UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – UFRGS

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

AVALIAÇÃO E VALIDAÇÃO DA UTILIDADE CLÍNICA DO SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS) PARA CONFIRMAÇÃO DO DIAGNÓSTICO DE DOENÇAS LISOSSÔMICAS SELECIONADAS

Diana Elizabeth Rojas Málaga

Orientador: Prof. Dr. Roberto Giugliani

Co-orientadora: Dra. Ana Carolina Brusius-Facchin

Porto Alegre, outubro de 2018.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – UFRGS

PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA E BIOLOGIA MOLECULAR

AVALIAÇÃO E VALIDAÇÃO DA UTILIDADE CLÍNICA DO SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS) PARA CONFIRMAÇÃO DO DIAGNÓSTICO DE DOENÇAS LISOSSÔMICAS SELECIONADAS

Diana Elizabeth Rojas Málaga

Tese submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para obtenção do grau de Doutor em Ciências (Genética e Biologia Molecular).

Porto Alegre, outubro de 2018.

Este trabalho foi desenvolvido no Laboratório de Genética Molecular do Serviço de Genética Médica do Hospital de Clínicas de Porto Alegre. Teve como instituições financiadoras: Fundo de Incentivo à Pesquisa e Eventos do Hospital de Clínicas de Porto Alegre (FIPE/HCPA) (Projetos 15-0465, 16-0435 e 18-0249) e o Instituto Nacional de Genética Médica Populacional- INaGeMP (CNpq/CAPES/FAPERGS). Bolsa de estudos financiada pela instituição Innovate Peru (Convenio N° 073-FINCYT-BDE-2014).

AGRADECIMENTOS

À Deus, pelo dom da vida, por tudo!

Ao meu orientador Roberto, pela confianza no meu trabalho;

À minha co-orientadora Ana, pela presença, envolvimento constante e integral neste trabalho, e por sempre me incentivar;

Àos amigos do Laboratório de Genética e Biologia Molecular do SGM-HCPA, em especial à Eriza, Rowena, Ândria e Francyne, por estarem sempre prontas a me ajudar, pelo carinho, amizade, e minutos/horas/dias e anos de altas conversa. Gurias, vocês acabaram se tornando parte da minha família;

Aos "Fabrellos", por serem por muitos anos minha segunda família em Porto Alegre e por me tratarem sempre com muito amor.

Ao Eliézer, pelo amor e parceria na reta final do doutorado.

À minha família no Perú, pelo apoio incondicional nestes anos longe deles, pelas palavras diárias de incentivo e por sempre colocarem os meus sonos acima dos seus;

Ao PPGBM pelos conhecimentos repassados;

Ao Innovate Peru, pelo apoio financeiro;

A todos, que de alguma forma fizeram toda a diferença no desenvolvimento deste trabalho o meu MUCHAS GRACIAS!

SUMÁRIO

LISTA DE ABREVIATURAS

a-CGH	*Array comparative genomic hybridization*
ACMG	*American College of Medical Genetics and Genomics*
ANVISA	Agência Nacional de Vigilância Sanitária
ARMS	*Amplification-refractory mutation system*
CLN2	Lipofuscinose ceróide neuronal tipo 2
CNV	C*opy Number Variation*
Cr.	Cromossomo
dHPLC	*Denaturing high performance liquid chromatography*
DL	Doença lisossômica
DLD	Doença lisossômica de depósito
ADN	Ácido desoxirribonucleico
EIM	Erros inatos do metabolismo
ExAC	*Exome Aggregation Consortium*
Gb	Gigabase
GID	Grandes indels
HCPA	Hospital de Clínicas de Porto Alegre
HRM	*High resolution melting*
IGV	*Integrative Genomics Viewer*
Indels	Inserções e deleções
ISFET	*Ion-sensitive field-effect transitor*
ISP	*Ion Sphere Particle*
LEIM	Laboratório de Erros Inatos do Metabolismo
LGM	Laboratório de Genética Molecular

M	Missense
MLPA	*Multiplex Ligation-dependent Probe Amplification*
MPS	Mucopolissacaridoses
MPSPS	*Mucopolysaccharidoses Plus Syndrome*
N	Nonsense
NGS	*Next-generation sequencing*
NIH	*National Institutes of Health*
NPC	Niemann-Pick
pb	Pares de bases
PCR	Reação em cadeia da polimerase
PGM	*Personal Genome Machine*
PID	Pequenas indels
R	Regulatórias
RC	Rearranjos complexos
RFLP	*Restriction fragment length polymorphism*
RT-PCR	*Real-time polymerase chain reaction*
SGM	Serviço de Genética Médica
SIPF	Sangue impregnado em papel filtro
SNC	Sistema nervoso central
SNPs	Single nucleotide polymorphisms
SSCP	*Single strand conformation polymorphism*
SUS	Sistema Único de Saúde
TN	Triagem neonatal
TRE	Terapia de reposição enzimática

TRS	Terapia de redução de substrato
UTR	*Untranslated region*
WES	Whole-exome sequencing
WGS	*Whole-genome sequencing*

LISTA DE FIGURAS

LISTA DE TABELAS

RESUMO

Introdução: As doenças lisossômicas (DLs) são patologias genéticas que, apesar de serem classificadas como raras, acometem uma significativa porcentagem da população. Muitos fatores tornam seu diagnóstico desafiador, entre eles a variabilidade no fenótipo, com poucos sinais e sintomas clínicos específicos. O desenvolvimento de ferramentas inovadoras de investigação, destacando-se os novos métodos de sequenciamento massivo, reduziria a "odisseia diagnóstica" enfrentada pelo paciente e sua família, proporcionaria um diagnóstico mais precoce com melhor resultado no tratamento nas situações em que há terapia disponível, além de um adequado aconselhamento genético. A tecnologia de sequenciamento de nova geração (*next-generation sequencing*, NGS) tem claras vantagens sobre as técnicas de sequenciamento convencional, oferecendo um alto rendimento diagnóstico ao permitir definir um espectro mutacional abrangente. O NGS permite o sequenciamento de vários genes simultaneamente com custo global relativamente baixo, tornando painéis de genes uma alternativa atrativa para o *screening* genético. Esta abordagem é capaz de detectar, de maneira altamente específica, variantes *missense*, *nonsense*, de sítio de *splicing*, e pequenas indels, e algumas grandes deleções em homozigose ou hemizigose. No entanto, para que essa tecnologia seja aplicável à prática clínica, é necessária uma etapa de validação prévia para a determinação dos parâmetros críticos de análise desde o processamento da amostra até a análise e interpretação de dados. Após a validação dos métodos, será possível a avaliação da utilidade clínica de painéis NGS para diversas aplicações, incluindo seu uso como método confirmatório nos casos de triagem neonatal alterada para doenças lisossômicas.

Objetivos: 1) Desenhar e validar uma estratégia baseada no NGS para a análise de 24 genes, incluídos em 3 painéis distintos de acordo com parâmetros pré definidos, e associados com 22 DLs; 2) Avaliar a qualidade e eficiência do ADN extraído de sangue impregnado em papel filtro (SIPF) para uso no NGS; 3) Avaliar a utilidade clínica do NGS para: a) diagnóstico molecular, b) diagnóstico diferencial e, c) confirmação diagnóstica de casos alterados na triagem neonatal para pelo menos 4 dessas condições (doença de Gaucher, Fabry, Pompe e Niemann-Pick tipo A/B).

Metodologia: Estudo descritivo, amostragem por conveniência, incluindo pacientes com diagnóstico clínico e bioquímico das DLs estudadas. Foi utilizada a plataforma NGS Ion

Torrent Personal Genome Machine (Thermo Fisher Scientific) para sequenciamento de três painéis de genes, desenhados através da ferramenta Ion Ampliseq™ Designer (Thermo Fisher Scientific). Para a validação foram incluídos pacientes com diagnóstico molecular prévio por sequenciamento de Sanger. Para a extração de ADN de SIPF, foi avaliado o método não comercial de fenol-clorofórmio. A utilidade clínica do NGS foi estabelecida através do: a) diagnóstico molecular de um grupo de pacientes com suspeita de lipofuscinose ceróide neuronal tipo 2 (CLN2) para o estabelecimento do genótipo (*TPP1*) associado à doença, assim como estabeleceu-se o diagnóstico molecular de outros pacientes com suspeita clínica e/ou bioquímica de algumas das DLs selecionadas, b) diagnóstico diferencial de um paciente utilizando um dos painéis NGS desenhados e, c) avaliação das amostras de recém-nascidos (obtidas em estudo paralelo) que tiveram resultados inicialmente alterados na triagem neonatal bioquímica para algumas das DLs selecionadas.

Resultados: 1) Três painéis foram desenhados, cada um consistindo em dois *pools* de *primers* que amplificam as regiões codificantes e 20pb da junção exon-intron. Os painéis A, B e C têm amplitude de cobertura de 97.74, 99.6 e 98.38%, respectivamente. Nesta validação foi possível estabelecer a sensibilidade, especificidade e limitações de cada painel (Artigo 1); 2) foi estabelecida a metodologia para a extração de ADN para seu uso em vários processos moleculares *downstream*: PCR convencional, Real-Time PCR, PCR-RFLP, MLPA, Sequenciamento de Sanger e NGS na plataforma Ion Torrent PGM™ (Artigo 2); 3) A utilidade clínica dos painéis desenhados foi estabelecida através de: a) Genotipagem de 48 pacientes com suspeita clínica e bioquímica de CLN2 (painel C) (Artigo 4), e de 3 outros casos: 1 caso de Niemann-Pick tipo B (*SMPD1*) (painel B) e 2 casos de doença de Danon (*LAMP2*) (painel A) (Artigo 1), b) o diagnóstico diferencial foi realizado através do caso de um paciente com suspeita inicial de Niemann-Pick C que, após a análise molecular, foi diagnosticado com Niemann-Pick tipo B utilizando o painel B, (Artigo 1) c) a aplicabilidade dos painéis de genes para a confirmação diagnóstica em casos de triagem neonatal foi realizada em 2 casos com resultados alterados na triagem neonatal: um caso de Pompe (painel A) e outro de Gaucher (painel B) (Artigo 3 e Anexo I).

Conclusão: A abordagem por NGS através do uso de painéis, foi capaz de identificar diferentes alterações genéticas nos genes estudados, incluindo variantes do tipo *missense*, *nonsense*, de sítio de *splicing*, e pequenas indels. Estes painéis oferecem uma estratégia de triagem dos 24 genes associadas às DLs selecionadas. Este trabalho foi inovador ao utilizar o NGS para a análise molecular dos genes associados às DLs no Brasil; foi desenvolvido no

Laboratório de Genética e Biologia Molecular do Serviço de Genética Médica do Hospital de Clínicas de Porto Alegre, conhecido como um centro de referência em diagnóstico e pesquisa de DLs, e que atualmente tem os painéis NGS como principal ferramenta de diagnóstico molecular.

ABSTRACT

Introduction: Lysosomal diseases (LDs) are genetic pathologies that, although classified as rare, affect a significant percentage of the population. Many factors make its diagnosis challenging, including variability in phenotype, with few specific clinical signs and symptoms. The development of innovative research tools, highlighting new methods of massive sequencing, would reduce the "diagnostic odyssey" faced by the patient and his family, would provide an earlier diagnosis with better treatment outcome in situations where therapy is available, in addition to adequate genetic counseling. Next-generation sequencing (NGS) technology has clear advantages over conventional sequencing techniques, offering a high diagnostic yield by enabling a comprehensive mutational spectrum to be defined. The NGS allows the sequencing of several genes simultaneously with relatively low overall cost, making gene panels an attractive alternative for genetic screening. This approach is able to detect, in a highly specific manner, missense, nonsense, splicing, and small indels variants, and some large deletions in homozygosis or hemizygosis. However, for this technology to be applicable to clinical practice, a prior validation step is required to determine the critical parameters of analysis from sample processing to data analysis and interpretation. After validation of the methods, it will be possible to evaluate the clinical usefulness of NGS panels for several applications, including their use as a confirmatory method in cases of altered neonatal screening for lysosomal diseases.

Objectives: 1) To design and validate an NGS-based strategy for the analysis of 24 genes, included in 3 different panels according to pre-defined parameters, and associated with 22 LDs; 2) To evaluate the quality and efficiency of DNA extracted from dried blood spots (DBS) for its use in NGS; 3) To evaluate the clinical use of NGS for: a) molecular diagnosis, b) differential diagnosis and c) diagnostic confirmation of altered cases in neonatal screening for at least 4 of these conditions (Gaucher, Fabry, Pompe and Niemann-Pick type A / B).

Methodology: Descriptive study, convenience sampling, including patients with clinical and biochemical diagnosis of selected LDs. The NGS Ion Torrent Personal Genome Machine (Thermo Fisher Scientific) platform was used for the sequencing of three gene panels, designed using the Ion Ampliseq ™ Designer tool (Thermo Fisher Scientific). For the validation, patients with previous molecular diagnosis by Sanger sequencing were included. For the extraction of DBS DNA, a non-commercial method of phenol-chloroform was evaluated. The clinical utility of NGS was established through: a) molecular diagnosis

of a group of patients with suspected type 2 neuronal cercoid lipofuscinosis (CLN2) to establish the disease-associated genotype (*TPP1*), as well as the molecular diagnosis of other patients with clinical and / or biochemical suspicion of some of the selected LDs, b) differential diagnosis of a patient using one of the designed NGS panels, and c) evaluation of the newborn samples (obtained in a parallel study) that initially had altered results in the neonatal screening for some of the selected LDs.

Results: 1) Three panels were designed, each consisting of two pools of primers that amplify the coding regions and 20pb of the exon-intron junction. Panels A, B and C have coverage range of 97.74, 99.6 and 98.38%, respectively. In this validation it was possible to establish the sensitivity, specificity and limitations of each panel (Article 1); 2) the methodology for the extraction of DNA for its use in several downstream molecular processes was established: conventional PCR, Real-Time PCR, PCR-RFLP, MLPA, Sanger Sequencing and NGS in the Ion Torrent PGM^{TM} platform; 3) The clinical use of the panels was established through: a) Genotyping of 48 patients with clinical and biochemical CLN2 (panel C) (Article 4), and of 3 other cases: 1 case of Niemann-Pick type B (*SMPD1*) (panel B) and 2 cases of Danon's disease (*LAMP2*) (panel A) (Article 1), b) he differential diagnosis was made through the case of a patient with initial suspicion of Niemann-Pick C who, after molecular analysis, was diagnosed with Niemann-Pick type B using panel B (Article 1), c) the applicability of the gene panels for diagnostic confirmation in cases of neonatal screening was performed in 2 cases with altered results in neonatal screening: one case of Pompe (panel A) and another case of Gaucher (panel B) (Article 3 and Annex I) .

Conclusion: The NGS approach using gene panels was able to identify different genetic alterations in the genes studied, including missense, nonsense, splicing variants, and small indels. These panels offer a screening strategy for the 24 genes associated with the selected LDs. This work was innovative by using NGS for the molecular analysis of genes associated with LDs in Brazil; it was developed in the Laboratory of Genetics and Molecular Biology of the Medical Genetics Service of the Hospital de Clinicas de Porto Alegre, known as a reference center in diagnosis and research of LDs, and that currently has the NGS panels as the main tool of molecular diagnosis.

CAPÍTULO 1: INTRODUÇÃO

1.1 ERROS INATOS DO METABOLISMO

O termo “erros inatos do metabolismo” (*inborn errors of metabolism*, EIM) foi proposto pela primeira vez em 1902 pelo doutor Archibald E. Garrod ao descrever doenças como alcaptonúria, albinismo, cistinúria, porfiria e pentosúria[1].

Os EIM são doenças metabólicas hereditárias resultantes do defeito de uma enzima envolvida na síntese, transporte ou degradação de moléculas, o que leva à interrupção da via metabólica ou a interferência de uma via alternativa[2,3]. Este bloqueio leva à falta ou excesso de um determinado substrato.

A incidência dos EIM varia muito e depende da população. Apesar de serem catalogados individualmente como doenças raras, em conjunto têm uma incidência combinada que varia de 1 em 784 nascidos vivos até 1 em 2555 nascidos vivos[4–6], representando cerca do 10% de todas as doenças monogênicas[7,8]

Aproximadamente 600 EIM têm sido reconhecidos até o momento, e esse número está aumentando constantemente à medida que novos conceitos e técnicas se tornam disponíveis para a identificação de fenótipos bioquímicos[9]. Aproximadamente 25% se manifestam no período neonatal, os recém-nascidos são usualmente sadios ao nascimento com sinais clínicos se desenvolvendo em horas até dias após o nascimento. Estes sinais são geralmente inespecíficos e podem incluir diminuição da atividade, má alimentação, dificuldade respiratória, letargia ou convulsões. Esses sinais são comuns a várias outras condições neonatais, como sepse e disfunção cardiopulmonar[10,11].

Devido a heterogeneidade dos EIM, a sua classificação se torna um desafio. A classificação pode se basear no órgão afetado (distúrbios neurológicos ou hepáticos), organela afetada (distúrbios mitocondriais, peroxissomais ou lisossômicos) e/ou na idade de início (desde início neonatal até início juvenil ou adulto)[9]. Estes podem ser classificados também de acordo com a fisiopatologia em três grupos principais (Tabela 1)[12]:

Tabela 1. Tipos de Erros Inatos do Metabolismo

I. Distúrbios do metabolismo intermediário
Metabolismo e transporte de aminoácidos
Oxidação de ácidos graxos e cetogênese
Metabolismo e transporte de carboidratos
Relacionado com vitaminas (cobalamina, folato)
Metabolismo peptídico
Metabolismo mineral
Metabolismo Energético Mitocondrial
II.Distúrbios da biossíntese e quebra de moléculas complexas
Metabolismo de purinas e pirimidinas
Armazenamento lisossomal
Peroxissomas
Metabolismo isoprenóide e esterol
Ácido biliar e metabolismo do heme
Glicosilação
Metabolismo das lipoproteínas
III.Distúrbios do metabolismo dos neurotransmissores
Metabolismo de glicina e serina
Metabolismo de amina biogênica
Metabolismo gama-aminobutirato
Outros (por exemplo, convulsões dependentes de piridoxina ou dependentes de ácido folínico, deficiência de sulfito oxidase)

O diagnóstico é realizado através de testes bioquímicos, enzimáticos ou moleculares. O diagnóstico bioquímico baseia-se na identificação de níveis elevados de um substrato ou níveis baixos de produtos, o qual é realizado por espectrometria de massas. O ensaio de atividade enzimática pode ser realizado para alguns distúrbios. Alternativamente, quando os testes bioquímicos não podem confirmar nem descartar um EIM, o sequenciamento do gene associado à doença pode ser realizado para a confirmação do diagnóstico[13]. No entanto, novas tecnologias estão sendo empregadas para este fim e para o descobrimento de novas doenças, entre elas encontram-se o *whole-exome sequencing* (WES) e o *untargeted metabolomics*. O WES, através da tecnologia do NGS, permite que os 20.000 genes do nosso exoma sejam sequenciados de maneira simultânea[14]. Por outro lado, a análise metabolômica permite o estudo e a caracterização de pequenas moléculas numa amostra biológica e pode fornecer uma visão geral do status bioquímico do indivíduo do qual a amostra foi obtida[13].

1.2 DOENÇAS LISOSSÔMICAS (DLs)

1.2.1 Definição, etiologia e manifestações clínicas

Os lisossomos foram descobertos por Christian de Duve em 1955, no seu laboratório em Leuven, Bélgica[15]. Estas organelas intracelulares, degradam e reciclam uma série de macromoléculas extracelulares, incluindo glicosaminoglicanos, esfingolipídios, fragmentos de glicogênio e proteínas. Esta função catabólica é realizada através da ação de aproximadamente 60 tipos diferentes de hidrolases ácidas, como glicosidases, sulfatases, fosfatases, lipases, fosfolipases, proteases e nucleases[16].

Dez anos após o descobrimento desta organela, o pesquisador H.G. Hers desenvolveu o conceito de doença lisossômica (DLs)[17]. Até o momento, aproximadamente 60 DLs têm sido descritas (Tabela 2). A maioria delas é herdada de forma autossômica recessiva, embora alguns estejam ligadas ao cromossomo X (Fabry, Hunter ou MPS II e Danon) ou de herança dominantes (CLN4).

As DLs são causadas por mutações em genes que codificam hidrolases ácidas solúveis, proteínas integrais de membrana, proteínas ativadoras, proteínas transportadoras ou proteínas não lisossomais que são necessárias para a função lisossomal, resultando no acúmulo de substratos específicos nestas organelas e na produção de grandes vacúolos intracelulares. O acúmulo é progressivo e causa perda da função celular e tecidual em diferentes fases do desenvolvimento da doença, sendo que as manifestações clínicas não são específicas deste grupo de doenças[16].

A maioria dos pacientes com DLs nasce aparentemente saudável e os sintomas aparecem progressivamente[18]. Alguns sintomas comuns incluem dismorfismo facial, atrasos de desenvolvimento, infecções recorrentes, problemas musculares, organomegalia e alterações esqueléticas. Muitas DLs afetam o sistema nervoso central (SNC) e a maioria dos pacientes tem uma redução significativa na expectativa de vida e aumento da morbidade[19]. A avaliação de tecidos através de histologia, análise enzimática ou de imagens podem ajudar no diagnóstico. Tecidos que são úteis no diagnóstico das DLs estão listados na Tabela 3[16].

Quase todas as DLs mostram um amplo espectro clínico com relação à gravidade dos sintomas, progressão e idade de início que podem variar acentuadamente dentro de um único distúrbio. Isso pode ser explicado, em parte, pela atividade enzimática residual associada ao genótipo do paciente. Assim, os fenótipos dependerão do tipo, quantidade e local de armazenamento do material não degradado, sendo que a sua diversidade dificulta o

diagnóstico precoce. O estabelecimento de uma correlação genótipo-fenótipo é difícil devido a vários fatores. Primeiro, existe um número elevado de mutações privadas, o que implica que a maioria de famílias possui diferentes mutações. Outra causa que dificulta a correlação é a variabilidade clínica entre pacientes portadores da mesma mutação, mesmo entre pacientes da mesma família, sugerindo que existam outros fatores responsáveis de levar à alteração da função da proteína e subsequentemente ao fenótipo clínico, entre eles variantes polimórficas e modificadores genéticos[20,21].

Embora as DLs sejam individualmente raras, a prevalência combinada é relativamente alta, variando de 7,5 por 100.000 na Colúmbia Britânica[4] a 23,5 por 100.000 nascidos vivos nos Emirados Árabes Unidos[22], com as esfingolipidoses como o grupo mais prevalente, seguido pelas mucopolissacaridoses. No Brasil, esta prevalência combinada foi estabelecida pelo grupo de Giugliani et al.[23], em 5.56 por 100.00 nascidos vivos[23].

Alguns distúrbios são mais prevalentes que outros em certas áreas geográficas ou em grupos populacionais específicos (por exemplo, as doenças de Gaucher, Tay-Sachs, Niemann-Pick tipo A e mucolipidose IV são mais comuns em judeus Ashkenazi), em grande parte como resultado do efeito fundador de mutações ancestrais[24–26].

As DLs são classificadas frequentemente de acordo com o substrato acumulado[27] (Tabela 2). Clinicamente esta classificação é muito útil e aceita. Assim, os distúrbios em que prevalece o acúmulo de glicosaminoglicanos são classificados como mucopolissacaridoses (MPS), aqueles dominados pelo armazenamento de lipídios como lipidoses. No entanto, deve ser destacado que na maioria de DLs mais de um composto é acumulado e, em alguns distúrbios, por várias razões, o material armazenado pode ser bastante heterogêneo[28].

