1 Ma	Section of the control of the contro	The second secon	ADMINISTRATION OF THE PROPERTY	ACTION CONTROL OF THE	House
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	10.00   10.0		are dept.	13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	### 150 A 100 A 10	Secretary of the Control of the Cont	Section of the control of the contro		ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION ACTION ACTION AND ACTION ACTIO	Home Plant P
SEE 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	10.00   10.0		To dear the control of the control o	13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	24 A 20 A	Secretary and se	Section of the control of the contro		ACTION CONTROL OF THE	The
3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	10.00   10.0		To dear the control of the control o	13 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	### 15	Section Annual Section Section (Inc.)  When the Committee of the Committee	Section of the control of the contro		ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION ACTION ACTION AND ACTION ACTIO	Home State S
6E 5 H Z Z Z Z Z Z Z 2 L Z L Z L Z L Z L Z L Z	100   100		To dear the control of the control o	134	24	Secretary and se	Section of the control of the contro		ACTION AND ACTION ASSESSMENT ASSE	Personal Property of the Prope
हिंद है	100   100		Franchischer (1997)	134	24 A 10 Month of the control of the	Secretary and se	Section of the control of the contro		ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ACTION AND ACTION	Name
62 8 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	619 629 629 620 620 620 620 620 620 620 620 620 620		Transfer days	1 N	24 A 2 Sheld A 2	Section of the control of the contro	The second secon		ACTION CONTROL OF THE	1000   1000
5 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6			Franchischer (1997)	3 M	24 A 10 Month of the control of the	Secretary and se	Section of the control of the contro		ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ACTION AND ACTION	Home Part of the P
60 0 0 1 2 2 2 2 2 2 2 2 1 2 1 2 1 2 1 2			Transfer days	1 N	24 A 2 Sheld A 2	Section of the control of the contro	Secretary of the control of the cont		ACTION AND ACTION ACTION AND ACTION ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ACTION AND ACTION AND ACTION AND ACTION ACTIO	1000   1000
62 8 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2	1		Transfer days	1 N	1	Security of the Control of the Contr	Secretary of the control of the cont		ACTION AC	1000   1000
2   2   2   3   3   3   3   3   3   3			one death	1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N	1	Section Annual Conference of C	Section of the control of the contro		ACTION CONTROL OF THE	1   1   1   1   1   1   1   1   1   1
2 2 3 2 2 3 2 3 2 3 2 3 3 3 3 3 3 3 3 3			and district the second	1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N 1 N	### 100 Miles   10	Security of the Control of the Contr	Secretary of the control of the cont		ACTION AC	1   1   1   1   1   1   1   1   1   1
	1		and district the second	19 19 19 19 19 19 19 19 19 19 19 19 19 1	### A PA P	Section of the control of the contro	The second secon		ACTION AC	The color
2 2 3 2 2 3 2 3 2 3 2 3 3 3 3 3 3 3 3 3	1		and district the second	19 19 19 19 19 19 19 19 19 19 19 19 19 1	### 100 Miles   10	Section of the control of the contro	Section of the control of the contro		ACTION AC	1000   1000
1	1		and district the second	19 19 19 19 19 19 19 19 19 19 19 19 19 1	20	Security of the Control of the Contr	The second secon		ACTION AND ACTION AND ACTION AND ACTION AND ACTION AND ACTION ACTION AND ACTION	The color

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://gen e.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 Furthermore, lack of coverage of coding regions

Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 6 of 7

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,	IVD (NM 002225)	p.S281R (c.841A>C)	Homo
P101	0-10	+	-		AD, de novo, Xlinked	Neurodevelopmental disorder	essential tremor  Psychomotor delay, developmental delay, microcephaly, retrognathism,	FASTKD2 (NM 014929)  ALDH18A1 (NM_002860.3)	p.R358* (c.1072C>T) c.2207-3C>T	Homo Htz
							testicular ectopia, dysmorphia	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) <sub>13</sub> )	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	NARS2 (NM_024678)	p.M1fs (c.1_2insG)	Compound Htz
							weight, tongue thrust		p.P20H (c.59C>A)	
P105	0-10	+	-		de novo	Leukodystrophy / Other: ceroid lipofuscinosis	Mild psychomotor delay, speaking delay, convulsive crisis,	SLC2A1 (NM_006516)	p.S82Y (c.245C>A)	Htz
							ataxia, leukodystrophy, hypotonia, talking regression	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.1523V (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 highthroughput 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,

Jalkh et al. BMC Medical Genomics (2019) 12:11 Page 7 of 7

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

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

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

### References

- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- 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.
- Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988:16:1215.
- 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.

- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The genome analysis toolkit: a MapReduce framework for analyzing nextgeneration DNA sequencing data. Genome Res. 2010;20:1297–303.
- Li H, Durbin R. Fast and accurate short read alignment with burrowswheeler transform. Bioinforma Oxf Engl. 2009;25:1754–60.
- 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.
- 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.
- Zhu X, Petrovski S, Xie P, Ruzzo EK, Lu Y-F, McSweeney KM, et al. Wholeexome sequencing in undiagnosed genetic diseases: interpreting 119 trios. Genet Med. 2015;17:774–81.
- 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.
- 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.
- 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.
- Thiagarajah JR, Donowitz M, Verkman AS. Secretory diarrhoea: mechanisms and emerging therapies. Nat Rev Gastroenterol Hepatol. 2015;12:446–57.
- 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