Tabela 2. Classificação das doenças lisossômicas (DLs)

DLs	Enzima deficiente ou defeito	Produto de depósito	Gene	Nº Mutações descritas*
Esfingolipidoses				
Fabry	α-Galactosidase A	Gb3	*GLA*	922
Farber	Ceramidase ácida	Ceramida	*ASAH1*	71
Gaucher	β-Glucocerebrosidase	Glicosilceramida	*GBA1*	472
Niemann-Pick A & B	Esfingomielinase ácida	Esfingomielina	*SMPD1*	251
GM1 gangliosidosis	β-Galactosidase	GM1	*GLB1*	215
GM2 gangliosidosis (Tay-Sachs)	β-hexosaminidase A	GM2	*HEXA*	184
GM2 gangliosidosis (Sandhoff)	β-hexosaminidase B	GM2	*HEXB*	112
GM2 gangliosidosis (GM2-activator deficiency)	Proteína Ativadora de GM2	Gangliósidos GM2 e outros glicolípidos	*GM2A*	10
Krabbe	β-galactocerebrosidase	Galactosilceramida	*GALC*	237
Leucodistrofia metacromática	Arilsulfatase A	Sulfatídeos	*ARSA*	227
Mucopolissacaridoses				
MPS I (Hurler, Scheie, Hurler/Scheie)	α-L-iduronidase	Dermatan e heparan sulfato	*IDUA*	279
MPS II (Hunter)	Iduronato-2-sulfatase	Dermatan e heparan sulfato	*IDS*	632
MPS III A (Sanfilippo)	Heparan N-sulfatase	Heparan sulfato	*SGSH*	147
MPS III B (Sanfilippo)	N-acetyl-α-glucosaminidase	Heparan sulfato	*NAGLU*	168
MPS III C (Sanfilippo)	α-Glucosaminida-acetilCoA transferase	Heparan sulfato	*HGSNAT*	70
MPS III D (Sanfilippo)	N-acetylglucosamina-6-sulfatase	Heparan sulfato	*GNS*	25
MPS IVA (Morquio)	N-acetilgalactosamina-6-sulfatase	Sulfato de condroitina e Queratan sulfato	*GALNS*	334

MPS IVB (Morquio)	β-galactosidase ácida	Queratan sulfato	*GLB1*	215
MPS VI (Maroteaux-Lamy)	Arilsulfatase B	Dermatan sulfato	*ARSB*	200
MPS VII (Sly)	β-glucuronidase	Sulfato de condroitina, dermatan sulfato, heparan sulfato	*GUSB*	64
MPS IX (Natowicz)	Hialuronidase	Ácido hialurônico	*HYAL1*	3
MPSPS (Mucopolysaccharidosis-plus syndrome)	desconhecido	desconhecido	*VPS33A*	1
Mucolipidoses				
Mucolipidosis IIα/β, IIIα/β	N-acetil 1-glucosamina fosfotransferase, subunidades α/β	Lipídos e oligossacarídeos	*GNPTAB*	183
Mucolipidosis IIIγ	GlcNAc-1-P transferase	Oligos, GAGs, lipídeos	*GNPTG*	45
Oligossacaridoses e glicoproteinoses				
α-mannosidosis	α-manosidase	α-manosides	*MAN2B1*	144
β-mannosidosis	β-manosidase	β-manosides	*MANBA*	20
Pompe	α-glucosidase ácida	glicogêneo	*GAA*	565
Fucosidoses	Fucosidase	glicolípidos de fucosídeos	*FUCA1*	33
Aspartylglucosaminuria	Aspartylglucosaminidase	aspartil-glucosamina	*AGA*	38
Schindler (Kanzaki)	α-galactosidase B	glicoesfingolipídos, glicoproteínas, oligossacarídeos	*NAGA*	9
Sialidosis	Sialidase, α-neurominidase1	oligossacarídeos	*NEU1*	60
Lipidoses				

Doença de Wolman e doença de armazenamento dos ésteres do colesterol (CESD)	Lipase ácida	Ésteres de colesterol, triglicerídeos	*LIPA*	86
Defeitos em proteínas integrais de membrana e estruturais				
Cistinose	Cistinosina	Cistina	*CTNS*	148
Danon	LAMP2	Glicogênio	*LAMP2*	98
Infantile sialic-acid-storage disease and Salla disease	Sialin	Ácido siálico	*SLC17A5*	53
Niemann-Pick Type C1	NPC1	Colesterol e esfingolipídos	*NPC1*	455
Niemann-Pick Type C2	NPC2	Colesterol e esfingolipídos	*NPC2*	27
Mucolipidosis IV	Mucolipin	Lipídos	*MCOLN1*	35
Outros				
Galactosialidosis	Proteína protetora catepsina A (PPCA)	sialiloligossacarídeos	*CTSA*	36
Deficiência múltipla de sulfatase	Enzima geradora de formilglicina	Sulfatídeos, mucopolissacarídeos	*SUMF1*	53
Lipofuscinose Ceróide Neuronal (Batten) CLN1	palmitoil tioesterase 1	Saposinas	*PPT1*	77
CLN2	tripeptidil-peptidase 1	subunidade c da ATP sintase mitocondrial	*TPP1*	110
CLN3	Battenina	subunidade c da ATP sintase mitocondrial	*CLN3*	71
CLN4	DnaJ homolog subfamily C member 5	subunidade c da ATP sintase mitocondrial	*DNAJC5*	2
CLN5	Proteína neuronal lipofuscinose 5	subunidade c da ATP sintase mitocondrial	*CLN5*	45

CLN6	Proteína neuronal lipofuscinose 6	subunidade c da ATP sintase mitocondrial	*CLN6*	82
CLN7	Proteína 8 Contendo o domínio da superfamília facilitadora maior	desconhecido	*MFSD8*	46
CLN8	Proteína neuronal lipofuscinose 9	subunidade c da ATP sintase mitocondrial	*CLN8*	42
CLN9	desconhecido	desconhecido	-	-
CLN10	Catepsina D	Saposinas	*CTSD*	19
CLN11	Granulinas	-	*GRN*	176
CLN12	Cation-transporting ATPase 13A2	-	*ATP13A2*	38
CLN 13	Catepsina F	-	*CTSF*	11
CLN 14	Proteína KCTD7 contendo domínio BTB / POZ	-	*KCTD7*	17
Cistinose	Transportador de cistina	cistina	*SLC3A1* *SLC7A9*	202 135
Cobalamina F	Transportador de cobalamina	cobalamina	*TCN2*	37
Picnodisostose	Catepsina K	proteínas ósseas	*CTSK*	54
Galactosialidose	Neuraminidase e proteína protetora β-galactosidase	oligossacáridos, ácido siálico	*CTSA*	36

Tabela adaptada de Filocamo et al. (2011)[27]. * HGMD Professional 2018.1, consultado: 7 junho 2018.

Tabela 3. Sistemas analisados no diagnóstico das DLs.

Orgão	Manifestação	Doença a considerar
Fígado	Hepatomegalia	CESD MPS Glicoproteinoses Mucolipidoses II, III Doença de Pompe Doença de Gaucher Niemann-Pick Wolman Disease
Baço	Esplenomegalia	MPS Doença de Gaucher Niemann-Pick
Osso e articulações	Disostose óssea multiplex Articulações inchadas, nódulos dos tecidos moles	MPS Glicoproteinoses Doença de Farber
Músculo-Cardíaco Esquelético	Cardiomegalia, insuficiência cardíaca, miopatia	Doença de Pompe Doença de Danon Doença de armazenamento de glicogênio tipo III e IV
Olho	Mancha macular vermelho cereja	Doença de Tay-Sachs Doença de Sandhoff Niemann-Pick GM1
	Opacidade da córnea	Sialidoses MPS Mucolipidoses II, III
	Cristais na córnea	Cistinoses
Glândula supra-renal	Calcificações adrenais bilaterais	Doença de Wolman
Cérebro	Disfunção mental e motora	Doença de Krabbe Leucodistrofia metacromática NCLs Niemann-Pick Doença de Gaucher MPS Glicoproteinases Doença de Tay-Sachs Doença de Sandhoff GM1

Tabela adaptada de Ferreira et al. (2017)[16].

1.2.2 Diagnóstico das DLs

O diagnóstico definitivo de uma DL requer uma colaboração entre especialistas clínicos e laboratoriais. O diagnóstico clínico é crítico, pois com base na avalição física e

resultados de exames adicionais, é possível direcionar os testes laboratoriais a serem realizados.

A investigação laboratorial tradicionalmente inclui uma análise preliminar (em urina ou plasma) para a detecção específica de macromoléculas não degradadas, seguida pela análise de atividade enzimática em plasma, leucócitos, fibroblastos ou sangue impregnado em papel filtro[23]; a demonstração da atividade deficiente ou ausente por um ensaio enzimático específico é um método eficaz e confiável de diagnóstico[27]. Cerca de 75% das DLs são causadas por uma deficiência na atividade hidrolase lisossomal.

Por outro lado, a presença da atividade normal da enzima lisossomal não pode excluir um diagnóstico específico se for acompanhado por sintomas clínicos sugestivos e/ou pela presença anormal de metabolitos na urina e/ou biopsia do tecido. Podem existir mutações em cofatores da enzima[29-30], por exemplo, mutações no gene *PSAP*, o qual codifica o cofator SapC, necessário para a função da β-glucosidase, no caso da doença de Gaucher[29]. Neste caso, a análise molecular pode complementar o diagnóstico enzimático.

Além disso, existem indivíduos que mostram uma redução significativa da atividade enzimática, mas permanecem clinicamente saudáveis. Esta condição, denominada "pseudodeficiência" enzimática, é conhecida em algumas hidrolases lisossomais (Tabela 4).

Tabela 4. Hidrolases ácidas com pseudodeficiência

Enzima lisossomal	Gene	Doença associada ao gene
arisulfatase A[31]	*ARSA*	Leucodistrofia metacromática
β-hexosaminidase[32]	*HEXA*	Tay-Sachs
α-iduronidase[33]	*IDUA*	MPS I
α-glucosidase[34]	*GAA*	Doença de Pompe
α-galactosidase[35,36]	*GLA*	Doença de Fabry
β-galactosidase[37]	*GLB1*	MPS IVB
		GM1
α-fucosidase[38]	*FUCA1*	Fucosidoses
β-glucuronidase[39,40]	*GUSB*	MPS VII

Embora algumas dessas condições genéticas sejam raras, estima-se que a pseudodeficiência da arisulfatase A tenha uma frequência de 7.3 a 15%[41–43]; e no caso da α-glucosidase, embora esteja presente em menos do 1% dos bebês nos Estados Unidos, a frequência dos alelos da pseudodeficiência pode ser bastante alta em algumas populações (3.9% no Japão e 3.3% em Taiwan)[44,45]. Por tanto é necessário realizar uma combinação de

análises enzimáticas e moleculares para determinar o genótipo dos indivíduos e seus familiares, a fim de diferenciar os indivíduos com pseudodeficiência daqueles portadores ou afetados por uma DL, com a finalidade de tomar uma decisão sobre a introdução de uma determinada terapia.

Por outro lado, o atraso no diagnóstico é frequente, principalmente nas DLs que apresentam formas atenuadas. No caso da doença de Fabry, têm sido descritos atrasos de 14-20 anos em homens e, 14-16 anos em mulheres[46]. Na doença de Pompe de início tardio tem sido descrito atrasos de 5 até quase 9 anos desde o início da doença até o diagnóstico definitivo[47,48]. No Brasil, no ano 2008 foi descrito um atraso de 4.8 anos no diagnóstico das mucopolissacaridoses[49].

1.2.3 Tratamento das DLs

Como as DLs compreendem uma ampla gama de doenças que afetam múltiplos órgãos, o tratamento é necessariamente multidisciplinar. Embora não exista ainda cura para as DLs, nos últimos 25 anos, o foco tem sido o desenvolvimento de terapias que corrijam os efeitos metabólicos destas doenças. Algumas dessas terapias são altamente experimentais, enquanto outras já são aplicadas[50–53].

1.2.3.1 Terapia de Reposição enzimática

Os primeiros estudos da terapia de reposição enzimática (*Enzyme replacement therapy*, TRE) foram realizados no ano 1990, os quais demonstraram a eficiência desta terapia na doença de Gaucher. No período de 2000 até o presente, a TRE também foi utilizada na doença de Fabry, Pompe, MPSI, II, IV, VI e VII. Para outras cinco DLs, a TRE está sendo testada em ensaios clínicos na Europa e nos Estados Unidos[54].

Atualmente no Brasil, o tratamento para sete DLs está aprovado pela Anvisa (Agência Nacional de Vigilância Sanitária) (Tabela 5).

Tabela 5. Status da Terapias de Reposição Enzimática no Brasil

Doença	Nome do produto	Nome genérico	Aprovado pelo Anvisa	Fornecidas pelo SUS
Pompe	Myozyme®	Alfa-alglicosidase	Sim	Não
Fabry	Fabrazyme® Replagal®	Betagalsidase Alfagalsidase	Sim	Não
Gaucher	Cerezyme VPRIV Uplyso™	Imiglucirase Alfavelaglucirase Alfataliglicerase	Sim	Sim
Niemann-Pick tipo B	-	Alfa-Olipudase	Protocolo aprovado para ensaio clínico- 2017	
MPS I	Aldurazyme ®	Alfa-idurinase	Sim	Não
MPS II	Elaprase®	Iduronate-2-sulfatase	Sim	Não
MPS IVA	Vimizin®	Elosulfasa alfa	Sim	Não
MPS VI	Naglazyme®	N-acetylgalactosamine 4-sulfatase	Sim	Não

Apesar do êxito da TRE, a eficácia é variável entre os pacientes. Uma limitação é a biodisponibilidade da enzima recombinante, sendo estas moléculas grandes, não são capazes de difundir livremente através da membrana e muitas vezes não atingem os níveis terapêuticos nos tecidos-alvos, especialmente no SNC. No entanto, existem estratégias em fase pré-clínica para aumentar a biodisponibilidade, por exemplo, a administração intratecal em pacientes com MPSI e VI [55,56].

1.2.3.2 Transplante de células-tronco hematopoéticas

As células-tronco hematopoéticas, derivadas da medula óssea do doador ou do sangue do cordão umbilical, podem ser terapêuticas ao repovoar os tecidos e enzimas lisossomais funcionais no espaço extracelular e no sangue. Estas enzimas funcionais atravessam a barreira hematoencefálica e, nos tecidos, podem ser internalizadas pelas células, corrigindo a atividade deficiente da enzima e, assim, reparando o defeito metabólico. Esta é a única terapia disponível para Krabbe[57] e foi demonstrado que funciona em algumas outras DLs como na MPS I[58] e na Leucodistrofia Metacromática[59], mas é ineficaz em outras como MPS III, enquanto não é mais recomendada para Gaucher[60]. Tanto os ossos como o cérebro são os tecidos menos suscetíveis aos benefícios clínicos deste tipo de terapia. Entre as suas desvantagens está a rejeição do enxerto e infecção.

1.2.3.3 Terapia de Redução de Substrato

Existe ainda pouco progresso com a terapia de redução de substrato (*substrate reduction therapy*, TRS). O miglustate é um agente redutor de substrato aprovado para tratar as doenças de Gaucher tipo I e Niemann-Pick, e alguns outros estão sendo testados em ensaios clínicos[50,53].

1.2.3.4 Terapia Gênica

Como a maioria das DLs são monogênicas, elas são excelentes candidatas à terapia gênica. Atualmente, a Leucodistrofia metacromática é o único exemplo bem-sucedido [57].

1.2.3.5 Chaperonas moleculares

As chaperonas moleculares ou farmacológicas é uma abordagem emergente para o tratamento das DLs mais prevalentes[61] (Fabry, Gaucher tipo I, Pompe, GM1, GM2, MPS IIIC e Batten). São moléculas pequenas capazes de atravessar a barreira hematoencefálica. Essas moléculas interagem com as enzimas mutantes, favorecendo sua conformação correta e aumentado sua estabilidade[62]. Em comparação com outras abordagens, esta mostra vantagens, particularmente em termos de administração oral, ampla biodistribuição, e impacto positivo na qualidade de vida dos pacientes[61]. As chaperonas são mutação-específica[61].

Alguns estudos clínicos mostraram que algumas alterações patológicas associadas às DLs não são reversíveis, mesmo com tratamento a longo prazo. Assim, um diagnóstico molecular preciso e precoce pode ser de grande importância. Um diagnóstico molecular preciso é requisito fundamental para que o paciente tenha acesso a estudos de novas estratégias terapêuticas, como as de pequenas moléculas para TRS ou as chaperonas moleculares[63,64].

1.2.4 Triagem Neonatal das DLs

A Triagem Neonatal (TN) é uma ação preventiva para a identificação de doenças em fase pré-sintomática em todos os recém-nascidos, possibilitando assim o tratamento precoce e, consequentemente, a redução da morbi/mortalidade gerada pelas doenças triadas[65,66].

Introduzida nos anos 60 nos Estados Unidos para a pesquisa de fenilcetonúria, a TN foi se expandindo ao longo dos anos em termos de cobertura geográfica e número de doenças

investigadas, chegando alguns deles a incluir mais de 60 condições. Hoje, a maior parte dos recém-nascidos do mundo são submetidos a algum tipo de triagem neonatal, e dezenas de milhares tem a história natural das suas doenças alterada positivamente[65].

Na TN são coletadas algumas gotas de sangue do calcanhar do recém-nascido (entre o terceiro e quinto dia de vida), as quais são impregnadas em papel-filtro (SIPF), após a coleta as amostras são encaminhadas para os Serviços de Referência em Triagem Neonatal para a execução dos diferentes testes de triagem. No caso de um resultado alterado, a família é contatada, e uma segunda amostra (soro, sangue total, urina ou sangue impregnado em papel filtro) é coletada para a realização de testes confirmatórios, diferenciando os resultados positivos dos falsos-positivos[67,68].

A possibilidade de tratamento específico para as doenças lisossômicas, iniciada com o TRE para a doença de Gaucher nos anos 90 e, ampliada para outras diversas condições ao longo dos anos seguintes, junto com os benefícios do tratamento precoce, trouxeram para discussão a possibilidade de incluir DLs selecionadas nos programas de TN[69]. No mundo vem se realizando diversos esforços, através da realização de estudos pilotos que determinem a viabilidade e necessidade da inclusão destas doenças nos programas de TN. Entre os países estão: Áustria[70], Taiwan[71–73], Itália[74,75], Hungria[76], e os Estados Unidos[77–79]. Nos Estados Unidos, tanto as doenças de Pompe, MPS I, Krabbe, Niemann-Pick A/B e a Fabry foram recomendadas para inclusão nos programas de TN[80], sendo que Pompe e MPS I já tiveram a recomendação aprovada (*Advisory Committee on Heritable Disorders in Newborns and Children*). O estado de Nova Iorque faz a TN para a doença de Krabbe desde o ano 2006[81].

Não existem programas de TN para as DLs na América Latina. No entanto, no Brasil foi implementado um programa piloto de TN, inicialmente para uma comunidade com alto risco de incidência de MPS VI, o que vem permitindo o atendimento precoce dos indivíduos afetados e o aconselhamento genético[69,82]. Entre outras iniciativas, encontram-se um estudo piloto para a triagem de 04 DLs (MPS I, Pompe Gaucher e Fabry) em 10,527 recém-nascidos, realizada no Sul do país, e um outro estudo piloto em desenvolvimento para a triagem de DLs tratáveis, realizado no SGM-HCPA.

Até o momento, os programas de TN de DLs se basearam em métodos bioquímicos de triagem simultânea de diferentes condições, seguidos da confirmação pela análise

molecular, usualmente pelo sequenciamento convencional (método de Sanger) do gene relacionado ao defeito enzimático [78].

1.3 ANÁLISE MOLECULAR DAS DLs

1.3.1 Espectro mutacional das DLs

Na maioria de DLs, grande parte das mutações são *de novo*, privadas e é observada uma alta frequência de heterozigosidade composta, além da presença de vários polimorfismos [83–87]. Desta maneira, a análise molecular de mutações específicas não é uma solução, uma vez que, na maioria dos casos, não existem mutações comuns nos genes associados às DLs (Tabela 6)[88].

É observado um amplo espectro mutacional nas DLs (Tabela 6). Entretanto, é importante destacar que em algumas populações, como nos judeus Ashkenazi, existem mutações recorrentes para alguns distúrbios como a doença de Gaucher Tipo 1. Nesta população, aproximadamente 90% das mutações podem ser detectadas através do *screening* das quatro mutações comuns: N370S, L444P, IVS2* e 84insG. No resto da população não judaica, estas quatro mutações contabilizam para o 50-60% dos alelos mutados sendo que o restante corresponde a mutações privadas. Da mesma maneira, quatro mutações são as responsáveis por mais de 95% dos alelos mutantes em pacientes judeus Ashkenazi com Niemann-Pick tipo A/B[84].

Tabela 6. Revisão das mutações encontradas na literatura nos genes incluídos no presente estudo.

Painel	Doença	Gene	Exons	Cr.	Mutações descritas*	Tipo de mutação (%)					
						M/N	S	R	PID	GID	RC
A	Fabry	*GLA*	07	Xq22	922	70.1	4.7	0.5	19.3	4.7	0.7
	Pompe	*GAA*	20	17q25.2-q25.3	565	59.8	12.7	0.2	23.0	3.4	0.9
	Danon	*LAMP2*	10	Xq24	98	31.6	19.4	2.0	36.7	10.3	-
	Schindler	*NAGA*	09	22q11	9	88.9	-	-	11.1	-	-
B	Niemann-Pick tipo C1	*NPC1*	25	18q11.2	455	67.0	8.2	-	22.6	2.2	-
	Niemann-Pick tipo C2	*NPC2*	05	14q24.3	27	70.4	14.8	-	11.1	3.7	-
	Gaucher	*GBA1*	12	1q21	472	76.7	5.3	0.2	11.9	1.5	4.4
	Niemann-Pick tipo A/B	*SMPD1*	06	11p15.4-q15.1	251	73.3	1.6	-	24.7	0.4	-
	Def. lipase ácida	*LIPA*	10	10q23.2-q23.3	86	57.0	16.3	1.2	18.5	5.8	1.2
	Def. quitotriosidase	*CHIT1*	11	1q31-q32	8	50.0	-	-	12.5	25.0	12.5
	Def. prosaposina	*PSAP*	15	10q21-q22	26	57.7	19.2	-	15.4	7.7	-
C	CLN1	*PPT1*	09	1p32	77	62.3	18.2	-	16.9	2.6	-
	CLN2	*TPP1*	13	11p15	110	59.1	16.4	-	23.6	0.9	-
	CLN3	*CLN3*	14	16p12.1	71	49.3	18.3	-	15.5	16.9	-
	CLN4	*DNAJC5*	05	20q13.33	2	50.0	-	-	50.0	-	-
	CLN5	*CLN5*	04	13q21.1-q32	45	64.4	2.2	-	28.9	4.5	-
	CLN6	*CLN6*	07	15q23	82	69.5	4.9	-	23.2	2.4	-
	CLN7	*MFSD8*	13	4q28.1-q28.2	46	71.7	13.0	-	15.3	-	-
	CLN8	*CLN8*	03	8p23	42	73.8	-	-	14.3	11.9	-
	CLN10	*CTSD*	09	11p15.5	19	73.7	5.3	-	21	-	-
	CLN11	*GRN*	13	17q21.32	176	42.6	11.9	3.5	36.9	5.1	-
	CLN12	*ATP13A2*	29	1p36	38	65.8	7.9	-	21.1	5.2	-
	CLN 13	*CTSF*	13	11q13	11	81.8	9.1	-	9.1	-	-
	CLN 14	*KCTD7*	04	7q11.21	17	64.7	5.9	-	23.5	5.9	-

Cr: cromossomo, M/N: *missense/ nonsense*, S: sítio de *splicing*, R: regulatória, PID: pequenas indels, GID: grandes indels, RC: rearranjos complexos. * HGMD Professional 2018.1, consultado 7 junho 2018.

1.3.2 Pseudogenes

Pseudogenes são sequencias com alta homologia aos genes funcionais, mas que não produzem uma proteína funcional[89]. A presença destes pode levar a uma genotipagem errada de um paciente, razão pela qual são necessárias técnicas adicionais para diferenciar o gene funcional do pseudogene e realizar uma chamada de variantes correta. Por exemplo, no caso de pacientes MPS II é necessária uma análise adicional para a detecção de rearranjos gene-pseudogene que não são detectados pelos testes convencionais[90]. Na doença de Gaucher, também é necessária a análise de deleções parciais/totais do gene e rearranjos gene-pseudogenes que podem não ser detectadas nas análises de mutações de rotina[91,92].

O gene *GBA1*, associado à doença de Gaucher, compreende 11 exons e 10 introns ao longo de 7.6kb. Está localizado no cromossomo 1q21 dentro de um locus complexo que contém sete genes e dois pseudogenes, provavelmente devido a um evento de duplicação desta região cromossômica. De fato, um pseudogene de 5.7kb altamente homólogo (*GBAP1*) está localizado aproximadamente há 16kb *downstream* do gene funcional *GBA1*[93]. A região exônica do *GBAP1* compartilha 96% de homologia com a região codificante do gene *GBA1*, enquanto a homologia alcança 98% na região entre o intron 8 e a UTR_3´. *GBAP1* é menor do que *GBA1*, devido à falta de várias inserções *Alu*, regiões intrônicas e com uma deleção de 55 pares de bases no intron 9[94,95]. A alta homologia e a proximidade física entre *GBAP1* e *GBA1* permite eventos de recombinação resultando na geração dos denominado "alelos complexos" ou alelos "Rec"[96,97]. Estes alelos "Rec" são gerados tanto por eventos de *crossovers* ou por conversão gênica entre o gene *GBA1* e seu pseudogene. Todos esses alelos "Rec" têm sido encontrados em heterozigose, sugerindo letalidade em homozigose[98]. Os poucos casos de homozigotos descritos na literatura, correspondem a abortos ou mortes neonatais[99–101].

Dentre as 472 mutações descritas no gene *GBA1*, existem mais de 20 alelos "Rec", sendo o RecNcil e Recdelta55 as mais frequentes, sozinho ou em combinação com outra mutação pontual[102,103]. O alelo RecNcil é produto de uma recombinação do intron 9 ao exon 10, resultando na incorporação de um segmento do *GBAP1* que inclui as mutações p.Leu483Pro (L444P), p.Ala495Pro (A495P) e p.Val499Val (V460V) no gene funcional *GBA1*[104]. Recdelta55 engloba uma deleção de 55pb no exon 9 do gene GBA, correspondente à porção do pseudogene[105].

1.3.3 Testes genéticos moleculares para diagnóstico das DLs

Apesar dos testes bioquímicos fornecerem um diagnóstico confiável para a maioria de DLs, a identificação de variantes patogênicas, através da análise genética, está sendo cada vez mais recomendada devido às informações adicionais que proporciona.

Um teste genético molecular pode confirmar o diagnóstico enzimático de uma DL, e é essencial para o diagnóstico definitivo de uma DL que é produto da deficiência de uma proteína lisossomal não enzimática, por exemplo, na doença de Danon e várias lipofuscinoses ceróides neuronais (Tabela 2). A confirmação do diagnóstico é crucial para avaliar as opções de tratamento, como a TRE disponível para algumas DLs, e para tratamentos alternativos como as chaperonas farmacológicas ou TRS, para as quais estão em andamento alguns estudos clínicos[20,106].

Estes testes também permitem distinguir casos de pseudodeficiência[32–34,36–41], a identificação de membros da família em risco e podem esclarecer o tipo de variação genética e seu impacto na proteína e na atividade residual da enzima, ajudando no estabelecimento de uma correlação genótipo-fenótipo para algumas DLs[27].

Algumas técnicas convencionais de análises de mutações concentram-se principalmente na identificação de uma mutação por vez. A mais usada por vários grupos de pesquisa e laboratórios clínicos é o sequenciamento de Sanger, o qual é considerado o método padrão ouro para o sequenciamento de ADN[107–111], mas devido à alta heterogeneidade alélica e ao fato de que esta metodologia só pode analisar um segmento de ADN/exon de cada vez, se torna um processo de trabalho intensivo, demorado e caro. Atualmente, o sequenciamento de Sanger é geralmente o método escolhido para investigar indivíduos de uma família com uma mutação conhecida num gene específico, e assim identificar mais pacientes, portadores ou para diagnóstico pré-natal. Em alguns centros, o sequenciamento de Sanger também é aplicado para preencher eventuais *gaps* (regiões com pouca ou sem cobertura) que a abordagem baseada no NGS pode deixar[112].

Em geral, a análise de sequência tem potencial de detectar ambas variantes patogênicas em uma grande porcentagem dos probandos (Tabela 6). Para a detecção de outras variantes, além das variantes *missense*, *nonsense*, de sítio de *splicing*, e pequenos indels, existem outros métodos, por exemplo, para análise de deleções/duplicações (PCR quantitativo, PCR *long-range*, MLPA, microarranjos desenhados para detectar deleções de exon único ou

duplicações) e, para a análise de rearranjos complexos (sequenciamento, análise de SNPs, análise cromossômica por microarranjos).

Entre outros métodos convencionais para análise genética utilizados, encontram-se o RFLP (*restriction fragment length polymorphis*m)[113–117], ARMS (*amplification-refractory mutation system*)[118–120], SSCP (*single strand conformation polymorphism*)[121–123], dHPLC (*denaturing high performance liquid chromatography*)[124,125], RT-PCR (*real-time polymerase chain reaction*)[126], HRM (*high resolution melting*)[127] e MLPA (*multiplex ligation-dependent probe amplification*)[128–135], todos eles utilizados na análise molecular de diferentes DLs. Também existem novas abordagens como o a-CGH (*array comparative genomic hybridization*)[114,136–138] e o sequenciamento de nova geração (*next-generation sequencing*). Algumas indicações e limitações gerais das diferentes técnicas moleculares recomendadas para o diagnóstico das DLs encontram-se em detalhe na Tabela 7.

Tabela 7. Indicações e limitações das diferentes técnicas moleculares utilizadas para o diagnóstico das DLs.

Método	Indicação	Limitação
Sequenciamento de Sanger	Para detectar variantes de ADN localizadas em regiões codificantes do gene associado e regiões adjacentes da junção exon/intron, incluindo mutações tipo SNPs, pequenas indels e variantes de sítios de *splicing*.	Não detecta Variação no Número de Cópias (C*opy Number Variations,* CNVs).
RT-PCR (Taqman)	Adequado para o *screening* de mutações pontuais específicas.	Requer optimização para cada locus analisado. Novas variantes não descritas não são detectadas.
HRM	Determinação do status de metilação do ADN e genotipagem	Não detecta CNVs
MLPA	Para estabelecer a presença de CNVs.	Necessário confirmar resultados com outras técnicas (sequenciamento de Sanger, tecnologias NGS e técnicas de a-CGH, entre outras) ou com sondas diferentes do probemix, devido à interferência de polimorfismos na região de alinhamentos dos *probes*.
CGH e a-CGH	Para detectar grandes deleções e duplicações, rearranjos não balanceados e limites de deleção que não são detectados por análise de sequencia	Não detecta rearranjos balanceados.
NGS (painéis)	Detecção de mutações pontuais e pequenas indels em muitos genes diferentes ao mesmo tempo e de uma maneira eficiente em termos de custo e tempo.	Análise de pseudogenes ou regiões altamente homólogas (ex. MPSII e Gaucher) O sequenciamento de Sanger é necessário para atender aos padrões de cobertura e para confirmação e variantes em regiões homopoliméricas. Não são detectadas translocações e inversões.

1.4 SEQUENCIAMENTO DE NOVA GERAÇÃO (NGS)

O sequenciamento de nova geração (*next-generation sequencing*, NGS) é uma tecnologia de sequenciamento de ADN de alto rendimento que permite a geração rápida de dados de milhares a milhões de pares de bases do ADN de um paciente através do sequenciamento de vários genes em uma única reação[139]. O conceito por trás da tecnologia NGS é semelhante ao sequenciamento tradicional por eletroforese capilar. A diferença está no fato de que, em vez de sequenciar um único fragmento de ADN, o NGS pode sequenciar milhões de fragmentos de forma massiva e paralela.

As diferentes abordagens do NGS permitem o sequenciamento rápido de todo o genoma (*whole-genome sequencing*, WGS) ou partes dele (via *whole-exome sequencing*- WES ou painéis de genes) (Figura 1).

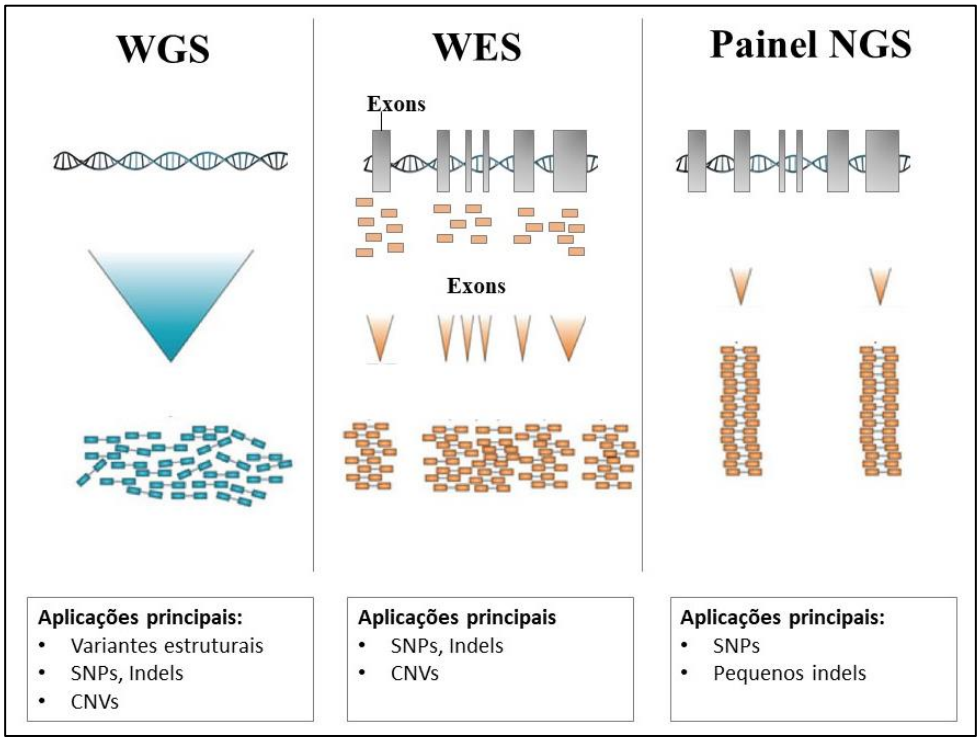

Figura 1. Comparação das regiões alvo sequenciadas no WGS, WES e Painel NGS e suas aplicações principais.

As diferentes abordagens NGS têm vantagens sobre o sequenciamento de Sanger, como a grande profundidade de cobertura (vezes que uma base é lida) e uma alta sensibilidade. Na Figura 2 (superior), encontra-se a visualização parcial das leituras geradas pelo NGS, no as barras cinzas verticais indicam a profundidade de cobertura e as barras cinzas horizontais o alinhamento de cada leitura. Todas as posições, na cor cinza, indicam que a base na leitura é igual à da referência, qualquer troca é indicada como um retângulo colorido. A figura 2 representa, na parte superior, uma mutação de ponto em heterozigese (G/C) com profundidade de cobertura de 295 vezes. Na parte inferior da é observado parte do eletroferograma gerado no sequenciamento automatizado (Sanger), indicando a presença da mesma mutação em heterozigose como 2 picos sobrepostos (1 pico correspondendo à base G em preto e outro à base C em azul).

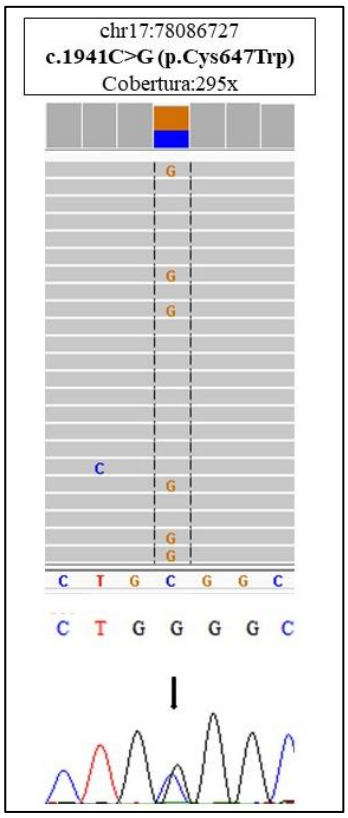

Figura 2. Comparação da profundidade de cobertura entre o NGS (parte superior, como visualizado no software IGV) e o Sequenciamento de Sanger (parte inferior) de uma mutação pontual.

1.4.1 Desenho de painéis NGS

É recomendado que os painéis NGS contenham genes com evidência científica suficiente de um papel causal da doença[141], pois variantes em genes que ainda não estão estabelecidos como causadores da doença são difíceis de interpretar e podem levar a resultados inconclusivos. O *Clinical Genome Resource* ou ClinGen (financiado pelo NIH) tem realizado uma revisão abrangente baseada em evidências das associações gene-doença, classificando-a em "definitiva", "forte", "moderada", "limitada", "nenhuma evidência declarada" ou "evidências conflitantes"[142].

Concentrar-se em um conjunto limitado de genes permite obter uma maior profundidade de cobertura, levando a uma sensibilidade e especificidade analítica superior. Uma maior profundidade de cobertura aumenta a confiança da chamada de heterozigotos[141]. As regiões que apresentam baixa cobertura (por exemplo, regiões ricas em GC ou regiões repetitivas), podem ser completadas utilizando o sequenciamento de Sanger ou outra metodologia, o que melhoraria a sensibilidade clínica do ensaio. Incluir menos genes no desenho também permite que o laboratório utilize sequenciadores menores e que sejam sequenciadas mais amostras por "corrida" em comparação com o WES ou WGS. A quantidade de dados gerados e requisitos de armazenamento também são mais manejável quando se trabalha com painéis gênicos[141].

Após a seleção dos genes que serão incluídos no painel, é importante realizar uma revisão completa das variantes descritas em cada um deles com a finalidade de identificar variantes patogénicas comuns ou *hotspots* e variantes patogênicas localizadas fora de regiões exónicas tipicamente cobertas, como regiões intrônicas profundas ou não traduzidas. Esta informação é essencial para determinar a região genômica alvo de interesse e poder ser importante para selecionar as amostras para o processo de validação. Também é importante para determinar a sensibilidade e especificidade do teste.

O próximo passo é determinar a região genômica de interesse. Geralmente são selecionadas as regiões codificantes e junções exon-intron. Não existe consenso sobre o número de bases da sequência intrônica que deve ser incluída na análise, embora a maioria de laboratórios inclua sequencias de ±10-20 bases da junção exon-intron para a detecção de mutações de sítios de splicing[143].

Informações sobre a doença, incluindo os principais indicadores clínicos, mecanismos de doença, prevalência, modo de herança, penetrância e expressividade, devem

ser investigadas no estágio de planejamento do teste. Todos esses fatores desempenham um papel crítico na interpretação de resultados e na elaboração do relatório final.

1.4.2 Ion Torrent Personal Genome Machine

A plataforma Ion Personal Genome Machine™ (PGM) (Thermo Fisher Scientific), foi lançada no ano 2011 e está catalogada como um sequenciador de segunda geração. Esta plataforma detecta os prótons que são liberados a medida que os nucleotídeos são incorporados na fita que está sendo sintetizada[144].

Ao comparar as especificações técnicas desta plataforma com outras disponíveis no mercado, podemos observar que o Ion PGM é a plataforma NGS de menor custo, proporciona diferentes rendimentos de sequenciamento dependendo do Ion chip utilizado (até 1Gb utilizando o Ion chip 318) e o tempo de corrida é um dos menores que existe (2 horas). Embora os dados gerados utilizando essa plataforma tenham uma taxa de erro (~1.8%) superior ao da Illumina (<0.4%), desde que haja profundidade de cobertura suficiente, a capacidade de chamada de variantes é bastante semelhante às das outras tecnologias[144].

Uma das limitações que esta plataforma apresenta é a detecção de homopolímeros (sequencias contínuas de bases iguais). Nesse caso, todas as bases do homopolímero vão ser incorporadas em um único fluxo de dNTP´s. O sensor acoplado à cada poço do Ion chip tem uma resposta bastante linear; por tanto, se um A tem um sinal x, um AA vai ter um sinal aproximado de 2x, e assim por diante. Na prática, é possível detectar com boa acurácia homopolímeros de até 6 bases. Além disso, o Ion Torrent possui um mapeador (TMAP) otimizado também para lidar com erros de homopolímeros[145].

Independentemente da plataforma utilizada, é possível identificar etapas comuns entre todos os sequenciadores: preparo da amostra (biblioteca), enriquecimento da biblioteca e sequenciamento, processamento de dados brutos, seguido da análise e interpretação dos dados gerados (Figura 3).

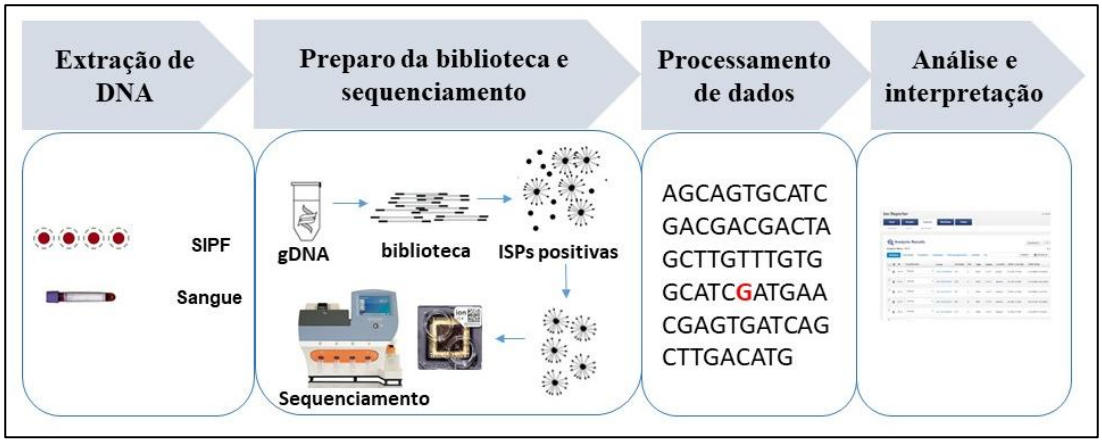

Figura 3. Fluxo de trabalho geral do NGS.

1.4.3 Preparo da amostra

O processo começa com a extração do ADN genômico de uma amostra do paciente. O NGS pode ser realizado a partir do ADN extraído de qualquer tipo de amostra, desde que a quantidade e qualidade sejam suficientes.

1.4.4 Preparo da biblioteca

Primeiro, o ADN é amplificado utilizando o painel NGS (pool de primers para a amplificação dos genes incluídos no painel) através de um PCR *ultrahigh-multiplex*. Desta maneira são gerados milhares de *templates*, de forma que todas nossas sequencias alvos sejam cobertas o mais uniforme possível. Uma vantagem é que é necessário relativamente pouco ADN do paciente nessa etapa. Trabalhando com o Ion PGM, é necessário apenas 20ng de ADN total.

Após a amplificação, adaptadores, sequencias artificias conhecidas, são incorporadas ao *template*. Nesta etapa é possível combinar diferentes amostras de uma mesma reação de sequenciamento através do uso de *barcodes* diferentes, normalmente de 5 ou 10 bases, reduzindo assim, o custo de processamento. O número de amostras que podem ser sequenciadas em simultâneo dependerá da profundidade de cobertura desejada[141]. Depois de ligados esses adaptadores-*barcodes*, as amostras são misturadas, amplificadas e sequenciadas juntas. Após isso, no processo de sequenciamento, essa parte do adaptador-*barcode* é lida e as amostras são separadas computacionalmente[145].

1.4.5 Enriquecimento da biblioteca

O enriquecimento da biblioteca tem como objetivo gerar em um pequeno espaço físico milhares de cópias de cada *template* de ADN produzido na etapa anterior. O objetivo dessa amplificação é aumentar a fonte de sinal iônico para o sequenciador Ion PGM, que será detectado na etapa de sequenciamento.

Um dos processos de amplificação desenvolvidos para esse propósito é o PCR em emulsão. Nele são criados milhões de micro reatores em uma emulsão de óleo. Esses reatores contêm todos os reagentes necessários para uma reação de PCR e pequenas esferas (também chamadas de *Ion Sphere Particle*-ISP) cobertas com a sequência complementar ao adaptador e com um diâmetro de 3-µm. Esses fragmentos servem para fixar os clones do *template* à esfera e também como primer para a reação de PCR. Ao final do processo a esfera pode estar em três estados: 1) caso ideal: um único *template* foi incorporado a uma única esfera; 2) ISP policlonais: múltiplos fragmentos foram incorporados a uma esfera e, 3) ISP vazias: nenhum fragmento foi à esfera[145]. Tanto as ISPs vazias como as policlonais são eliminadas através de diferentes processos.

1.4.6 Sequenciamento

As ISPs são carregadas num Ion chip que contém milhões de poços (3.5µm de diâmetro) com sensores de prótons e, o sequenciamento é iniciado a partir da sequência do adaptador. Atualmente são oferecidos 3 tipos diferentes de chips para sequenciamento: Ion 314^{TM} Chip v2 com rendimento de até 100 Mb, Ion 316^{TM} Chip v2 com rendimento de até 500Mb e, o Ion 318^{TM} Chip v2 com rendimento de até 1Gb. O tamanho das leituras pode alcançar 400pb, com comprimento médio de aproximadamente 200pb.

Cada esfera é carregada num poço com um sensor individual pela ação da centrifugação do chip utilizando uma centrífuga especial. Durante o sequenciamento, os quatro nucleotídeos são fornecidos de forma gradual durante a corrida. Quando o nucleotídeo fornecido é complementar à base, este é incorporado na fita pela ação de uma polimerase. Isso aumenta o comprimento da fita que está sendo sintetizada em uma base (ou mais, se trata-se de um homopolímero) e resulta na hidrólise do nucleotídeo trifosfato, que causa a liberação de um único próton (H^+) por cada nucleotídeo incorporado, alterando o pH do poço (ΔpH). Esta alteração do pH é detectada por um sensor (*ion-sensitive field-effect transitor*, ISFET) e convertida em um sinal elétrico durante o todo o processo de sequenciamento[144]. Após o fluxo de cada nucleotídeo, é realizada uma lavagem para garantir

que estes não permaneçam no poço. A medida que o sequenciamento prossegue, cada uma das quatro bases é introduzida sequencialmente. Se as bases desse tipo forem incorporadas, os prótons são liberados e um sinal é detectado proporcionalmente ao número de bases incorporadas[144]. Esses dados de sinal bruto são transformados em chamadas de bases, que é a identificação do nucleotídeo presente em cada posição em uma leitura. Isto é realizado utilizando o software Torrent Suite v5.0 (Thermo Fisher Scientific).

O principal tipo de erro desta plataforma são as inserções e deleções pois, em regiões homopoliméricas maiores de 6 bases, não existe linearidade entre a intensidade do fluxo de íons hidrogênios detectados e o número de nucleotídeos incorporados, fazendo com que erros na determinação do tamanho de tais regiões sejam frequentes[146].

1.4.7 Análise e interpretação de dados

O NGS produz uma quantidade extensa de dados de sequência que normalmente é processada e analisada em três etapas principias (Figura 4).

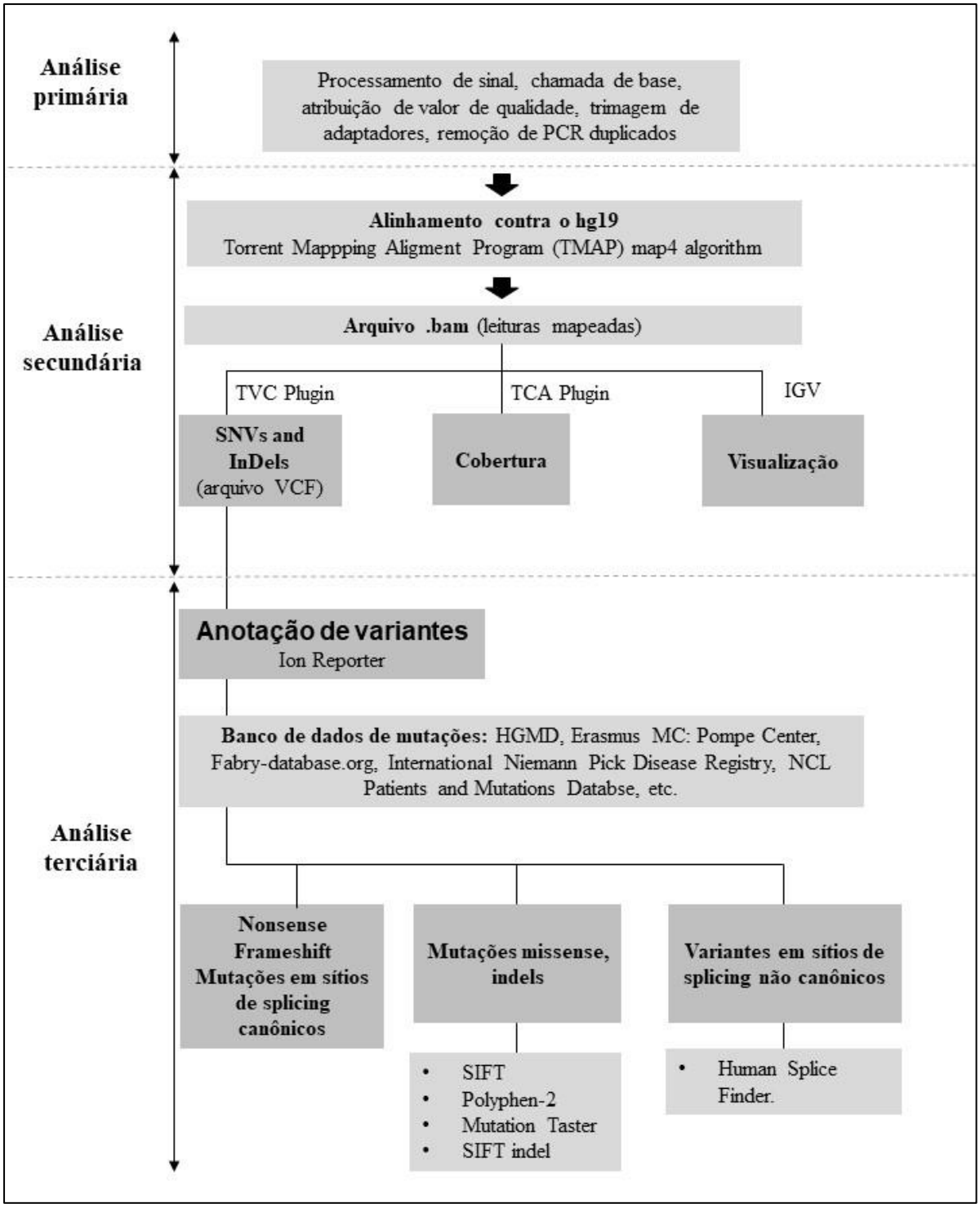

Figura 4. Fluxograma da análise de dados gerados pelo NGS.

O primeiro passo, executado pelo software do instrumento (Torrent Suite Software™), é o processamento dos sinais em sequencias com valores de qualidade associadas a cada base. Também é realizada a demultiplexação das amostras. Essas

informações são compiladas em um formato de .fastaq[147], que é o arquivo de entrada para a análise secundária durante a qual as leituras de sequência são alinhadas a uma sequência referência (hg19, Genome Reference Consortium GRCh37). As leituras alinhadas são compiladas em um formato de arquivo denominado .bam. As principais informações no arquivo .bam incluem a localização do alinhamento de leitura em relação à referência, a qualidade do mapeamento de leitura, a profundidade de cobertura por base mapeada e a orientação das leituras (*forward* ou *reverse*) quando o sequenciamento é bidirecional[143]. Os arquivos .bam podem ser visualizados no programa IGV (*Integrative Genomics Viewer* do Broad Institute).

A terceira etapa utiliza o arquivo .bam como entrada no software (Ion Reporter™) que determina as diferenças entre as leituras alinhadas e a sequencia de referência e compila essas diferenças em um formato de arquivo de chamada de variantes (*variant caller file*, VCF). O arquivo VCF (*Variant Call Format*) é um arquivo de texto que contém informações relacionadas a: cromossomo, posição da variante iniciando a contagem por 1, bases de referência, bases alteradas, score da variante em escala phred, se a variante passou pelos controles de qualidade, e informação com estatísticas sobre as variantes. A precisão da chamada de variantes depende da profundidade de cobertura e melhora com o aumento da cobertura[141]. Esta última etapa também inclui a anotação de variante, na qual é adicionada informação sobre cada variante detectada (por exemplo, atribuição da nomenclatura c. e p., localização exonica ou intronica) e a associação da variante com meta-dados (por exemplo, frequencia populacional)[143,145,148].

Cada etapa é complexa e para realiza-las requer uma combinação de algoritmos e softwares. As escolhas desses são influenciadas pela química do sequenciamento e da plataforma, o tipo de variantes que serão detectadas (por exemplo, SNVs CNVs) e a experiência bioinformática do laboratório.

A interpretação da variante é geralmente realizada usando a informação de frequência nas diferentes bases de dados, análise de segregação, banco de dados de mutações, literatura e programas de predição *in silico*.

É recomendado que as variantes sejam interpretadas seguindo as diretrizes de classificação de variantes publicadas recentemente[149]. A *American College of Medical Genetics and Genomics* (ACMG) publicou pela primeira vez recomendações para a interpretação de variantes no ano 2005, depois no 2008[150], com a revisão mais recente de

2015[149], introduzindo um novo sistema de classificação: “patogênica”, “provavelmente patogênica”, “de significado incerto”, “provavelmente benigna” e “benigna”, com base em critérios usando determinadas evidências, por exemplo, dados populacionais, dados computacionais, dados funcionais, dados de segregação, etc).

As variantes com uma frequência populacional alta (geralmente maior que 1%, no caso de doenças raras) são frequentemente filtradas. Para relatar uma variante, é importante determinar se o efeito da variante é consistente com o fenótipo do paciente e também examinar a segregação da variante na família do probando (quando for possível). Para frequências populacionais, existem dados disponíveis gratuitamente para seu uso em pesquisa e diagnóstico, como os do projeto *1000 Genomes*[151] e *Exome Aggregation Consortium* (ExAC)[152], dbSNP[153], Online Archive of Brazilian Mutations[154], Human Gene Mutation Database[88], e outras bases de dados específicas de doenças como Pompe Center at Erasmus Medical Center (http://www.pompecenter.nl/), Fabry-database.org (http://fabry-database.org/), International Niemann-Pick Rare Disease Registry (https: // inpdr .org /) e o NCL Mutation and Patient Database (https://www.ucl.ac.uk/ncl/mutation.shtml).

A avaliação da patogenicidade de novas variantes (não encontradas nas bases de dados de mutações ou descritas anteriormente na literatura) são analisadas através de ferramentas *in silico*. Para a predição do potencial efeito deletério na função da proteína, são utilizados os programas SIFT[155,156], Polyphen-2[157] e Mutation Taster[158]. Para avaliar o possível efeito de uma variante de sítio de splicing, pode ser utilizado o Human Splice Finder[159]. As inserções e deleções, podem ser analisadas através do Mutation Taster e SIFT Indel[160]. Mutações *nonsense*, *frameshift* e de sítios canônicos de splicing geralmente são classificadas automaticamente como patogênicas[149].

Estudos funcionais são úteis para determinar se a variante afeta a função da expressão normal. No entanto, esses estudos podem ser difíceis de interpretar porque não existem sistemas modelo perfeitos e os resultados podem ser contraditórios entre as diferentes análises[143].

1.4.8 Limitações do NGS

Durante o desenho do painel NGS, é importante considerar as limitações do NGS. Muitas dessas limitações são inerentes a todas as tecnologias. Entre as limitações técnicas temos:

1) Interferência de sequencias homólogas: Alguns genes ou partes de genes, podem não ser sequenciados adequadamente, o que não permite total confiança na qualidade dos dados. Isto inclui genes com sequencias complexas como pseudogenes, rearranjos genéticos e regiões com alto conteúdo de GC. Regiões de alta homologia com outras regiões genômicas, como pseudogenes ou eventos de duplicação gênica, podem levar a resultados falsos-positivos e/ou falsos negativos devido as leituras mapeadas incorretamente[141]. Chamada de variantes em regiões altamente homólogas que não podem ser detectadas com precisão pelo NGS frequentemente podem ser resolvidas por outros métodos, como o sequenciamento de Sanger, através do desenho de primers específicos, ou por outros métodos, como os bioinformáticos, através do realinhamento digirido[141,143].

Por exemplo, além da análise NGS do gene *GBA*, associado à doença de Gaucher, alguns investigadores analisam pelo menos a região 3´ do gene por sequenciamento de Sanger[161,162] ou apenas os alelos "Rec" mais frequentes[163]. Muitos estudos que utilizaram a tecnologia NGS sem o apoio do Sanger, não descreveram a presença de alelos "Rec", desestimando-os. De fato, apesar dos avanços da tecnologia de sequenciamento, o sequenciamento de Sanger continua sendo o método padrão ouro para analisar o gene *GBA*, a não ser em grandes grupos de pacientes[164–168]. Outra abordagem para resolver o problema dos alelos "Rec", em especial do alelo RecNciI é o desenho de primers NGS específicos seguido do realinhamento dirigido das leituras com a sequência do gene *GBA* e não contra genoma completo[169].

2) Duplicações e deleções de exons: Deleções e duplicações em nível de exon podem ser detectadas por NGS usando algumas ferramentas bioinformáticas disponíveis comercialmente; no entanto, a sensibilidade e especificidade devem ser estabelecidas pelo laboratório[170].

3) Regiões repetitivas: regiões homopoliméricas e expansão de trinucleotídeos não são detectadas pelo NGS e podem ser analisadas usando outros métodos[171].

1.4.9 Confirmação das variantes detectadas pelo NGS

Dentro da comunidade científica se discute sobre a necessidade de confirmar as variantes chamadas no NGS utilizando o sequenciamento de Sanger, isto devido a que vários laboratórios têm descrito uma alta especificidade para os dados gerados pelo NGS. A confirmação por Sanger dos resultados do NGS pode ser ineficiente, redundante e cara[172].

O grupo de Nelson et al.[173] determinou que 80% dos dados gerados poderian ser liberados sem confirmação pelo método de Sanger, isto reduziria o tempo de resposta e

manteria a especificidade. Mu et al[174], relataram 98.7% de concordância entre o NGS e o Sanger, 1.3% foram identificados como falsos-positivos, localizados principalmente em regiões genômicas complexas (regiões ricas em AT, regiões ricas em GC, regiões homopoliméricas e regiões de pseudogenes) e para os quais é indispensável a confirmação por Sanger.

O grupo de Nelson et al, também estabeleceu certos critérios para descrever uma variante genética detectada pelo NGS sem confirmação pelo método de Sanger, entre eles: 1) se é uma substituição de nucleotídeo, 2) "PASS" no filtro, 3) score de qualidade adequada, 4) frequência alélica da variante (heterozigoto = 0.3-0.6, homozigoto/hemizigoto= >0.9). É recomendado a confirmação da variante se: 1) inserções/deleções (devido a inconsistências no mapeamento), 2) score de qualidade baixo e, 3) frequência alélica da variante fora dos intervalos estabelecidos.

1.5 VALIDAÇÃO ANALÍTICA

Uma vez que a metodologia foi estabelecida e o protocolo otimizado no laboratório, o teste deve ser validado, incluindo todas as etapas do processo (tanto o trabalho de bancada como a parte bioinformática) usando todos os tipos de amostra que serão aceitos para o painel (por exemplo: sangue total, sangue impregnado em papel filtro, saliva, tecido embebido em parafina, *swab* bucal, amniócitos e vilosidades coriônicas) (Figura 5). É exigido que os laboratórios determinem as características de desempenho do ensaio, incluindo 3 medidas: sensibilidade analítica, especificidade analítica e precisão ou reprodutibilidade[141,175,176]. Estas 3 medidas são determinadas testando amostras provenientes de indivíduos com variantes de sequência conhecidas e controles negativos conhecidos. Para a validação de um painel NGS, não é viável identificar e analisar controles para cada uma das mutações descritas nos genes; por tanto é realizada uma validação baseada nos métodos. Esta abordagem de validação baseada nos métodos, utiliza amostras com mutações conhecidas, particularmente direcionada para mutações comuns e tipos de variantes ou regiões genômicas que podem ser mais difíceis de detectar, como indels, regiões ricas em GC e regiões repetitivas[143].

A sensibilidade analítica é a probabilidade de o ensaio detectar uma variante quando está presente na região alvo (1- taxa de falsos negativos). Isso é determinado dividindo o número de variantes conhecidas (positivos verdadeiros) detectados pelo painel NGS pelo

número total de variantes conhecidas detectadas por um método de referência. A especificidade analítica é a probabilidade de o ensaio ser negativo quando nenhuma variante estiver presente (1-taxa de falsos positivos). Essa medida é estabelecida dividindo o número de verdadeiros negativos pela soma de verdadeiros negativos e falsos positivos, tipicamente obtidos por comparação com os resultados obtidos por um método de referência como o sequenciamento de Sanger. Por último, a precisão refere-se à reprodutibilidade ou "robustez" do ensaio, ou seja, a capacidade de obter os mesmos resultados quando o ensaio é repetido. Para a reprodutibilidade, tanto a reprodutibilidade intra como inter-corridas devem ser avaliadas[143].

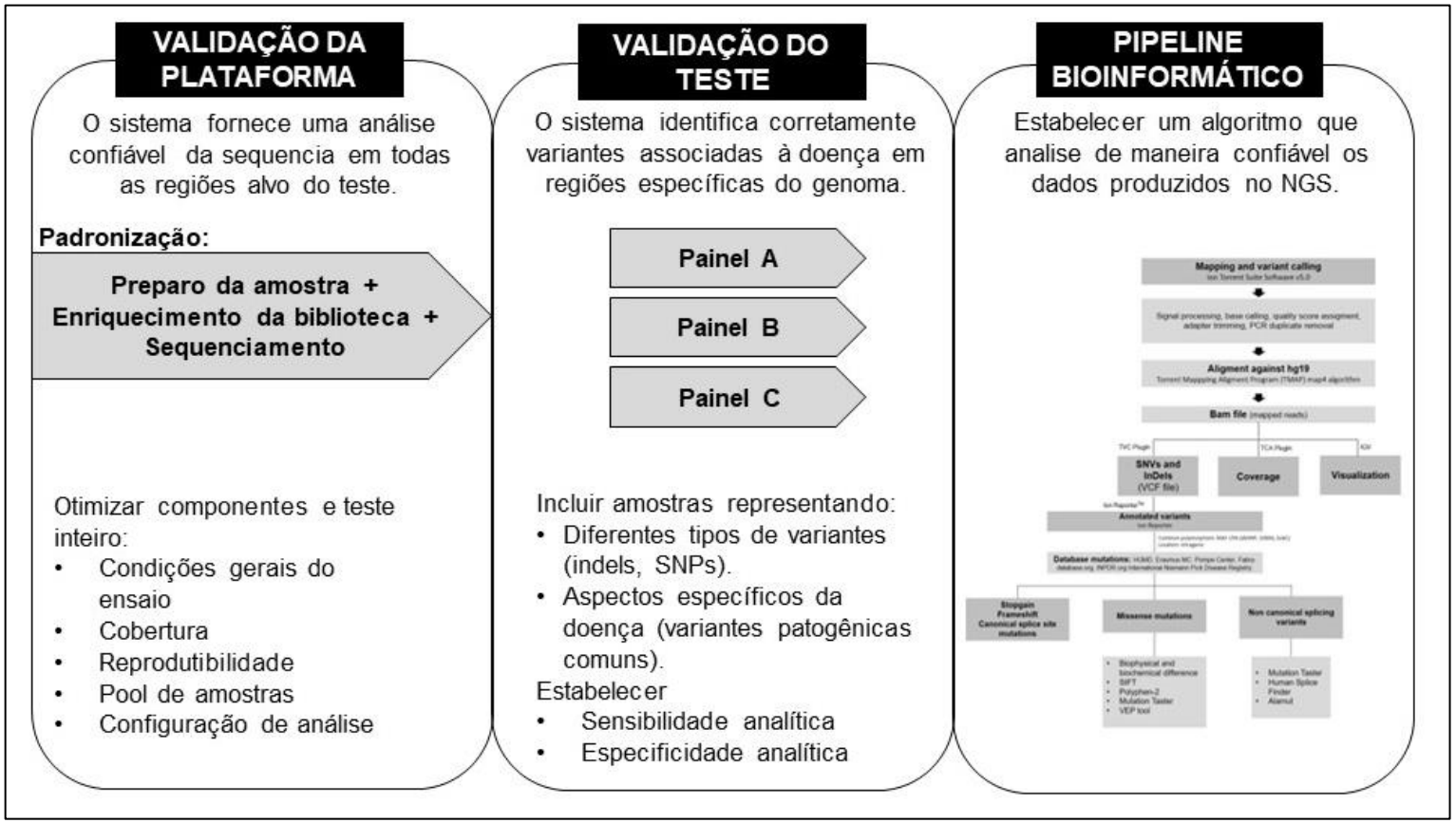

Figura 5. Fluxograma do processo de validação de um painel NGS. Adaptado de Rehm et al[141]

1.6 PAINÉIS NGS, *WES* E *WGS* NO DIAGNÓSTICO E PESQUISA DAS DLs

Existem poucos estudos sobre a aplicação de painéis NGS no diagnóstico das DLs (Tabela 8).

O estudo de Fernandez-Marmiesse et al[177] descreve pela primeira vez os resultados de um projeto piloto para avaliar a aplicação do NGS para o *screening* de mutações no diagnóstico das DLs. Utilizando uma abordagem baseada no NGS para a análise simultânea de 57 genes lisossomais, foram capazes de diagnosticar corretamente 18 controles positivos cegos e 25 pacientes com suspeita de DL, os pacientes eram da Espanha e de Portugal. Foram analisados 84 probandos, incluindo 18 controles positivos com diagnóstico clínico e bioquímico prévio de alguma DL e 66 pacientes com suspeita, de moderada a alta, de alguma DL. Foi realizada a preparação da biblioteca através da amplificação das regiões codificantes e junções exon-intron dos 54 genes incluídos no painel. O sequenciamento foi feito utilizando a plataforma SOLiD4 e HiSeq2000 Illumina, sendo que a última plataforma mencionada rendeu uma cobertura maior (99.97% das bases cobertas pelo menos 20x). Os autores encontraram duas dificuldades técnicas relacionadas ao enriquecimento das bibliotecas; a primeira devido à composição de bases de alguns exons, o qual interferiria com a eficiência do enriquecimento, assim, os exons próximos a regiões repetitivas não foram completamente cobertos. Para exons localizados em ilhas CpG, a cobertura também diminuiu drasticamente, produzindo *gaps* na cobertura de determinados exons dos genes associados às DLs, principalmente nos genes *IDUA*, *GBA* e *GAA*. Os autores também apontam dificuldades na filtragem e interpretação de dados gerados pelo NGS. A interpretação de todas as variantes detectadas é um dos principais desafios, além disso, as irregularidades na cobertura contribuiram com resultados falsos-positivos.

Outro exemplo da utilidade clínica dos painéis NGS no diagnóstico das DLs é o do diagnóstico da doença de Pompe de início tardio realizado pelo grupo de Lèvesque et al[178]. Os autores desenvolveram um painel para analisar os exons e junções exon-intron do gene *GAA* junto com outros 77 genes associados a diversas desordens neuromusculares com fenótipos sobrepostos. Eles alcançaram uma cobertura média de 200x com todos os exons do *GAA* sendo cobertos 20x e apenas 0.03% dos exons dos genes do painel com coberturas menores de 20x. Obtiveram uma sensibilidade de 100% e uma especificidade de 98%. Analisaram 34 pacientes com desordens musculares de etiologia desconhecida, o estudo demonstrou um rendimento diagnóstico de 32%. Esse painel NGS demostrou facilitar o diagnóstico em pacientes com fraqueza muscular inespecífica o com fenótipos atípicos.

No ano 2016, o grupo de Yubero et al[179] desenvolveu um painel NGS incluindo 171 genes associados a um grupo selecionado de EIM, entre os quais se encontravam genes associados a algumas DLs como a doença de Pompe, doença de Danon, MPS I, II, III IV, VI, VII, IX, doença de Fabry, doença de Farber, GM1, GM2, doença de Gaucher, doença de Krabbe, leucodistrofia metacromática, entre outros. O objetivo do estudo foi comparar os achados de rendimento diagnóstico de um grupo com diagnóstico clínico e bioquímico prévio de algum EIM (grupo 1) com um grupo com suspeita inespecífica de algum EIM (grupo 2). Foram analisados um total de 146 pacientes, 81 do grupo 1 e 65 do grupo 2. O rendimento médio foi de 50%, sendo que o rendimento para o grupo 1 foi de 78% e para o grupo 2 de 15.4%. Foi utilizada a plataforma MiSeq da Illumina, e se obteve uma cobertura média de 355x com 93% das bases cobertas mais de 20x. Uma das limitações mais críticas que este grupo descreveu foram as limitações metodológicas. Uma fração das regiões codificantes permaneceu sem ser sequenciada, assim levando a falsos negativos. O grupo concluiu que as avaliações clínicas em combinação com achados bioquímicos consistentes, são ferramentas úteis para aumentar o rendimento diagnóstico em pacientes com EIM. Nesses casos, o uso do painel NGS é altamente produtivo e custo-efetivo.

Outros dois exemplos de aplicação do NGS para LSD tem sido descritos através de relatos breves e correspondem aos grupos de Wood et al[180] e Coutinho et al[181].

De acordo com as "orientações para o diagnóstico por NGS" da *EuroGentest and the European Society of Human Genetics* e outros guidelines[141,143,182], "para fins de diagnóstico, apenas os genes com uma relação conhecida (isto é, publicada e confirmada) entre o genótipo e a patologia devem ser incluídos na análise" através do NGS. No entanto, existem algumas aplicações com foco em pesquisa que se aventuram a aplicar a tecnologia NGS para identificar variações de sequência em genes candidatos conhecidos, mas também potencialmente novos genes candidatos para as DLs. Entre as primeiras publicações de um extenso painel de pesquisa para a identificação de DLs está o estudo de Di Fruscio et al[183] em 2015, com o objetivo de detectar variantes genéticas na via endocítica no lisossomo. Detalhes deste painel NGS podem ser encontrados na Tabela 8.

Nos últimos anos, as abordagens de WES e WGS passaram a ser consideradas como testes universais para a identificação da maioria de distúrbios mendelianos[184–186] com a exceção daquelas causadas por variações estruturais complexas. O WES tem demonstrado ser uma solução económica com um maior rendimento de sequenciamento e que favorece a

identificação de novos genes associados a doenças. Três publicações recentes demostram a complexidade e variabilidade clínica das DLs, bem como o papel fundamental do WES em expor lacunas no conhecimento clínico devido à raridade de muitas dessas doenças. O primeiro exemplo corresponde ao de três irmãos, entre 40 e 60 anos de idade, que desenvolveram osteólise severa dos ossos nas mãos e pés. Os genes conhecidos de displasia esquelética foram sequenciados, mas nenhuma mutação foi encontrada. O WES eventualmente revelou variantes patogênicas no *ASAH*1, causando deficiência da ceramidase ácida ou doença de Farber, o que foi confirmado posteriormente por testes enzimáticos. Os pacientes tinham sintomas "típicos" da doença na infância, mas eram tão leves que não suspeitavam de uma DL. Estes se tornaram os pacientes mais idosos com diagnóstico de doença de Farber. Além disso, o gene *ASAH1* agora pode ser incluído na lista de genes testados em pacientes com osteólise[187].

O segundo exemplo, é o descrito pelo grupo de Kondo et al [188] no ano passado. Um grupo de pacientes com manifestações clínicas (características faciais grosseiras, anormalidades esqueléticas, hepatoesplenomegalia, problemas respiratórios e retardo mental) e achados bioquímicos (acumulação excessiva de GAGs urinários, níveis extremamente altos de heparan sulfato no plasma) muito similares aos apresentados por pacientes afetados com mucopolissacaridoses (MPS) foram investigados para este grupo de doenças. No entanto, não puderam ser diagnosticados enzimaticamente como MPS. Usando o WES e o sequenciamento de Sanger, foi identificada uma mutação em homozigose (c.1492C>T, p.Arg498Trp) no gene *VPS33A* dos 13 pacientes. Este gene está envolvido na via endocítica e autofágica, mas a mutação não afetou nenhuma destas duas vias. A acidificação lisossômica e o acúmulo de heparan sulfato foram detectados nas células dos pacientes, sugerindo um novo papel desses genes na função lisossomal. Por tanto, este grupo propôs um novo tipo de MPS que não é causado por deficiência enzimática, denominando-a MPSPS (*Mucopolysaccharidoses Plus Syndrome*).

Certamente, o WES está se mostrando como uma ferramenta com grande potencial e sua utilização é fundamental para o diagnóstico de doenças conhecidas com "apresentações atípicas"[189]. Isto é crítico na maioria das DLs, distúrbios genéticos raros que podem não ser diagnosticados por anos, uma vez que todo o espectro da variação fenotípica não é bem caracterizado, dado o número reduzido de pacientes descritos na literatura e a baixa frequência em que eles ocorrem[190–194].

Tabela 8. Parâmetros utilizados nos estudos selecionados para a análise NGS das DLs.

Parâmetros do estudo	Di Fruscio et al.[183] (Lysoplex)	Fernandez-Marmiesse et al.[177] (Painel LSD)	Levesque et al.[178] (Painel Muscular)	Wood et al.[180] (Painel LSD)	Coutinho et al.[181] (Painel LSD)	Yubero et al.[179] (Painel IEM)
Número de genes incluídos no painel	891	57	78	81	96	171
Conteúdo de genes	194 genes lisossomais, 106 genes associados à autofagia, 627 genes com papel na via endocítica	57 genes associados às DLs	*GAA* e 77 genes associados a doenças musculares	81 genes associados às DLs	96 genes codificadores de proteínas lisossomais, reguladores lisossomais e proteínas não lisossomais envolvidas na biogênese lisossomal.	171 genes associados a EIM, dentro dos quais encontram-se as DLs.
Doença	Lipofuscinose ceróide neuronal (NCL)	DLs	Doença de Pompe de início tardio	DLs	DLs	DLs
Número de pacientes analisados	48 pacientes com NCL	18 controles positivos, 66 pacientes com suspeita de DLs	34 pacientes com distúrbio muscular não estabelecido	14 controles positivos com 11 DLs diferentes	12 pacientes com DLs variadas: Mucolipidoses II α/β, MPS VI, Gaucher, Fabry, Picnodisostoses, Krabbe, GM2 e NCLs.	81 pacientes com suspeita clínica e bioquímica de algum EIM, 65 com suspeita inespecífica de EIM.
Média da cobertura	40x 95%	20x 94.5%	20x 97.7%	372x	Não informado	355x (MiSeq)

	(HiSeq1000) 100x 80%	(SOLiD) 20x 99.97% (HiSeq2000)	(MiSeq)	50x 95%		20x 93%
Rendimento diagnóstico	67% de pacientes com NCL analisados	51% dos pacientes analisados	32% dos pacientes analisados	Não determinado.	100%, devido ao uso de amostras com diagnóstico bioquímico prévio.	50% dos pacientes analisados
Valor do método	Mais robusto que WES, pode ser utilizado de maneira efetiva para o descobrimento de novos genes associados à doença	Diagnósticos inesperados (GM1, GM2), menor tempo para diagnóstico.	O diagnóstico foi estabelecido em apresentações atípicas.	Diagnóstico estabelecido, painel especial para DLs com ensaios bioquímicos disponíveis.	Painel robusto, envolvendo 96 genes associados às DLs.	Diagnóstico rápido e eficaz, particularmente no grupo que apresentava indicações clínicas e bioquímicas para o diagnóstico.

1.7 O SERVIÇO DE GENÉTICA MÉDICA DO HOSPITAL DE CLÍNICAS DE PORTO ALEGRE E SEU PAPEL NAS LSD NO BRASIL

O Serviço de Genética Médica (SGM) do Hospital de Clínicas de Porto Alegre (HCPA), criado no ano de 1982, é um centro de referência no Brasil e na América Latina para o diagnóstico e pesquisa nas diferentes áreas médicas na região sul do Brasil, especialmente de Erros Inatos do Metabolismo, e conta com profissionais treinados em avaliação clínica, diagnóstico laboratorial e manejo de doenças genéticas, EIM e DLs.

Desde 2004 é reconhecido como Centro Colaborador da OMS para o desenvolvimento de Serviços de Genética Médica na América Latina e foi habilitado como Serviço de Referência em Doenças Raras pelo Ministério da saúde no ano de 2016[195].

No SGM-HCPA, no período do 1982-2015, amostras de um total de 72,797 pacientes foram encaminhadas para investigação de algum EIM, tendo sido confirmada uma DL em 3.211 casos (4.41% ou 1 em 22), sendo 3,099 pacientes brasileiros (96.5%). Esses casos, vindo de diferentes regiões brasileiras, forneceram informações sobre a frequência relativa das DLs (Figura 6) no país e permitiram estimar a frequência mínima das DLs no país[23].

O número de diagnósticos neste serviço aumentou ao longo do tempo, em parte devido à implementação de diferentes testes enzimáticos em vários tipos de amostra (leucócitos, fibroblastos, plasma e sangue impregnado em papel filtro) e à criação de diferentes redes de diagnóstico, entre elas a “Rede EIM Brasil”, “Rede MPS Brasil”, “Rede NPC Brasil” e a “Rede DLD Brasil”, as quais envolvem muitos serviços brasileiros.

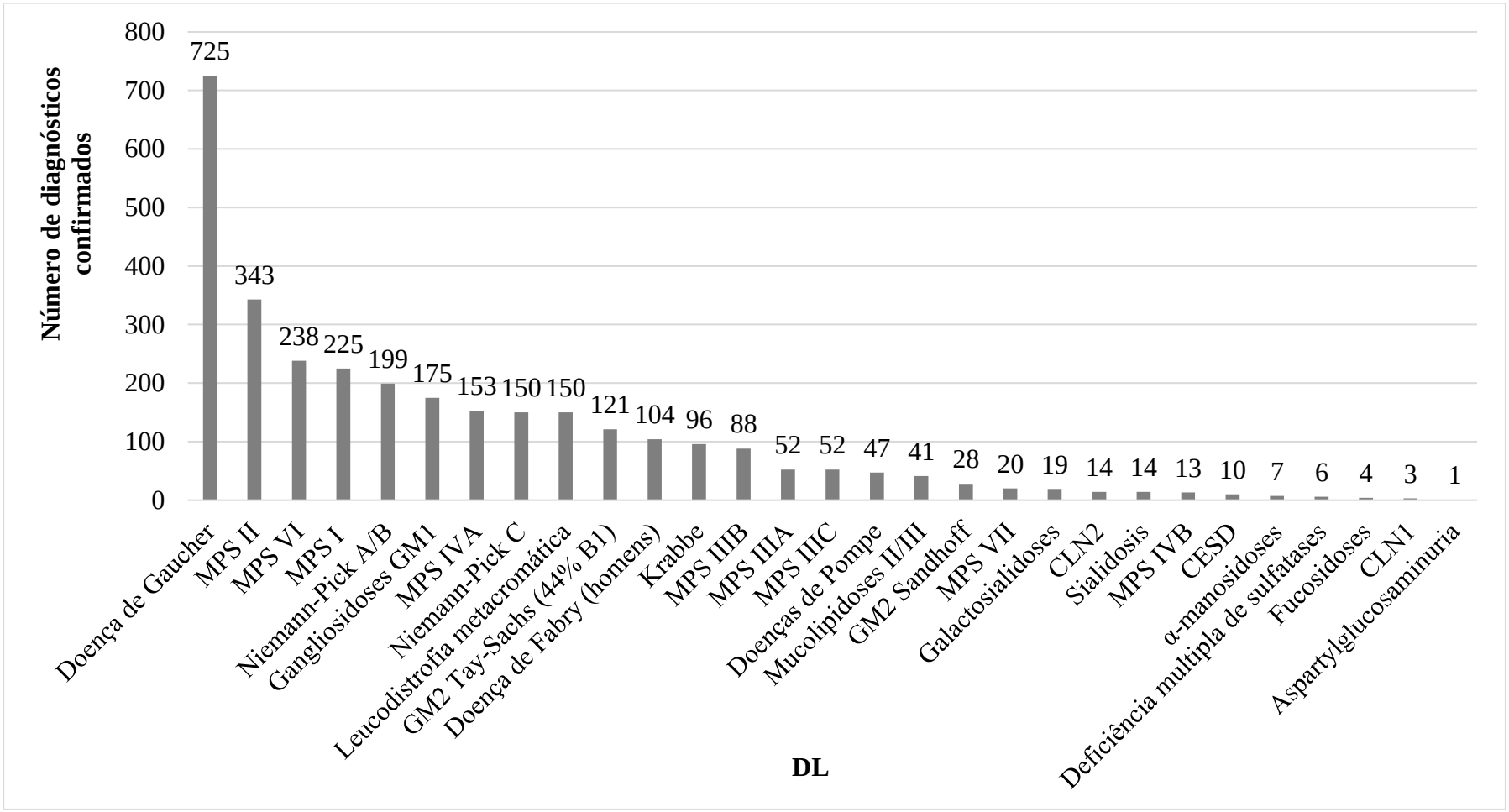

Figura 6. DLs diagnosticadas no Serviço de Genética Médica do Hospital de Clínicas de Porto Alegre no período de 1982-2015. Figura baseada no trabalho de Giugliani et al. 23.

CAPÍTULO 2: JUSTIFICATIVA

As doenças lisossômicas constituem um problema de saúde importante, com impacto enorme sobre os indivíduos afetados e suas famílias. Muitos fatores dificultam o diagnóstico destes distúrbios genéticos, sendo a variabilidade fenotípica e genotípica a principal causa e levando a um atraso significativo do diagnóstico.

Existem poucos centros especializados no diagnóstico deste grupo de doenças na América Latina, um deles é o Serviço de Genética Médica do Hospital de Clínicas de Porto Alegre, no sul do país, o qual dispõe de vários laboratórios incluindo o de Erros Inatos do Metabolismo (LEIM) e o de Genética Molecular (LGM), ambos situados entre os mais completos do país. Por este motivo, o SGM-HCPA tem se consolidado como um serviço de referência, concentrando uma grande quantidade de pacientes, referidos de diferentes regiões brasileiras e de outros países, para diagnóstico de DLs.

As mucopolissacaridoses (MPS) são o principal grupo de DLs caracterizadas no LGM através de diferentes testes moleculares, incluindo o NGS. A implementação desta abordagem para o diagnóstico de outros grupos de DLs, permitirá sua identificação e reduzirá a "odisseia diagnóstica" que essas famílias enfrentam, levando a um aconselhamento genético apropriado e uma introdução mais precoce das medidas terapêuticas disponíveis, provavelmente com melhor resultado.

Por outro lado, amostras provenientes de outros países ou de regiões distantes do país são enviadas na forma de sangue impregnado em papel filtro (SIPF) ao SGM-HCPA, devido à fácil coleta e transporte deste tipo de material. A implementação de um método de extração de ADN, a partir do SIPF, que seja eficiente e custo-efetivo, é fundamental para o diagnóstico molecular por NGS ou outros testes moleculares, especialmente em um laboratório assistencial.

Além disso, um estudo piloto de um programa de triagem neonatal (TN) está sendo desenvolvido no SGM-HCPA para doenças lisossômicas tratáveis. A validação dessa tecnologia permitirá também sua avaliação para utilização como método confirmatório para casos de triagem neonatal para doenças lisossômicas alterada, utilizando o ADN extraído do mesmo SIPF coletado para a triagem neonatal. A possibilidade de realizar a análise molecular, por NGS, a partir de ADN extraído do SIPF coletado na TN, reduziria significativamente o tempo de confirmação diagnóstica.

Neste sentido, sendo o SGM-HCPA um serviço de referência para DLs, a implementação destes painéis NGS e a demonstração da sua utilidade clínica, representará um aporte inestimável para o diagnóstico deste tipo de doenças.

CAPÍTULO 3: OBJETIVOS

3.1 Objetivo Geral

Validar e avaliar a utilidade clínica de uma estratégia baseada no sequenciamento de nova geração (NGS), estabelecendo a sensibilidade, especificidade e limitações, para a detecção de mutações em pacientes com suspeita de doenças lisossômicas selecionadas;

3.2 Objetivos Específicos

1. Desenhar e validar 3 painéis NGS, contendo 24 genes, para o diagnóstico de doenças lisossômicas selecionadas, utilizando a plataforma Ion PGM^{TM} e ADN extraído de sangue periférico de pacientes com diagnóstico bioquímico e molecular prévio de alguma DL.

2. Avaliar a qualidade e eficiência do ADN extraído de sangue impregnado em papel filtro (SIPF) para seu uso no NGS.

3. Avaliar a utilidade clínica da abordagem NGS para:

   a) diagnóstico molecular de DLs,

   b) diagnóstico diferencial de DLs e,

   c) confirmação diagnóstica de casos alterados na triagem neonatal para pelo menos 3 DLs (doença de Gaucher, Fabry e Pompe).

CAPÍTULO 4: RESULTADOS

O trabalho desenvolvido nesta tese resultou em 4 artigos científicos: um já publicado, um aceito para publicação, e dois em preparo para serem submetidos a publicação.

Na seção Anexos: um quinto artigo (Anexo I) contém resultados parciais desta tese, a respeito da Triagem Neonatal de DLs, e outros 4 artigos refletem o trabalho em paralelo realizado durante meus estudos de doutorado.

4.1 Artigo 1: Sensitivity, Advantages, Limitations and Clinical Utility of Targeted Next-Generation Sequencing Panels for the Diagnosis of Selected Lysosomal Storage Disorders

Artigo aceito para publicação na revista *GENETICS AND MOLECULAR BIOLOGY*, 2018.

Sensitivity, Advantages, limitations and Clinical Utility of Targeted Next-Generation Sequencing Panels for the Diagnosis of Selected Lysosomal Storage Disorders

Diana Rojas Málaga[1,2], Ana Carolina Brusius-Facchin[2], Marina Siebert[3], Gabriela Pasqualim[1,3], Maria Luiza Saraiva-Pereira[2], Carolina F.M de Souza[2], Ida V.D. Schwartz[2], Ursula Matte[3,4], Roberto Giugliani[2,]

1 Postgraduate Program in Genetics and Molecular Biology, UFRGS, Porto Alegre, Brazil

2 Medical Genetics Service, Hospital de Clinicas de Porto Alegre, Brazil

3 Experimental Research Center, Hospital de Clinicas de Porto Alegre, Brazil

4 Department of Genetics, UFRGS, Porto Alegre, Brazil

Running title: Targeted NGS panel for selected LSDs.

Corresponding author

Ana Carolina Brusius Facchin

Medical Genetics Service, Hospital de Clinicas de Porto Alegre

Rua Ramiro Barcelos 2350, Porto Alegre, RS 90035-903, Brazil.

E-mail: afacchin@hcpa.edu.br

Abstract

Lysosomal storage disorders (LSDs) constitute a heterogeneous group of approximately 50 genetic disorders. Variability in phenotype penetrance, similar clinical manifestations, and a high allelic heterogeneity are aspects that turn LSDs diagnosis challenging. The development of a powerful tool for the diagnosis could reduce the "diagnostic odyssey" for these families, leading to an appropriate genetic counseling, and a better outcome for current therapies, since enzyme replacement therapy have been approved in Brazil for Gaucher, Fabry and Pompe disease and is under development for Niemann-Pick Type B. However, application of NGS technology in the clinical diagnostic setting require a previous validation phase. Here we assess the application of this technology as a fast, accurate, and cost-effective method to determine genetic diagnosis in selected LSDs. We have designed two panels for testing simultaneously 11 genes known to harbor casual mutations of LSDs. A cohort of 58 patients were used to validate those two panels, and the clinical utility of these gene panels was tested in four novel cases. We report the assessment of a NGS approach as a new tool in the diagnosis of LSDs in our service.

Keywords: Ion Torrent, molecular diagnostics, next-generation sequencing, lysosomal storage disorders, validation.

Introduction

Lysosomal Storage Disorders (LSDs) comprise a heterogeneous group of at least 50 rare genetic disorders caused by progressive accumulation of specific substrates, generally due to a deficiency of a lysosomal enzyme (Filocamo *et al*., 2011). One main factor related to diagnosis delay is the variable severity of a wide spectrum of clinical manifestations, which are not specific of the disorder, and can overlap with symptoms of other LSDs (Vieira 2008; Martins 2013). Another challenge in this group is the high allelic heterogeneity for genetic screening. Early diagnosis is important since enzyme replacement and other available therapies improve the natural course of many of these diseases (Giugliani et al, 2016; Tajima et al, 2013; Franco et al, 2016; Muenzer et al, 2014).

Canonical approach to the diagnosis of patients with LSDs include the detection of the accumulated substrate, whenever possible, and the activity assay of the deficient enzyme, followed by Sanger sequencing of the gene associated to the disorder, which can be expensive and time consuming (Wang 2011). Fortunately, new technologies are becoming more accessible and relative affordable for the diagnostic routine. Targeted next-generation sequencing (TNGS) allows the simultaneous screening of several LSDs-related genes, with great depth of coverage, manageable interpretation, and relatively low risk of finding variants of unknown significance, with short turn-around times for the final report (Bhattacharjee 2015; Rehm 2013).

However, before using TNGS technologies as a diagnostic tool, the validation of each test offered in the clinical setting is required. This validation is essential for stablishing critical parameters, from sample processing to the analysis and interpretation steps, following the recommendations of published guidelines (Rehm 2013, Gargis 2012).

Here, we present the development and validation of two different TNGS panels of genes related to a subgroup of LSDs, offered as a diagnostic alternative by a Brazilian reference service for rare diseases. The sensitivity, advantages, drawbacks, and clinical utility of these TNGS panels are then reported.

Methods

Gene panel design

Genes associated with LSDs with overlapping clinical manifestations as well as related deficiencies were included in our panels (Figure 1). The two panels comprised 11 genes:

Panel A: *GLA* (Fabry disease), *NAGA* (Schindler disease), *GAA* (Pompe disease), *LAMP2* (Danon disease), and Panel B: *NPC1* (Niemann-Pick disease type C1), *NPC2* (Niemann-Pick disease type C2), *GBA1* (Gaucher disease), *LIPA* (Lysosomal acid lipase deficiency), *SMPD1* (Niemann-Pick disease type A/B), *CHIT1* (Chitotriosidase deficiency) and *PSAP* (Prosaposin deficiency and saposin B deficiency). Custom primers were designed using Ion Ampliseq™ Designer v3.4 (Thermo Fisher Scientific) to generate a pool of primers for amplification of genomic regions of interest. Each one consists of two primer pools that target the entire coding region, including 20pb of intron-exon junction. Missed areas in the design were fill in with Sanger sequencing to reach a 100% breadth of coverage.

Samples

The validation phase was performed using gDNA, extracted from whole blood using a standard saline extraction method (Miller 1988), from 55 previously diagnosed patients (22 for panel A and 33 for panel B), who underwent previous investigation with biochemical tests and Sanger sequencing (with known mutations and polymorphisms, including SNPs and small *indels*). Samples from 3 healthy adults were also analyzed. All probands derived from Medical Genetic Service, Hospital de Clinicas de Porto Alegre, Brazil. All samples were anonymized, sequenced, and analyzed in a blind manner. TNGS was performed using Ion Torrent Personal Genome Machine™ (PGM™) System (Thermo Fisher Scientific). Clinical utility assay of the validated tests was assessed by evaluating four patients with suspected LSDs. This study was approved by the institutional Ethics Committee of HCPA, Brazil (#15-0165).

Multiplex PCR enrichment, library construction, and massive parallel sequencing

Twenty nanograms of each gDNA sample were used for PCR enrichment of targets by applying the two custom Ampliseq™ panel (Thermo Fisher Scientific). Each panel consisted of 2 separated PCR primer pools. Library was constructed using Ion AmpliSeq™ Library kit 2.0 (Thermo Fisher Scientific). Eight to nine samples barcoded with Ion Xpress™ Barcode Adapters kit (Thermo Fisher Scientific) were included in each set of library preparation. Unamplified libraries were purified with Agencourt AMPure XP kit (Beckman Coulter). Libraries were prepared in equimolar concentrations using the Ion Library Equalizer™ kit (Thermo Fisher Scientific) or quantified using the Qubit® dsDNA HS kit (Thermo Fisher Scientific), followed by dilution to the same concentration. For template preparation, the barcoded libraries were pooled in equimolar concentrations of 100

pM each and were subsequently submitted to emulsion PCR (emPCR) using the Ion PGM™ Template OT2 200 kit (Thermo Fisher Scientific) on the Ion OneTouch2™ Instrument (Thermo Fisher Scientific). The percent of positive Ion Sphere Particles (ISPs) was defined with flow cytometry performed on Attune® Acoustic Focusing Flow Cytometer (Thermo Fisher Scientific) according to the demonstrated protocol (Part. no. 4477181, Thermo Fisher Scientific). Positive ISPs were enriched using Ion OneTouch™ ES (Enrichment System; Thermo Fisher Scientific).

All barcoded samples were loaded onto Ion 314™ chips v2 (Thermo Fisher Scientific) taking up to 8-9 samples on a single chip per sequencing run. Chip loading procedure was performed according to the user guide for the Ion PGM sequencing 200 kit v2 (Thermo Fisher Scientific), following the manufacturer´s instruction.

Data analysis

Raw signal data were analyzed using Torrent Suite Software v.5.0 (Thermo Fisher Scientific). Primary analysis includes signal processing, base calling, demultiplexing, read alignment to human genome 19 reference (Genome Reference Consortium GRCh37), quality control of mapping quality, coverage analysis, and variant calling. Afterwards, a list of detected sequence variants, including SNPs and small insertions/deletions, was imported into Ion Reporter™ Software (Thermo Fisher Scientific) for annotation. Alignments were visually verified with the Integrative Genomics Viewer (IGV) v2.3 (Robinson et al. 2011).

Candidate variants met the following criteria: detected on both strands and to account for 20% of total reads at that site, quality score ≥20, minimum read depth of 100X and variant frequency in the population ≤1%. The filtered variants were then compared to mutation databases, including dbSNP (htpp://www.ncbi.nlm.nih.gov/projects/SNP/), 1000G (http://browser.1000genomes.org), ExAC (http://exac.broadinstitute.org), Online Archive of Brazilian Mutations (http://abraom.ib.usp.br/), HGMD (http://www.hgmd.cf.ac.uk/ac/), Pompe Center at Erasmus Medical Center (http://www.pompecenter.nl/), Fabry-database.org (http://fabry-database.org/) and the International Niemann-Pick Rare Disease Registry (https://inpdr.org/). All databases were last accessed in September 2017. Evaluation of the pathogenicity of the novel variants of unknown significance (VUS) (i.e., not found in any of the mutation databases or not previously described in the literature) were analyzed with *in silico* web tools, such as SIFT (Sim 2012, Kumar 2009), Polyphen-2 (Adzhubei 2009), Mutation Taster (Schwarz et al., 2014), to predict potential protein

deleterious effect on protein function. To evaluate the possible effect of synonymous variant in gene splicing, we used the Human Splicer Finding web tool (Desmet 2009). Indels were analyzed by using VEST (Variant Effect Scoring Tool), VEP (Vep Effect Preditor) as well as Mutation Taster (Carter *et al.*, 2013, Douville *et al.*, 2016, McLaren *et al.*, 2016). Nonsense, frameshift, and canonical splice mutations were classified automatically as pathogenic (Richards *et al.*, 2015).

Performance characteristics

Run metrics and coverage analyses were performed to identify systematic deficiencies. We analyzed depth of coverage (DoC) in the targeted amplicons to assess target enrichment across all 58 samples data sets, to establish an acceptable reference range for key measures.

Two coverage analysis were generated: (1) High-level DoC overview plot, based on Tayoun *et al*., 2013, with relative DoC in the y-axis and amplicons on the x-axis, highlighting in red, the amplicons with significant lower coverage (Figure 2); (2) relative DoC over exons per gene plot (Supplementary Figure S1-S2). Direct visual inspection of amplicons reads on IGV v2.3 and evaluation of high-level DoC overview of coverage were used to established the reportable ranges for each panel.

Sensitivity and specificity were calculated (overall and for each gene) and compared with results obtained by standard testing by Sanger sequencing. False negative and positive overall rates were also calculated. To assess reproducibility of the assay, we measured concordance between independent runs using relative DoC.

Sanger sequencing

Sanger sequencing was performed for confirmation of all variants, to fill the regions missed by the custom panel design and low-coverage regions, and in the analysis of clinical utility. gDNA was amplified using specific primers designed for the free software Primer3 v.0.4.0. (available upon request). Amplicons were sequenced by both ends using the Big Dye Terminator v3.1 cycle sequencing kit, and fragments were resolved on an ABI 3500 DNA Analyzer (Thermo Fisher Scientific). Analysis of results was performed with the BioEdit v7.2.5 package free software.

Results

Run metrics

Our designs generate a total of 73 and 118 amplicons for Panel A and B, respectively. Mean amplicon read length was 150-180pb. Sequencing of genes generated reads in the range of 69000 to 7600 per sample. An evenly distributed mean depth of coverage for both panels was achieved and a mean of 95% of targeted bases were covered at least 100X. Another run metrics are summarized in Table 1.

Coverage analysis

An overview coverage for all analyzed samples is shown in Figure 1. Although the coverage for *GAA* and *SMPD1* was expected to be 94.54 and 100%, respectively, the actual mean coverage was found to be 92.46 and 97.22%. The coverage analysis demonstrates two regions poorly covered in these genes that is shown in more detail in Figure S1-S2 in Supplementary Material. Unfortunately, the low covered region is the location of c.573delT (p.Ser192fs) *SMPD1* mutation. The actual coverage for all other genes was as expected to be, based on probe design (100%).

Sensitivity

To assess the analytical sensitivity of the panels (Table 2 and 3), we compared results obtained by Sanger sequencing to those obtained by TNGS, including as many different types of variations as possible: nonsense, missense, small deletions, small insertions, splicing, and intronic variants. A total of 57 variants (pathogenic and polymorphisms) were analyzed (Table 4). We have also identified their correct zygosity status (data not shown).

Our assay identified precisely all casual mutations for LSDs, excepting 2 in Panel B. For this panel, one limitation was the inability to detect: (1) *SMPD1* c.573delT, p.Ser192fs, located in a region with low coverage, and (2) *GBA1* c.[1448T>G; 1483G>C; 1497G>C], p.[Leu444Pro;Ala456Pro,Val460Val] (Table 2 and 4).

Specificity

Of all sequenced DNA samples, we have identified 3 false positives in panel A (3 into 7476 true negative) and 5 false positives in panel B (5 into 15054 true negative), resulting in a specificity value of 99.96% (95%CI=0.998-0.999) and 99.97% (95%CI=0.9992-0.999), respectively (Table 2). Specificity by gene is shown in Table 3. These false positives were located in low coverage regions which are prone to sequencing errors.

Reproducibility

To determine the reproducibility of our assay, we sequenced 24 samples divided in 3 independent runs for Panel A, and 34 samples divided in 4 independent runs for Panel B (Figure 3).

Clinical Utility Assessment

Case 1

A 15-year-old male patient suspected of having a LSD, referred to our service via NPC Brazil Network. The main clinical findings were unexplained hepatosplenomegaly and myelogram with presence of numerous hystiocytes. Several biochemical assays were performed to reach a diagnosis, including measurement of oxysterol and activity of chitotriosidase, lysosomal acid lipase and b-galactosidase as reference enzyme, that were all within normal ranges. Filipin test resulted inconclusive. Eventually was suspected NPA/B and ASM enzyme activity in cultured skin fibroblast was tested, resulting in 1.25 nmol/h/mg prot (reference value: 49-72), this result indicates NPA/B disease. Due to several factors, like request of new samples for the biochemical assays, it took approximately 12 months to reach this biochemical diagnosis. Panel B, which includes genes related to LSD with hepatosplenomegaly as common clinical manifestation, was utilized as second-tier diagnostic approach. We found 2 pathogenic variants in *SMPD1*, both confirmed by Sanger sequencing: p.Arg610del (c.1826_1828delGCC) (rs120074118) and p.Asp420fs (c.1259delA), the last one being a novel, unreported mutation and not found in controls (n=32).

Case 2

A 21-month-old patient, daughter of consanguineous parents, who presented macrocephaly and hepatosplenomegaly as main clinical features, high cholesterol (228mg/dl) and triglycerides (492 mg/dl) levels, elevated liver enzymes (GGT: 137UI/L; TGP:256UI/L) as well as low levels of sphingomyelinase activity, was referred for molecular analysis of *SMPD1* gene. TNGS (Panel B) reveals the homozygous small deletion p.Leu474fs (c.1420_1421delCT), which was reported previously as being pathogenic (rs398123476).

Case 3

A 21-year-old female, child of a non-consanguineous marriage, diagnosis of hypertrophic cardiomyopathy at 18 years of age and with previous diagnosis of Danon disease, was

referred to our service to mutation analysis of *LAMP2* gene. TNGS (Panel A) detected the hemizygous variant p.Asn242fs (c. 725delA), a novel pathogenic variant.

Case 4

A 16-year-old male with suspicion of Danon disease due to hypertrophic cardiomyopathy with anomalous pathway, also presented intellectual deficiency, proximal myopathy and alterations in liver test. As case 3, was used Panel A, detecting the hemizygous variant c.741+1G>A, described as pathogenic (HGMD CS003703).

Discussion

Due to various reasons, such as a wide clinical and genetic heterogeneity, LSDs are difficult to diagnose, and it can take several years to reach a final diagnosis (Vieira 2008, Martins 2013). Even if no treatment is available for many of these disorders, genetic diagnosis has potential benefits as predict the prognosis, genetic counselling, and family screening. Recent studies highlight the clinical utility of TNGS technology for genetic diagnosis of LSDs (Wood 2013, Fernandez-Marmiesse 2014, Lévesque Sébastin 2016). Although some TNGS restrictions, such as the inability to detect large indels and structural variants, there are several advantages to consider this approach to be early applied in the investigation of patients with LSDs like high coverage, completeness, low rate of incidental finding, and potential to reduce diagnostic delay. TNGS assay involve various technical steps, starting from sample preparation to analysis and data interpretation, and each one requires a full validation. We have presented data on the developing and validation of two gene panels, designed following the criteria of overlapping clinical manifestations, to be offered as a diagnostic option by a reference center of rare disease (Figure 1). Prior to TNGS is necessary an enrichment step of the genes included in the panel through capture approaches based on hybridization or PCR-based strategies. The last one, especially suitable for the investigation of regions less than 100kb, versatile design and the most convenient for analysis of genes with pseudogenes due to its high specificity, sensitivity, and reproducibility (Claes & De Leener 2014). Despite some disadvantages have been reported for this approach (time-consuming, uneven coverage of the target regions due to unequal PCR efficiency across the various amplicons, allelic dropout and difficulties to detect large deletion/insertion events), Ion Ampliseq targeted technology utilized an PCR-based method (high throughput multiplex PCR) for this purpose, overcoming some of the limitations and

providing a high specificity (here, 99.96% and 99.97%) and uniformity (91%-98% for both panel) as those obtained in this study.

From run metrics results, we can conclude that all samples were uniformly covered at depths that exceed the minimum coverage required (100 X) for accurate calling of variants. Bioinformatics pipeline applied here demonstrated high sensitivity for panel A (sensitivity 100%) and for panel B (sensitivity 93.75%) (Table 1). The use of normal controls (n=3) allowed the identification of 8 platform-specific false positive variants and then filter these variants from subsequent analyses. A high reproducibility was observed revealing a high concordance between independent runs.

Inadequate coverage regions were identified by coverage plots (Figure 2), being those regions completed with Sanger sequencing. A technical difficulty encountered was related to enrichment of some targets and, as a consequence, low sequence coverage was found. This was observed at two targets corresponding to *GAA* and *SMPD1* genes. A high GC-content region (70%) was probably the main reason why the GAA amplicon was poorly covered (~20X). For *SMPD1* low covered amplicon (~30X) was identified both a GC-content of 66% and a homopolymeric region within the target. These are well-recognized limitations of NGS sequencing. As recommended by the ACMG, both tests achieved a 100% breadth of coverage when complemented with gold-standard DNA sequencing that improve clinical sensitivity.

Another major limitation of Panel B was the inability to detect *RecNci*I allele, c.[1448T>G; 1483G>C; 1497G>C]/ p.[Leu444Pro;Ala456Pro,Val460Val]. Highly-sequence similarity between functional genes and their pseudogenes can make it difficult the detection of genuine mutation due to the ambivalent mapping in the analysis of NGS data, which cannot always be avoided. Sanger sequencing is generally used to elucidate the correct variant mapping (Kathleen Clases & Kim De Leener). In our study, the presence of *GBAP1*, highly homologous *GBA1* pseudogene, complicated sequencing analysis by NGS, being the *RecNci*I allele particularly difficult to assess since mutant bases in *GBA1* (exon 10) are the wild type sequence in its pseudogene. Panel B failed to detect this allele (Table), representing a case of false negative when present because variant-containing reads align to homologous loci. Our strategy to infer the presence of *RecNci*I was based on the employment of a global alignment strategy, analyzing then the DoC of exon 10 *GBA1* and homologous *GBAP1* region. We observed that in the presence of *Rec* allele, an uneven reads

distribution was observed, due to the exclusive alignment of variant-containing reads with *GBAP1*. As examples: (a) For homozygous N370S, we observed a DoC of 369 X for *GBA1* exon 10 and 370X for *GBAP1* homologous region; (b) In the case of compound heterozygous (N370S/*RecNci*I) for exon 10 *GBA1* was observed a DoC of 144X and 313X for *GBAP1*. Therefore, presence of this *Rec* allele was infered, but Sanger sequencing using specific primer pair for exon 10 was required for confirmation of this inference.

Clinical utility assessment was performed. Two pathogenic variants were found, *SMPD1* p.Asp420fs (c.1259delA) and *LAMP2* p.Asn242fs (c. 725delA), demonstrating that our TNGS panel is a sensitive tool, with faster TAT for provision of results, and relatively low cost (~USD 320 per sample) when compared with Sanger sequencing of individual genes, and showing the potential role for diagnosis of LSDs in our Medical Genetics Service.

In conclusion, TNGS technology could be used for the simultaneous testing of a broad range of SNPs and indels, being fast, accurate as well as a cost-effective method for the diagnosis of selected LSDs. It could allow faster diagnosis and earlier treatment of patients, contributing to reduce the morbidity of the diseases and to improve patient survival and quality of life.

Acknowledgments

This work was funded by the FIPE/HCPA (Project number: 15-0165). DRM has a fellowship from Innovate Peru. The investigation was partially funded by INAGEMP (CNpq/CAPES/FAPERGS).

References

1. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR et al. (2010) A method and server for predicting damaging missense mutations. Nat Methods 7: 248–249.
2. Bhattacharjee A, Sokolsky T, Wyman S, Reese MG, Puffenberger E, Strauss K, Morton H, Parad RB, Naylor EW et al. (2015) Development of DNA Confirmatory and High-Risk Diagnostic Testing for Newborns Using Targeted Next-Generation DNA Sequencing. Genet Med 17(5):337-47.
3. Carter H, Douville C, Stenson PD, Cooper DN and Karchin R (2013) Identifying Mendelian disease genes with the Variant Effect Scoring Tool. BMC Genomics 14(3): S3.
4. Claes K and De Leener K (2014) Dealing with Pseudogenes in Molecular Diagnostics in the Next-Generation Sequencing Era. Methods Mol Biol 11671: 303-15.
5. Desmet FO, Hamroun D, Lalande M, Collod-Béroud G, Claustres M and Béroud C (2009) Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acid Res 37(9): e67.
6. Douville C, Masica DL, Stenson PD, Cooper DN, Gygax D, Kim R, Ryan M and Karchin R. (2016) Assessing the pathogenicity of insertion and deletion variants with the Variant Effect Scoring Tool (VEST-indel). Hum Mutat 37(1):28-35.
7. Fernandez-Marmiesse A, Morey M, Pineda M, Eiris J, Couce ML, Castro-Gago M, Fraga JM, Lacerda L, Gouveia S, Pérez-Poyato MS et al. (2014) Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. Orphanet J Rare Dis 9:59.
8. Filocamo, M. and, & Morrone, A. (2011) Lysosomal storage disorders: molecular basis and laboratory testing. Hum Genomics 5(3): 156–69.
9. Franco JF, Soares DC, Torres LC, Leal GN, Cunha MT, Honjo RS, Bertola DR and Kim CA (2016) Impact of early enzyme-replacement therapy for mucopolysaccharidosis VI: results of a long-term follow-up of Brazilian siblings. Genet Mol Res 15(12).
10. Gargis AS, Kalman L, Berry MW, Bick DP, Dimmock DP, Hambuch T, Lu F, Lyon E, Voelkerding KV, Zehnbauer BA et al. (2012) Assuring the quality of next-generation sequencing in clinical laboratory practice. Nat Biotechnol 30:1033-1036.
11. Giugliani R, Federhen A, Vairo F, Vanzella C, Pasqualim G, Silva LMR, Giugliani L, Boer APK, Souza CFM, Matte U et al. (2016) Emerging drugs for the treatment of Mucopolysaccharidoses. Expert Opin on Emerg Drugs 21(1):9-26.
12. Kumar P, Henikoff S and Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4(7):1073-81.
13. Lévesque S, Auray-Blais C, Gravel E, Boutin M, Dempsey-Nunez L, Jacques PE, Chenier S, Larue S, Rioux MF, Al-Hertani W et al. (2016) Diagnosis of late-onset Pompe disease and other muscle disorders by next-generation sequencing. Orphanet J of Rare Dis 11:1-10.
14. Martins AM, Kyosen SO, Garrote J, Marques FM, Guilhem JG, Macedo E, Sobral Neto J and Ura S et al. (2013) Demographic characterization of Brazilian patients enrolled in the Fabry Registry. Genet Mol Res 12(1): 136-142.
15. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P and Cunningham F (2016) The Ensembl Variant Effect Predictor. Genome Biol 17(1):122.
16. Miller SA, Dykes DD and Polesky HF (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16(3): 1215.

17. Muenzer J (2014). Early initiation of enzyme replacement therapy for the mucopolysaccharidoses. Mol Genet Metab 111(2): 63-72.
18. Rehm HL, Bale SJ, Bayrak-Toydemir P, Berg JS, Brown KK, Deignan JL, Friez MJ, Funke BH, Hegde MR, Lyon E et al. (2013) ACMG Clinical Laboratory Standards for Next-Generation Sequencing. Genet Med 5(9): 733-47.
19. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E et al. (2015) Standards and guideline for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association of Molecular Pathology. Genet Med 17(5): 405-424.
20. Schwarz JM, Cooper DN, Schuelke M and Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11(4):361-2.
21. Sim NL, Kumar P, Hu J, Henikoff S, Schneider G, and Ng PC (2012) SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res 40:452–457.
22. Tajima G, Sakura N, Kosuga M, Okuyama T and Kobayashi M (2013) Effects of idursulfase enzyme replacement therapy for Mucopolysaccharidosis type II when started in early infancy: Comparison in two siblings. Mol Genet Metab 108(3): 172-177.
23. Tayoun ANA, Tunkey CD, Pugh TJ, Ross T, Shah M, Lee CC, Harkins TT, Wells WA, Tafe LJ, Amos CI et al. (2013) A comprehensive assay for CFTR mutational analysis using next-generation sequencing. Clin Chem 59(10):1481-8.

24. Vieira T, Schwartz I, Muñoz V, Pinto L, Steiner C, Ribeiro M, Boy R, Ferraz V, de Paula A, Kim C, Acosta A and Giugliani R (2008) Mucopolysaccharidoses in Brazil: What happens from birth to biochemical diagnosis? Am J Med Genet A 146A: 1741–1747.
25. Wang RY, Bodamer OA, Watson MS, Wilcox WR and ACMG Work Group on Diagnostic Confirmation of Lysosomal Storage Diseases et al. (2011) Lysosomal storage diseases: Diagnostic confirmation and management of presymptomatic individuals. Genet Med 13(5):457-84.
26. Wood T, Basehore M, Jones J, Friez M, Cathey S and Pollard L (2013) Development of a next generation sequencing panel for lysosomal storage disorders. Mol Genet Metab 108 (2): S100.

Figure 1. Overlapping of clinical manifestations among LSDs. Venn´s four-set diagram represented by causal genes. **A**, Panel A. **B**, Panel B.

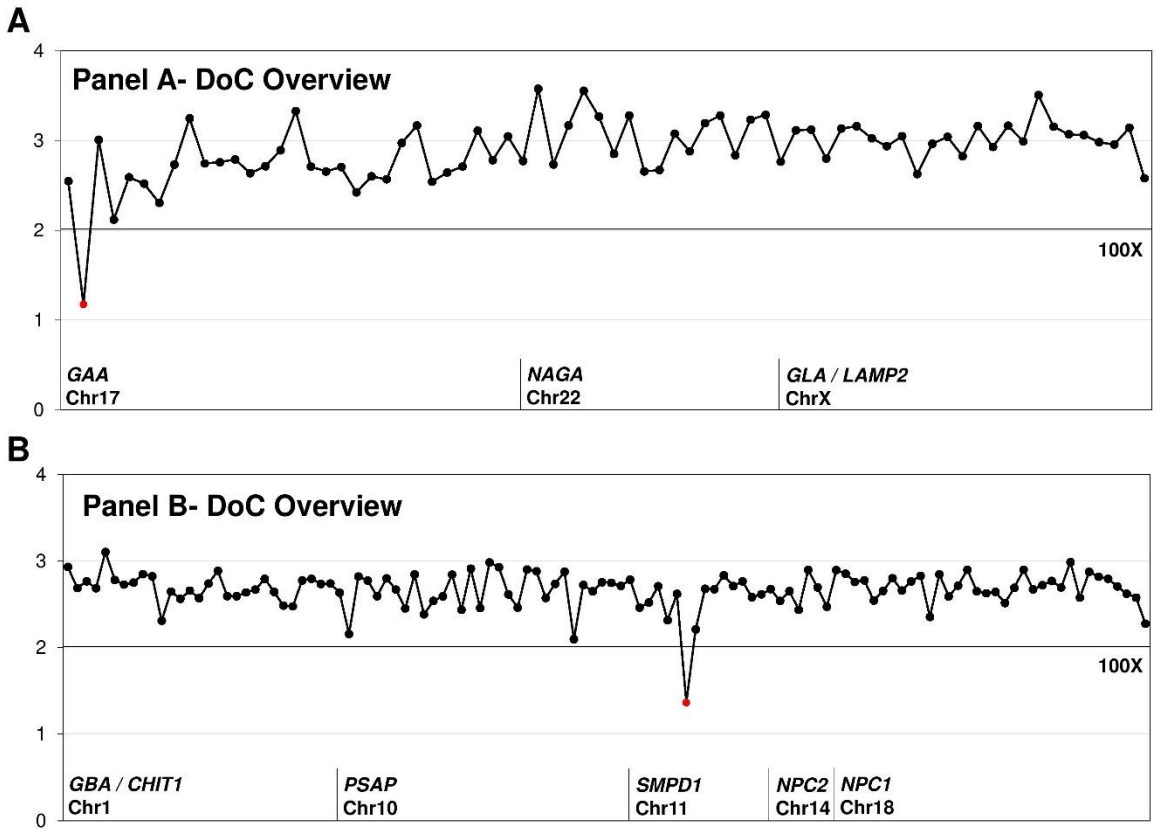

Figure 2. **Depth of Coverage (DoC) of panels presented in this study**. Overview for all 191 custom amplicons designed for Panel A (73 amplicons) (upper panel) and Panel B (118 amplicons) (lower panel) for TNGS. Line indicates a DoC of 100X. Relative DoC on the y-axis and amplicons on the x-axis.

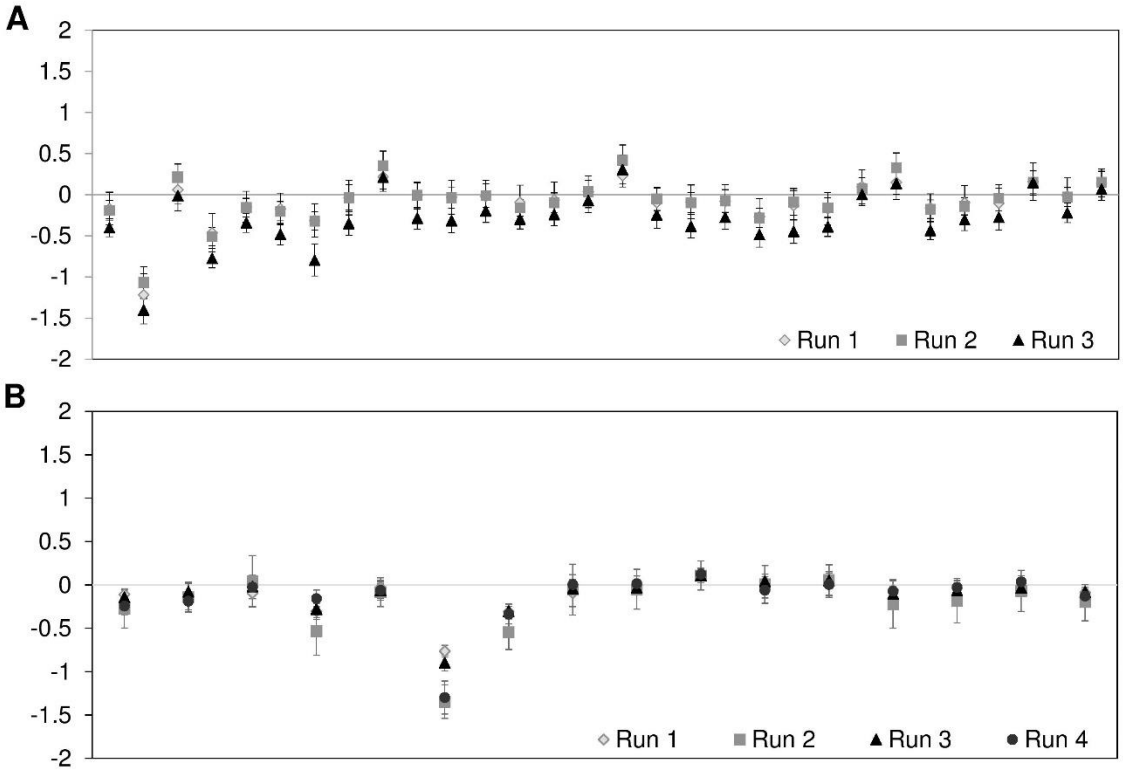

Figure 3. Reproducibility of assays. (A) Panel A: corresponds to the mean relative DoC at 30 amplicons (*GAA* gene) of 8 different samples sequenced in 3 different assay runs (B) Panel B: corresponds to the mean relative DoC at 16 amplicons (*SMPD1* gene) of 8-9 different samples sequenced in 4 different assay runs. Error bars represent standard deviation.

Table 1. Performance characteristic for both LSDs panels

Panel	Breadth of coverage	Mapped reads per sample	On Target	Mean Depth (X)	Uniformity	% Target bases covered		
						20x	100x	500x
A	97.74%	76729±31195	0.95±0.02	812±339	0.91±0.06	98.66±0.49	95.55±2.17	55.06±18.08
B	99.67%	69044±26566	0.86±0.05	498±198	0.98±0.01	99.72±0.14	94.71±6.41	56.42±25.39

Table 2. Analytical sensitivity, specificity, FN and FP rates for both TNGS panels.

Panel	Pathogenic variants	Polymorphism	Sensitivity	Specificity	FN rate	FP rate
	NGS/Sanger					
A	17/17	8/8	100% (25/25)*	99.96%	0.000%	0.040%
B	20/22	10/10	93.75% (30/32)**	99.97%	0.063%	0.033%

*95% CI=0.875-1; **95%CI=0.8091-0.9837, FN: false negative, FP: false positive

Table 3. Analytical sensitivity and specificity for each gene contained in our gene panels.

Panel	Gene	Sensitivity	Specificity
A	*GLA*	100% (10/10)	100% (1398/1398)
	NAGA	100% (3/3)	100% (1381/1381)
	GAA	100% (12/12)	99.89% (2987/2990)
	LAMP2	n.d*	100% (1707/1707)
B	*NPC1*	100% (8/8)	100% (4265/4265)
	NPC2	100% (1/1)	100% (5521/5521)
	GBA1	83.3% (5/6)	100% (1875/1875)
	LIPA	100% (4/4)	100% (1462/1462)
	SMPD1	91.7% (11/12)	99.8% (2775/2780)

	CHIT1	100% (1/1)	100% (1936/1936)
	PSAP	n.d*	100% (2184/2184)

n.d*: not determined because there were not positive controls with variant pathogenic in this gene.

Table 4. Variants detected in this study by TNGS and Sanger sequencing.

Gene	Sequence reference	Location	cDNA change	Protein change	dbSNP	Mutation type	NGS detected
GLA	NM_000169	Exon 01	c.32delG	p.Gly11fs	-	Deletion	Yes
		Exon 01	c.4C>T	p.Gln2Ter	-	Nonsense	Yes
		Exon 01	c.167G>A	p.Cys56Tyr	-	Missense	Yes
		Exon 02	c.334C>T	p.Arg112Cys	rs104894834	Missense	Yes
		Exon 03	c.456C>A	p.Tyr152Ter	-	Nonsense	Yes
		Exon 04	c.605G>A	p.Cys202Tyr	rs869312344	Missense	Yes
		Exon 05	c. 644A>G	p.Asn215Ser	rs28935197	Missense	Yes
		Exon 05	c.776C>G	p.Pro259Arg	-	Missense	Yes
		Exon 05	c. 790G>T	p.Asp264Tyr	rs190347120	Missense	Yes
		Exon 07	c.1102G>A	p.Ala368Thr	rs144994244	Missense	Yes
NAGA	NM_000262.2	Exon 03	c. 279G>A	p.Pro93Pro	rs133369	Missense	Yes
		Exon 06	c.720G>A	p.Gln240Gln	-	Missense	Yes
		Exon 08	c.973G>A	p.Glu325Lys	rs121434529	Missense	Yes
GAA	NM_001079804	Intron 01	c.-32-13T>G	-	rs386834236	Splicing	Yes
		Exon 03	c.596A>G	p.His199Arg	rs1042393	Missense	Yes
		Exon 03	c.668G>A	p.Arg223His	rs1042395	Missense	Yes
		Intron 8	c.1327-18A>G	-	rs2278619	Intron variant	Yes
		Exon 09	c.1374C>T	p.Tyr458Tyr	rs1800305	Missense	Yes
		Exon 10	c.1465G>A	p.Asp489Asn	rs398123169	Missense	Yes

		Exon 10	c.1504A>G	p.Met502Val	rs376067362	Missense	Yes
		Exon 14	c.1905C>A	p.Asn635Lys	-	Missense	Yes
		Exon 14	c.1941C>G	p.Cys647Trp	-	Missense	Yes
		Intron 14	c.2040+20A>G	-	rs2304836	Intron variant	Yes
		Exon 15	c.2065G>A	p.Glu689Lys	rs1800309	Missense	Yes
		Exon 18	c.2560C>T	p.Arg854Ter	rs121907943	Nonsense	Yes
NPC1	NM_000271.4	Exon 02	c.114_122delGAGGTACAA	p.Lys38_Tyr40del	-	Deletion	Yes
		Exon 05	c.530G>A	p.Cys177Tyr	rs80358252	Missense	Yes
		Exon 5, 8, 12	c.[547G>A;1093T>C;1937G>A]	p.[Ala183Thr;Ser365Pro;Arg646His]	rs111256741,-, rs112387560	Missense	Yes
		Exon 20	c.3019C>G	p.Pro1007Ala	rs80358257	Missense	Yes
		Exon 21	c.3104C>T	p.Ala1035Val	rs28942107	Missense	Yes
		Exon 21	c.3182T>C	p.Ile1061Thr	rs80358259	Missense	Yes
		Intron 22	c.3477+3 insCA	-	-	Insertion	Yes+
		Exon 24	c.3662_3662delT	p.Phe1211fs	-	Deletion	Yes
NPC2	NM_006432	Exon 01	c.58G>T	p.Glu20Ter	rs80358260	Nonsense	Yes
GBA1	NM_001005742	Exon 07	c.850C>A	p.Pro245Thr	-	Missense	Yes
		Exon 07	c.982_983insTGC	p.Leu327dup	rs121908298		Yes
		Exon 09	c.1226A>G	p.Asn370Ser	rs76763715	Missense	Yes
		Exon 09	c.1251G>C	p.Trp378Cys	-	Missense	Yes
		Exon 10	c.1448T>G	p.Leu444Pro	rs421016	Missense	Yes
		Exon 10	c.[1448T>G;1483G>C;1497G>C]	p.[Leu444Pro; Ala456Pro; Val460Val]	-	Missense	No
LIPA	NM_001127605	Exon 02	c.67G>A	p.Gly23Arg	rs1051339	Missense	Yes
		Exon 08	c.894G>A	p.Glu298Glu	rs116928232	Missense	Yes
		Exon 10	c.1204G>A	p.Gly342Arg	-	Missense	Yes
		Intron 05	c.539-5C>T	-	rs2297472	Intron variant	Yes
SMPD1	NM_000543	Exon 01	c.107T>C	p.Val36Val	rs1050228	Missense	Yes

		Exon 02	c.338G>A	p.Arg113His	rs149770879	Missense	Yes
		Exon 02	c.573delT	p.Ser192fs	rs727504167	Deletion	No
		Exon 02	c.636T>C	p.Asp212Asp	rs7951904	Missense	Yes
		Exon 02	c.690C>G	p.Arg230Arg	-	Missense	Yes
		Exon 02	c.714A>G	p.Ala238Ala	rs2682091	Missense	Yes
		Exon 02	c.739G>A	p.Gly247Ser	rs587779408	Missense	Yes
		Exon 06	c.1522G>C	p.Gly508Arg	rs1050239	Missense	Yes
		Exon 06	c.1749G>A	p.Ser583Ser	rs35098198	Missense	Yes
		Exon 06	c.1805G>C	p.Arg602Pro	-	Missense	Yes
		Exon 06	c.1805G>A	p.Arg602His	rs370129081	Missense	Yes
		Exon 06	c.1826_1828delGCC	p.Arg608del	rs120074118	Deletion	Yes
CHIT1	NM_003465.2	Exon 04	c.304G>A	p.Gly102Ser	rs2297950	Missense	Yes

+: new mutation, confirmed by Sanger.

4.3 Artigo 3: Pseudodeficiency of acid α-glucosidase: a challenge in the newborn screening for Pompe disease

Artigo publicado na revista *GENETICS AND MOLECULAR RESEARCH*, 2017.

GMR

Pseudo deficiency of acid α-glucosidase: A challenge in the newborn screening for Pompe disease

Diana Rojas Málaga[1,2], Ana C. Brusius-Facchin[2], Kristiane Michelin-Tirelli[2], Têmis M. Félix[2], Jaqueline Schulte[3], Jamile Pereira[3], Eurico Camargo Neto[3], Claudio Sampaio Filho[4], Roberto Giugliani[1,2,5].

[1]Postgraduate Program in Genetics and Molecular Biology, UFRGS, Porto Alegre, Brazil

[2]Medical Genetics Service, Hospital de Clinicas de Porto Alegre, Brazil

[3]Newborn Screening Center, Porto Alegre, Brazil

[4]Intercientifica, São José dos Campos, Brazil

[5]Department of Genetics, UFRGS, Porto Alegre, Brazil

Corresponding author: Roberto Giugliani

E-mail: rgiugliani@hcpa.edu.br

Genet. Mol. Res. 16 (4): gmr16039844

Received October 14, 2017

Accepted November 08, 2017

Published December 01, 2017

DOI http://dx.doi.org/10.4238/gmr16039844

ABSTRACT. When a low activity of acid α-glucosidase (GAA) is found, particularly in newborn screening programs, to differentiate α-glucosidase pseudo deficiency from true Pompe disease is important and urgent, as the result generates parental stress and also because this differentiation drives decisions related to the management of the case. Here, we report a case of GAA pseudo deficiency detected in a newborn screening performed by a private laboratory in Brazil. The confirmatory laboratory investigation performed at our service showed reduction of GAA activity on the dried blood spot, with inconclusive results when GAA activity was assayed in leukocytes. Genotyping of the GAA gene with next-generation sequencing revealed the common pathogenic mutation c.-32-13T>G and the "pseudo deficiency allele" p. [Gly576Ser; Glu689Lys], each one in heterozygous state and in trans. This report illustrates the need of newborn screening programs to have the adequate support to perform a comprehensive investigation whenever an abnormality is found in the initial screening test.

Key words: Alfa-glucosidase, Pseudo deficiency, Pompe disease, Newborn screening, Lysosomal diseases

INTRODUCTION

Pompe disease (PD, OMIM 232300) is an autosomal recessive lysosomal disorder (LD) caused by mutations in the GAA gene (17q) that encodes the lysosomal enzyme α-glucosidase (GAA, EC 3.2.1.20). GAA is responsible for glycogen degradation within lysosomes and its deficiency leads to accumulation of lysosomal glycogen especially in cardiac and skeletal muscle (Hirschhorn and Reuser, 2001). To date, over 550 distinct variations in GAA have been identified, although not all are considered pathogenic (www.pompecenter.nl). A specific enzyme replacement therapy was developed for Pompe disease, and is already approved in many countries (Chien and Hwu, 2007).

There is a worldwide interest in newborn screening (NBS) for LDs that, like PD, are amenable by enzyme replacement therapy as increasing evidence shows that early treatment intervention results in better outcomes (Chien and Hwu, 2007, 2013). However, some factors affect the detection of PD, including the presence of carriers and pseudo deficiency. The so-called pseudo deficiency allele p.(Gly576Ser; Glu869Lys), or c.[1726A; 2065A], causes, in homozygous state, reduction of GAA activity, which could be as low as the observed in patients affected by PD, but does not lead to the development of the disease (Tajima et al., 2002, Labrousse et al., 2010). We report here a case that illustrates the challenge of adequately classifying the patient with GAA deficiency as Pompe disease or pseudo deficiency, as this has a very important impact in the decision about the therapeutic measures to be taken.

CASE REPORT

Patient and methods

The male patient is the first child of a young and non-consanguineous couple. His father has diagnosis of type I Diabetes Mellitus. He was born by C-section after an uneventful pregnancy. His birth weight was 3175g, length 50 cm, OFC 33 cm and Apgar score 9 at first minute. At 4 days old, a routine newborn screening test was performed at a private laboratory and showed GAA activity below the lower reference range (patient: 4.4 µmol/L/hour; reference value: >5.9). The patient was referred to the Medical Genetics Service of HCPA for further investigation as described below. He evolved with normal motor development, held his head at 4 months old and walked unsupported at 11 months old.

Biochemical studies

GAA activity in the DBS was performed using a digital microfluidics platform for multiplexing enzyme assays (Neto et al., 2017). GAA activity in leukocytes was assayed with the substrate 4-methylumbelliferyl-α-D-glucoside, as previously reported (Li et al., 2004). GAA activity of parents' patient was assayed as well. For the assays in leukocytes, β-galactosidase was assayed as reference enzyme (Suzuki, 1977).

Molecular genetic studies

Peripheral blood sample was collected in tubes containing EDTA. DNA was obtained from blood samples by standard procedures (Millet et al., 1988). All coding exons (exons 2 through 20) and as well as the flanking intron/exon junctions (20pb) of the GAA gene were sequenced using next-generation sequencing method on Ion Torrent PGM platform employing a prior validated NGS panel that includes the GAA gene. The results were visualized in the Integrative Genomics Viewer (IGV) v2.3 (Broad Institute) and Ion Reporter Software v5.0 (Thermo Fisher Scientific). Parental segregation of the mutations was also determined by Sanger sequencing, using primers developed by Oba-Shinjo et al. (2009).

RESULTS

Enzyme analysis

GAA activity in the patient's leukocytes was 0.96 ± 0.03 nmol/h/mg protein, very close to the lower reference limit [reference range: 1.0-7.6] (Figure 1). GAA activities of the parents' leukocytes were within normal range (mother: 2.7; father: 1.9). Activity of β-galactosidase in leukocytes, assayed as a reference enzyme, was within normal range in all cases, confirming that the samples were adequate.

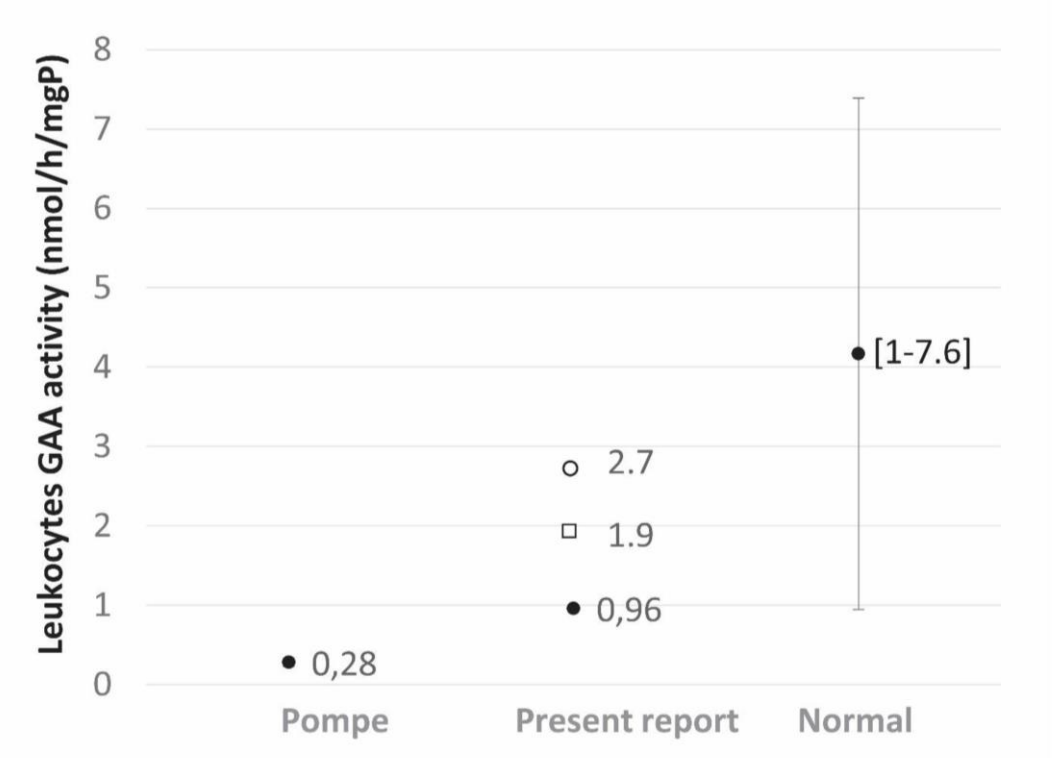

Figure 1. GAA activity in leukocytes. Pompe: Results in Pompe patients previously diagnosed; Present report: ○, mother; □ father; ● patient; Normal: Normal reference range.

Molecular genetic analysis

After sequencing the entire coding sequence and splice sites of the GAA gene, we identified 3 heterozygous variants when compared with reference sequence NM_001079804.1: the mutation c.-32-13T>G (rs386834236) in the intron 1 and the missense mutations p.Gly576Ser (c.1726G>A) (rs1800307) and p.Glu689Lys (c.2065G>A) (rs1800309) in exon 12 e 15, respectively (Figure 2). These mutations were also confirmed by Sanger sequencing. Parental analysis of the mutations determined that variant c.-32-13T>G was inherited from the father and c. [1726A; 2065A] (or p.Gly576Ser; Glu689Lys]) from the mother, both being heterozygous for the respective mutations.

These three variants are well-described in the literature as pathogenic or causative of pseudo deficiency. Nonetheless, we present in Table 1 the in-silico analyses results using SIFT, Polyphen-2, CADD, Mutation Taster, and Human Splicing Finder v.3 as pathogenicity prediction tools (Kumar et al., 2009, Adzhubei et al., 2010, Schwarz et al., 2014, Desmet et al., 2009).

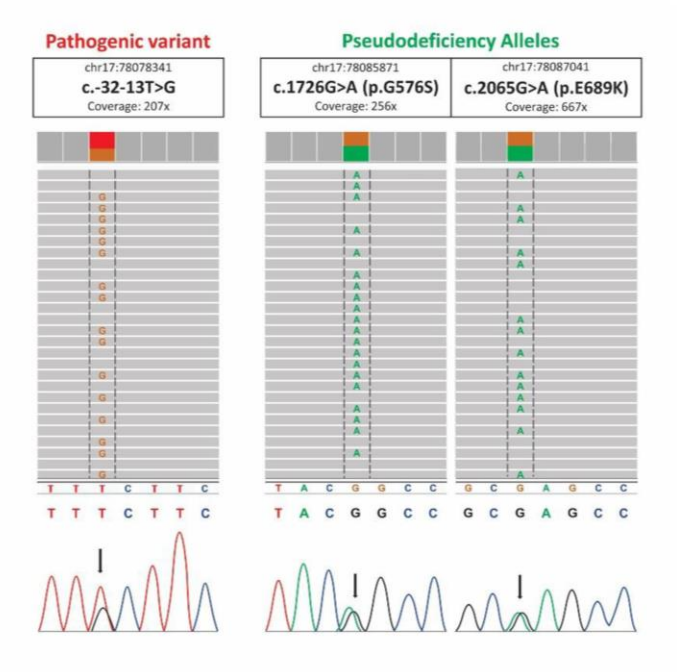

Figure 2. Identification of mutations by NGS panel, as visualized in Integrative Genomic Viewer (top panel) and confirmed by Sanger sequencing (bottom panel). A, c.-32-13T>G. B, c.1726G>A (p.G576S). C, c.2065G>A (p.E689K) in GAA.

Table 1. In silico analysis of GAA three variants found in our patient, using different prediction tools.

Variants	Clinvar	In silico Pathogenicity Prediction Tools				
		SIFT	Polyphen-2	CADD	Mutation Taster	HSF
c.-32-13T>G	Pathogenic	N.A	N.A	N.A	Disease causing	Modifier
p.Gly576Ser	Bening/ Likely bening	79%	81%	82%	Polymorphism	N.A
p.Glu689Lys	Bening/ Likely Bening/Pathogenic	82%	73%	65%	Polymorphism	N.A

Note: N.A: Not Applicable, HSF: Human Splicing Finder.

DISCUSSION

Differentiation between Pompe disease and GAA pseudo deficiency is crucial for reaching a correct diagnosis and taking a decision about the introduction of enzyme replacement therapy, available in many countries, including Brazil. Although it is present in less than 1% of babies in the United States, frequency of pseudo deficiency allele can by quite high in some populations (3.9% in Japan and 3.3% in Taiwan) (Kemper, 2013, Labrousse et al., 2010, Kumamoto et al., 2009). Data on pseudo deficiency frequency for the Brazilian population are not available. Positive results in NBS require second-tier confirmation to address false-positive results. Biochemical assay of the enzyme activity in leukocytes provides additional data, but molecular analysis of the GAA gene was needed to discriminate between PD and pseudo deficiency.

We used a prior validated NGS panel (analytical sensitivity of 100%) that includes GAA and other genes associated with different lysosomal disorders (LDs). NGS panels proved to be a valid method to genotype samples of suspected patients, in addition to being rapid, accurate and cost effective. The GAA intronic variant detected in our patient (c.-32-13T>G) is the most common variant causing late-onset PD (~40-70% of the alleles), and it is well-described as pathogenic (Huie et al., 1994a, Boerkoel et al., 1995). This variant causes three aberrant spliced transcripts: the partial or complete exon 2 skipping and a leaky wild-type splicing, which

allows a low level of GAA activity, preventing the most severe classic infantile form of PD and leading to a childhood/adult phenotype (Boerkoel et al., 1995, Dardis et al., 2014, Bergsma et al., 2015). Even though variants p.Gly576Ser and p.Glu689Lys, when analyzed independently, presented inconsistent results after in silico analysis for predicting pathogenicity as well as conflicting interpretation of pathogenicity at Clinvar, it has been reported that p.Gly576Ser reduces the GAA activity by about ~80%, while p.Glu689Lys results in 50% of normal GAA activity. The combined effect of both variants, p.[Gly576Ser, Glu689Lys] is like the effect of p.Gly576Ser alone (Kroos et al., 2008).

Our results were consistent with previous findings in NBS programs, indicating that newborn with mutation heterozygosity, pseudo deficiency homozygosity without GAA mutations and pseudo deficiency heterozygosity with and without a GAA mutation, present GAA activity significant lower than in controls, but could be distinguished from patients with Pompe disease, with few exceptions (Labrousse et al., 2010, Scott et al., 2013). Heterozygotes for the pseudo deficiency allele (like patient´s mother) are not at risk for developing PD. However, it has been hypothesized that in some cases the pseudo deficiency allele could modify the effect of another mutation (i.e. p.D645E, p.W746C, p.W746X) (Labrousse et al., 2010, Kroos et al., 2008). It has also been suggested that compound heterozygotes with one p.[Gly576Ser; Glu689Lys] allele and one pathogenic allele may develop a “Pompe-like” disease symptoms late in life due to the very low GAA activity, being a medical follow-up recommended (Kroos et al., 2008).

CONCLUSION

Our finding, and the fact that an individual homozygous for the pseudo deficiency allele was already reported in Brazil (Turaça et al., 2015) suggests that this allele is present in our population. This report of GAA pseudo deficiency detected by newborn screening illustrates and reinforces the need that such programs have a comprehensive protocol including further biochemical and genetic analysis, to provide a final diagnosis to the cases who had a positive result in the initial screening test. This process should be fast and efficient, as it was in the present case, in order to avoid parental stress in the false-positive cases and to enable the prompt start of therapy in confirmed cases.

ACKNOWLEDGEMENTS

This work was funded by the FIPE/HCPA (Project number 15-0165). DRM has a fellowship from Innovate Peru. The investigation was partially funded by INAGEMP (CNpq/CAPES/FAPERGS).

REFERENCES

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, et al. (2010). A method and server for predicting damaging missense mutations. Nat Methods. 7(4):248-249. https://doi.org/10.1038/nmeth0410-248

Bergsma AJ, Kroos M, Hoogeveen-Westerveld M, Halley D, et al. (2015). Identification and characterization of aberrant GAA pre-mRNA splicing in pompe disease using a generic approach. Hum Mutat. 36(1):57-68. https://doi.org/10.1002/humu.22705

Boerkoel CF, Exelbert R, Nicastri C, Nichols RC, et al. (1995). Leaky splicing mutation in the acid maltase gene is associated with delayed onset of glycogenosis type II. Am. J. Hum. Genet. 56: 887–897.

Chien YH and Hwu WL (2007). A review of treatment of Pompe disease in infants. Biologics. 1(3):195-201.

Chien YH, Hwu WL and Lee NC (2013). Pompe Disease: Early Diagnosis and Early Treatment Make a Difference. Pediatrics & Neonatology. 54: 219-227. https://doi.org/10.1016/j.pedneo.2013.03.009

Dardis A, Zanin I, Zampieri S, Stuani C, et al. (2014). Functional characterization of the common c.-32–13T>G mutation of GAA gene: identification of potential therapeutic agents. Nucleic Acids Res. 42:1291–1302. https://doi.org/10.1093/nar/gkt987

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, et al. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acid Research. 37(9):e67. https://doi.org/10.1093/nar/gkp215

Hirschhorn R and Reuser A (2001). Glycogen storage disease type II: acid a-glucosidase (acid maltase) deficiency. In The Metabolic and Molecular Bases of Inherited Diseases (Scriver, C. R., Baudet, A. L. Sly, W. S, and Valle, D., eds.), pp. 3389–3420, Mc Graw-Hill, New York.

Huie ML, Chen AS, Tsujino S, Shanske S, DiMauro S, et al. (1994a). Aberrant splicing in adult onset glycogen storage disease type II (GSDII): molecular identification of an IVS1 (-13T>G) mutation in a majority of patients and a novel IVS10 (+1GT>CT) mutation. Hum.Mol. Genet. 3: 2231–2236. https://doi.org/10.1093/hmg/3.12.2231

Kemper A (2013). Evidence Report: Newborn Screening for Pompe Disease. In Book: evidence report: newborn screening for Pompe disease. Available at [http://www.hrsa.gov/advisorycommittees/mchbadvisory/heritabledisorders/nominatecondition/reviews/pompereport2013.pdf].

Kroos MA, Mullaart RA, Van Vliet L, Pomponio RJ, et al. (2008). p.[G576S; E689K]: pathogenic combination or polymorphism in Pompe disease?. Eur J Hum Genet. 16:875-879. https://doi.org/10.1038/ejhg.2008.34

Kumamoto S, Katafuchi T, Nakamura K, Endo F, et al (2009). High frequency of acid alpha-glucosidase pseudodeficiency complicates newborn screening for glycogen storage disease type II in the Japanese population. Mol Genet Metab. 97: 190-195. https://doi.org/10.1016/j.ymgme.2009.03.004

Kumar P, Henikoff S, Ng PC. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 4(7):1073-81. https://doi.org/10.1038/nprot.2009.86

Labrousse P, Chien YH, Pomponio RJ, Keutzer J, Lee NC, et al (2010). Genetic heterozygosity and pseudodeficiency in the Pompe disease newborn screening pilot program. Molecular Genetics and Metabolism. 99: 379-383. https://doi.org/10.1016/j.ymgme.2009.12.014

Li Y, Scott CR, Chamoles NA, Ghavami A, Pinto BM, et al (2004). Direct Multiplex Assay of Lysosomal Enzymes in Dried Blood Spots for Newborn Screening. Clin Chem. 50:10, 1785–1796. https://doi.org/10.1373/clinchem.2004.035907

Miller SA, Dykes DD, Polesky HF (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16: 1215. https://doi.org/10.1093/nar/16.3.1215

Neto E, Schulte J and Pereira J (2017). Evaluation of a digital microfluidic platform for newborn screening of four lysosomal storage diseases: results of a pilot project. Manuscript submitted for publication.

Oba-Shinjo SM, Silva R, Andrade FG, Palmer RE, Pomponio RJ, et al (2009). Pompe disease in a Brazilian series: clinical and molecular analyses with identification of nine new mutations. J Neurol. 256:1881-1890. https://doi.org/10.1007/s00415-009-5219-y

Schwarz JM, Cooper DN, Schuelke M and Seelow D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 11(4):361-2. https://doi.org/10.1038/nmeth.2890

Scott CR, Elliot S, Buroker N, Thomas LI, Keutzer J, et al (2013). Identification of infants at risk for developing Fabry, Pompe, or mucopolysaccharidosis-I from newborn blood spots by tandem mass spectrometry. J Pediatr. 163: 498-503.

https://doi.org/10.1016/j.jpeds.2013.01.031

Suzuki, K. (1977) Globoid cell leukodystrophy (Krabbe disease) and GM1 gangliosidosis. In: Glew, R. H. & Peters, S. P., eds. Pratical enzymology of the sphingolipidoses. New York, Alan R. Liss. https://doi.org/10.1007/3-540-27660-2_8

Tajima Y, Matsuzawa S, Aikawa S, Okumiya T, et al. (2007). Structural and biochemical studies on Pompe disease and a 'pseudodeficiency of acid alpha-glucosidase. J. Hum. Genet. 52: 898–906. https://doi.org/10.1007/s10038-007-0191-9

Turaça LT, de Faria DO, Kyosen SO, Teixeira VD, Motta FL, et al (2015). Novel GAA mutations in patients with Pompe disease. Gene. 561(1):124-131. https://doi.org/10.1016/j.gene.2015.02.023

CAPÍTULO 5: DISCUSSÃO

Desde o ano 2013, os painéis NGS têm sido aplicados no diagnóstico das doenças lisossômicas[177–181,183], demonstrando a sua maior sensibilidade devido à profundidade de cobertura alcançada, sua utilidade clínica, seu potencial para reduzir o atraso no diagnóstico e proporcionando reduções de custo e de prazo em comparação com o sequenciamento de Sanger[177]. Com a tecnologia NGS, um diagnóstico genético-molecular pode ser estabelecido em 4-6 semanas, sendo fundamental no início do curso da doença uma vez que novas terapias estão se tornando disponíveis, ademais de outros benefícios como o aconselhamento genético, planejamento familiar e a escolha de opções reprodutivas para as famílias afetadas (por exemplo, diagnóstico pré-implantacional).

O SGM-HCPA é um serviço de referência nacional e internacional no diagnóstico de erros inatos do metabolismo, sendo que a sua maior especialidade é a de doenças lisossômicas. Para essa finalidade, utilizam-se diferentes técnicas bioquímicas de triagem (qualitativas, semi-quantitativas e quantitativas), técnicas mais específicas (na maioria ensaios enzimáticos colorimétricos e fluorométricos), e a investigação laboratorial de alterações genéticas por PCR, PCR-RFLP, PCR em tempo real e o sequenciamento automatizado.

Até poucos anos atrás, as MPS eram o principal grupo de DLs caracterizadas através dos diferentes testes moleculares no nosso serviço. No ano 2014, com a aquisição do sequenciador de nova geração, Ion Torrent PGM, inicio-se o processo de desenho e validação de painéis contendo diferentes genes associados às DLs, ampliando assim o rol de doenças investigadas no SGM-HCPA além das MPS.

Neste estudo, 03 painéis NGS (“A”, “B” e “C”) foram desenhados, validados e a sua utilidade clínica avaliada para o diagnóstico de DLs selecionadas. O critério de desenho destes painéis foram as manifestações clínicas sobrepostas que incluem hepatoesplenomegalia (Painel B: Niemann-Pick A, B e C, doença de Gaucher, Deficiência da Lipase Ácida e Deficiência da Prosaposina), disfunção do tecido muscular (Painel A: doenças de Fabry, Pompe, Schindler e Danon) e disfunção neurológica (Painel C: Lipofuscinoses ceróides neuronais).

Além disso, os 24 genes associados às DLs selecionadas apresentam uma alta heterogeneidade alélica. Até o momento, 3655 mutações[88] foram relatadas nestes genes, sendo as mais frequentes as mutações *missense/nonsense*, seguidas das pequenas indels , sítios de *splicing*, grandes indels, rearranjo complexos e as regulatórias (Tabela 6). Nesses genes, as mutações pontuais (*missense*/*nonsense*/sítios de *splicing*) e pequenas indels podem chegar a representar de 83.1% (*CLN3*) até 100% (*CTSF*) das mutações totais (Tabela 6). Por isto, a análise de sequência por NGS é a abordagem ideal para obter um alto rendimento diagnóstico, devido a que os painéis NGS são especificamente sensíveis à detecção deste tipo de variantes.

No primeiro artigo, mostramos o desenho e validação de 02 painéis NGS ("A" e "B"), determinando a sensibilidade, especificidade e limitações para cada um deles, assim como a demonstração da sua utilidade clínica. A detecção das limitações desta metodologia foi crítica, pois algumas regiões não foram cobertas pelo teste (regiões com alto conteúdo GC e regiões homopoliméricas), levando a falsos-negativos (variante c.573delT/p.Ser192fs no gene *SMPD1*). Os falsos-negativos são o maior problema enfrentado no diagnóstico genético, embora a proporção de falsos-negativos e falsos-positivos seja menor que as obtidas no WES e WGS[196]. Outra limitação foi a incapacidade de detectar a variante c.[1448T>G; 1483G>C; 1497G>C]/ p.[Leu444Pro;Ala456Pro,Val460Val], presente no gene *GBA1*, devido à presença do seu pseudogene *GBAP1*, mas isto foi superado parcialmente graças a uma abordagem bioinformática. Além das limitações, a sensibilidade analítica de ambos painéis foi alta (100 % para o painel A e 93,75% para o painel B). Conforme o recomendado pela ACMG[141], foi alcançada uma sensibilidade de 100% quando complementada com o sequenciamento de Sanger. A utilidade clínica foi demonstrada após o diagnóstico de 4 casos de pacientes encaminhados ao SGM-HCPA. A validação do painel "C" (Lipofuscinoses ceróides neuronais) foi realizada, mas não foi publicada.

O desenvolvimento do NGS para o diagnóstico das doenças genéticas utilizando DNA extraído de sangue impregnado em papel filtro (SIPF) tem sido um desafio, mas, na última década, tem sido alcançado um progresso considerável[197–202]. Contudo, são poucos os estudos que utilizaram metodologias não comerciais para a extração de DNA a partir deste tipo de material, entre elas, uma baseada na utilização de hidróxido de sódio[198] e outra baseada numa solução de lise celular, detergente alcalino e tampões neutralizantes de ácido[201].

Com o NGS se implementando na rotina do SGM-HCPA, era necessário um método de baixo custo e confiável para a extração de DNA a partir do SIPF, uma vez que a coleta, transporte e armazenamento deste tipo de material é mais simples. No artigo 2, avaliamos a qualidade e eficiência do DNA extraído de SIPF para seu uso no diagnóstico molecular de doenças genéticas, entre elas as DLs. O método não comercial de fenol-clorofórmio apresentou resultados aceitáveis, com desempenho similar ao obtido com DNA extraído de sangue periférico. Este método demonstrou a sua efetividade com volumes pequenos de DNA e é compatível com futuras aplicações *downstream*, incluindo o NGS. Isto torna possível o diagnóstico genético dos pacientes cujas amostras são encaminhadas ao nosso serviço como SIPF por causa da distância da cidade ou país dos quais são remetidas. Este método também está sendo utilizado para a confirmação diagnóstica dos casos com resultado alterado na triagem neonatal para DLs (de um estudo em paralelo), realizando a extração de DNA a partir do SIPF coletado durante a TN, sem necessidade de coleta de novo material e diminuindo o tempo para o diagnóstico confirmatório.

No artigo 3, a descrição de um caso de pseudodeficiência de α-glucosidase detectado num estudo-piloto brasileiro de triagem neonatal de DLs[203], frisa a necessidade de complementar os testes bioquímicos com testes moleculares, a fim de diferenciar os indivíduos com pseudodeficiência (considerados falso-positivos[78]) daqueles portadores ou afetados pela doença de Pompe. A análise genética do paciente foi realizada utilizando nossa abordagem NGS e, o genótipo dos pais estabelecido através da análise mutação-específica por sequenciamento de Sanger. Um segundo caso alterado para Gaucher também foi analisado através do NGS, confirmando o estado de portador/heterozigoto (c.1226A>G/p.Asn409Ser) (Anexo I). Ambos pacientes apresentaram valores baixos para a atividade enzimática correspondente, mas que não correspondiam aos observados nos casos afetados. Um outro caso, proveniente do estudo paralelo realizado no SGM-HCPA para TN das DLs, com resultado alterado para doença de Pompe foi investigado, se tratando de um falso-positivo. Portanto é necessário realizar uma combinação de análises enzimáticas e moleculares para distinguir indivíduos com pseudodeficiência daqueles portadores ou afetados por uma DL, com a finalidade de tomar uma decisão sobre a introdução do tratamento específico, como a terapia de reposição enzimática, uma vez que são doenças progressivas e a introdução precoce melhora a história natural da doença.

Finalmente, no quarto artigo, foi utilizado o painel NGS "C" para a análise genético-molecular de indivíduos latino-americanos com suspeita de Lipofuscinose ceróide neuronal

tipo 2 (CLN2). Este painel tem cobertura e sensibilidade de 100% para o gene associado (*TPP1*). Nossa abordagem permitiu identificar pacientes e portadores de CLN2, assim como ampliar o conhecimento das variantes patogênicas no nosso grupo de pacientes. A análise molecular por NGS foi realizada usando DNA extraído de SIPF, extraído através do método descrito no artigo 2. A utilização de DNA extraído de SIPF foi essencial, devido à que os pacientes foram encaminhados de diferentes países latino-americanos (Argentina, Brazil, Chile, Colômbia, Equador e México), e o transporte de amostras de sangue líquido não teria sido viável em razão das barreiras alfandegárias e sanitárias, que retardam a entrega do material. Graças a este estudo, foi possível implementar no SGM-HCPA a extração de DNA partir de SIPF para seu uso no NGS. Não apenas para este painel, se não para todos os painéis para DLs disponíveis no LGM do SGM-HCPA (8 painéis contendo 56 genes associados às DLs).

Os resultados do presente estudo cumpriram todos os objetivos estabelecidos e permitiram o diagnóstico genético-molecular de 45 pacientes com suspeita de DL.

Este estudo demostra também que o diagnóstico bioquímico deve permanecer como uma parte fundamental do algoritmo diagnóstico das DLs, uma vez que garante que as variantes patogênicas detectadas sejam a causa real do fenótipo observado. Além disso, o diagnóstico bioquímico direciona a escolha do painel NGS a ser utilizado, elevando assim o rendimento diagnóstico desta abordagem no nosso serviço.

CAPÍTULO 6: CONCLUSÃO

Até onde sabemos, este estudo é o primeiro a usar o NGS no diagnóstico de DLs no Brasil. Nossa abordagem demostra que o NGS pode ser usado para a triagem paralela de múltiplos genes associados às DLs e pode identificar com sucesso variantes patogênicas de mutação pontual e pequenas indels. Este ensaio pode ser utilizado como ferramenta de suporte, que em combinação com dados bioquímicos e clínicos podem facilitar o diagnóstico.

Os painéis validados foram implementados no SGM-HCPA e são oferecidos através da "Rede DLD Brasil" para o diagnóstico molecular de DLs, o que reflete o caráter translacional deste estudo.

A mesma abordagem NGS pode ser aplicada para o diagnóstico molecular de outras doenças genéticas. Utilizando a experiência obtida no desenvolvimento deste trabalho, será realizado um projeto piloto para avaliar a utilidade clínica do NGS na confirmação diagnóstica das doenças incluídas no Programa Nacional de Triagem Neonatal.

A implementação desta tecnologia de vanguarda permitirá um diagnóstico mais rápido e mais preciso das as DLs, permitindo que milhares de pacientes e suas famílias possam se beneficiar da introdução mais precoce do tratamento específico (disponível para muitas dessas condições), bem como das medidas gerais de manejo e das medidas preventivas como a detecção de portadores, o aconselhamento genético e o diagnóstico pré-natal e pré-implantacional.

CAPÍTULO 7: REFERÊNCIAS

1. Argmann, C. A., Houten, S. M., Zhu, J. & Schadt, E. E. A Next Generation Multiscale View of Inborn Errors of Metabolism. *Cell Metab.* **23,** 13–26 (2016).

2. El-Hattab, A. W. Inborn Errors of Metabolism. *Clin. Perinatol.* **42,** 413–439 (2015).

3. Romão, A. *et al.* Apresentação Clínica Inicial dos Casos de Erros Inatos do Metabolismo de um Hospital Pediátrico de Referência: Ainda um Desafio Diagnóstico. *Rev. Paul. Pediatr.* **35,** 258–264 (2017).

4. Applegarth, D. A., Toone, J. R. & Lowry, R. B. Incidence of inborn errors of metabolism in British Columbia, 1969-1996. *Pediatrics* **105,** e10 (2000).

5. Dionisi-Vici, C. *et al.* Inborn errors of metabolism in the Italian pediatric population: a national retrospective survey. *J. Pediatr.* **140,** 321–7 (2002).

6. Sanderson, S., Green, A., Preece, M. A. & Burton, H. The incidence of inherited metabolic disorders in the West Midlands, UK. *Arch. Dis. Child.* **91,** 896–899 (2006).

7. Tiwari, S., Kallianpur, D. & DeSilva, K. A. Communication Impairments in Children with Inborn Errors of Metabolism: A Preliminary Study. *Indian J. Psychol. Med.* **39,** 146–151 (2017).

8. Deodato, F. *et al.* Inborn errors of metabolism: an update on epidemiology and on neonatal-onset hyperammonemia. *Acta Paediatr. Suppl.* **93,** 18–21 (2004).

9. Alfadhel, M. *et al.* Thirteen year retrospective review of the spectrum of inborn errors of metabolism presenting in a tertiary center in Saudi Arabia. *Orphanet J. Rare Dis.* **11,** 126 (2016).

10. Saudubray, J.-M., Sedel, F. & Walter, J. H. Clinical approach to treatable inborn metabolic diseases: an introduction. *J. Inherit. Metab. Dis.* **29,** 261–74 (2006).

11. Leonard, J. & Morris, A. Diagnosis and early management of inborn errors of metabolism presenting around the time of birth1. *Acta Paediatr.* **95,** 6–14 (2006).

12. Sutton V Reid. Inborn errors of metabolism: Classification - UpToDate. Available at: https://www.uptodate.com/contents/inborn-errors-of-metabolism-classification. (Accessed: 6th June 2018)

13. Vernon, H. J. Inborn Errors of Metabolism. *JAMA Pediatr.* **169,** 778 (2015).

14. Choi, M. *et al.* Genetic diagnosis by whole exome capture and massively parallel DNA sequencing. *Proc. Natl. Acad. Sci.* **106,** 19096–19101 (2009).

15. Sabatini, D. . and A. M. B. The Online Metabolic and Molecular Bases of Inherited Disease | OMMBID | McGraw-Hill Medical. Available at: https://ommbid.mhmedical.com/book.aspx?bookID=971. (Accessed: 11th June 2018)

16. Ferreira, C. R. & Gahl, W. A. Lysosomal storage diseases. *Transl. Sci. Rare Dis.* **2,** 1–71 (2017).

17. Hers, H. G. Inborn Lysosomal Diseases. *Gastroenterology* **48,** 625–633 (1965).

18. Platt, F. M., Boland, B. & van der Spoel, A. C. The cell biology of disease: lysosomal storage disorders: the cellular impact of lysosomal dysfunction. *J. Cell Biol.* **199,** 723–34 (2012).

19. Wang, R. Y., Bodamer, O. A., Watson, M. S., Wilcox, W. R. & ACMG Work Group on Diagnostic Confirmation of Lysosomal Storage Diseases. Lysosomal storage diseases: Diagnostic confirmation and management of presymptomatic individuals. *Genet. Med.* **13,** 457–484 (2011).

20. Lualdi, S. *et al.* Enigmatic in vivo iduronate-2-sulfatase (IDS) mutant transcript correction to wild-type in Hunter syndrome. *Hum. Mutat.* **31,** E1261-85 (2010).

21. Gieselmann, V. *Cellular pathophysiology of lysosomal storage diseases*. *Fabry Disease: Perspectives from 5 Years of FOS* (Oxford PharmaGenesis, 2006).

22. Al-Jasmi, F. A. *et al.* in *JIMD reports* **10,** 1–9 (2012).

23. Giugliani, R., Federhen, A., Michelin-Tirelli, K., Riegel, M. & Burin, M. Relative frequency and estimated minimal frequency of Lysosomal Storage Diseases in Brazil: Report from a Reference Laboratory. *Genet. Mol. Biol.* **40,** 31–39 (2017).

24. Mistry, P. K. Gaucher's disease: a model for modern management of a genetic disease. *J. Hepatol.* **30 Suppl 1,** 1–5 (1999).

25. Natowicz, M. R. & Prence, E. M. Heterozygote screening for Tay-Sachs disease: past successes and future challenges. *Curr. Opin. Pediatr.* **8,** 625–9 (1996).

26. Schuchman, E. H. & Miranda, S. R. Niemann-Pick disease: mutation update, genotype/phenotype correlations, and prospects for genetic testing. *Genet. Test.* **1,** 13–9 (1997).

27. Filocamo, M. & Morrone, A. Lysosomal storage disorders: Molecular basis and laboratory testing. *Human Genomics* **5,** 156–169 (2011).

28. Ballabio, A. & Gieselmann, V. Lysosomal disorders: From storage to cellular damage. *Biochim. Biophys. Acta - Mol. Cell Res.* **1793,** 684–696 (2009).

29. Tylki-Szymańska, A. *et al.* Non-neuronopathic Gaucher disease due to saposin C deficiency. *Clin. Genet.* **72,** 538–542 (2007).

30. Grossi, S. *et al.* Molecular analysis of ARSA and PSAP genes in twenty-one Italian patients with metachromatic leukodystrophy: identification and functional characterization of 11 novel ARSA alleles. *Hum. Mutat.* **29,** E220-30 (2008).

31. Ricketts, M. H., Goldman, D., Long, J. C. & Manowitz, P. Arylsulfatase A pseudodeficiency-associated mutations: Population studies and identification of a novel haplotype. *Am. J. Med. Genet.* **67,** 387–392 (1996).

32. Cao, Z., Petroulakis, E., Salo, T. & Triggs-Raine, B. Benign HEXA mutations, C739T(R247W) and C745T(R249W), cause beta-hexosaminidase A pseudodeficiency by reducing the alpha-subunit protein levels. *J. Biol. Chem.* **272,** 14975–82 (1997).

33. Aronovich, E. L., Pan, D. & Whitley, C. B. Molecular genetic defect underlying alpha-L-iduronidase pseudodeficiency. *Am. J. Hum. Genet.* **58,** 75–85 (1996).

34. Nishimoto, J. *et al.* A family with pseudodeficiency of acid alpha-glucosidase. *Clin. Genet.* **33,** 254–61 (1988).

35. Froissart, R. Fabry disease: D313Y is an α-galactosidase A sequence variant that causes pseudodeficient activity in plasma. *Mol. Genet. Metab.* **80,** 307–314 (2003).

36. Hoffmann, B., Georg Koch, H., Schweitzer-Krantz, S., Wendel, U. & Mayatepek, E. Deficient α-galactosidase A activity in plasma but no Fabry disease – a pitfall in diagnosis. *Clin. Chem. Lab. Med.* **43,** (2005).

37. Gort, L., Santamaria, R., Grinberg, D., Vilageliu, L. & Chabás, A. Identification of a novel pseudodeficiency allele in the *GLB1* gene in a carrier of GM1 gangliosidosis. *Clin. Genet.* **72,** 109–111 (2007).

38. Wauters, J. G., Stuer, K. L., Van Elsen, A. & Willems, P. J. alpha-L-fucosidase in human fibroblasts. I. The enzyme activity polymorphism. *Biochem. Genet.* **30,** 131–41 (1992).

39. Chabas, A., Giros, M. L. & Guardiola, A. Low beta-glucuronidase activity in a healthy member of a family with mucopolysaccharidosis VII. *J. Inherit. Metab. Dis.* **14,** 908–14 (1991).

40. Vervoort, R. *et al.* A pseudodeficiency allele (D152N) of the human beta-glucuronidase gene. *Am. J. Hum. Genet.* **57,** 798–804 (1995).

41. Gieselmann, V. An assay for the rapid detection of the arylsulfatase A pseudodeficiency allele facilitates diagnosis and genetic counseling for metachromatic leukodystrophy. *Hum. Genet.* **86,** 251–5 (1991).

42. Chabás, A. *et al.* Frequency of the arylsulphatase A pseudodeficiency allele in the Spanish population. *Clin. Genet.* **44,** 320–3 (1993).

43. Regis, S., Filocamo, M., Stroppiano, M., Corsolini, F. & Gatti, R. Molecular analysis of the arylsulphatase A gene in late infantile metachromatic leucodystrophy patients and healthy subjects from Italy. *J. Med. Genet.* **33,** 251–2 (1996).

44. Kemper, A. R. *et al.* Evidence report: Newborn screening for pompe disease. (2013).

45. Labrousse, P. *et al.* Genetic heterozygosity and pseudodeficiency in the Pompe disease newborn screening pilot program. *Mol. Genet. Metab.* **99,** 379–83 (2010).

46. Martins, A. M. *et al.* Demographic characterization of Brazilian patients enrolled in the Fabry Registry. *Genet. Mol. Res.* **12,** 136–142 (2013).

47. Musumeci, O. *et al.* LOPED study: looking for an early diagnosis in a late-onset Pompe disease high-risk population. *J. Neurol. Neurosurg. Psychiatry* **87,** jnnp-2014-310164 (2015).

48. Montagnese, F. *et al.* Clinical and molecular aspects of 30 patients with late-onset Pompe disease (LOPD): unusual features and response to treatment. *J. Neurol.* **262,** 968–978 (2015).

49. Vieira, T. *et al.* Mucopolysaccharidoses in Brazil: What happens from birth to biochemical diagnosis? *Am. J. Med. Genet. Part A* **146A,** 1741–1747 (2008).

50. Parenti, G., Andria, G. & Ballabio, A. Lysosomal Storage Diseases: From Pathophysiology to Therapy. *Annu. Rev. Med.* **66,** 471–486 (2015).

51. Boustany, R.-M. N. Lysosomal storage diseases—the horizon expands. *Nat. Rev. Neurol.* **9,** 583–598 (2013).

52. Kingma, S. D. K., Bodamer, O. A. & Wijburg, F. A. Epidemiology and diagnosis of lysosomal storage disorders; challenges of screening. *Best Pract. Res. Clin.*

Endocrinol. Metab. **29,** 145–157 (2015).

53. Kelly, J. M., Bradbury, A., Martin, D. R. & Byrne, M. E. Emerging therapies for neuropathic lysosomal storage disorders. *Prog. Neurobiol.* **152,** 166–180 (2017).

54. Schielen, P. C. J. I., Kemper, E. A. & Gelb, M. H. Newborn Screening for Lysosomal Storage Diseases: A Concise Review of the Literature on Screening Methods, Therapeutic Possibilities and Regional Programs. *Int. J. neonatal Screen.* **3,** (2017).

55. Muenzer, J. *et al.* A phase I/II study of intrathecal idursulfase-IT in children with severe mucopolysaccharidosis II. *Genet. Med.* **18,** 73–81 (2016).

56. Munoz-Rojas, M.-V. *et al.* Intrathecal enzyme replacement therapy in a patient with mucopolysaccharidosis type I and symptomatic spinal cord compression. *Am. J. Med. Genet. Part A* **146A,** 2538–2544 (2008).

57. Wasserstein, M. P. *et al.* Clinical outcomes of children with abnormal newborn screening results for Krabbe disease in New York State. *Genet. Med.* **18,** 1235–1243 (2016).

58. Aldenhoven, M. *et al.* Hematopoietic Cell Transplantation for Mucopolysaccharidosis Patients Is Safe and Effective: Results after Implementation of International Guidelines. *Biol. Blood Marrow Transplant.* **21,** 1106–1109 (2015).

59. Biffi, A. *et al.* Lentiviral Hematopoietic Stem Cell Gene Therapy Benefits Metachromatic Leukodystrophy. *Science (80-.).* **341,** 1233158–1233158 (2013).

60. Wyatt, K. *et al.* The effectiveness and cost-effectiveness of enzyme and substrate replacement therapies: a longitudinal cohort study of people with lysosomal storage disorders. *Health Technol. Assess. (Rockv).* **16,** 1–543 (2012).

61. Parenti, G., Andria, G. & Valenzano, K. J. Pharmacological Chaperone Therapy: Preclinical Development, Clinical Translation, and Prospects for the Treatment of Lysosomal Storage Disorders. *Mol. Ther.* **23,** 1138–1148 (2015).

62. Parenti, G., Moracci, M., Fecarotta, S. & Andria, G. Pharmacological chaperone therapy for lysosomal storage diseases. *Future Med. Chem.* **6,** 1031–1045 (2014).

63. Beck, M. New therapeutic options for lysosomal storage disorders: enzyme replacement, small molecules and gene therapy. *Hum. Genet.* **121,** 1–22 (2007).

64. Geraets, R. D. *et al.* Moving towards effective therapeutic strategies for Neuronal Ceroid Lipofuscinosis. *Orphanet J. Rare Dis.* **11,** 40 (2016).

65. Therrell, B. L. *et al.* Current status of newborn screening worldwide: 2015. *Semin. Perinatol.* **39,** 171–187 (2015).

66. Vargas, P. R., Gtatn, –, Ms, /, Pimentel, H. & Santos, D. Programa Nacional de Triagem Neonatal: Ofi cinas Regionais de Qualifi cação da Gestão.

67. Da, M., Brasília, S. & Df, –. TRIAGEM NEONATAL TRIAGEM NEONATAL Manual de Normas Técnicas e Rotinas Operacionais do Programa Nacional de Triagem Neonatal.

68. Triagem Neonatal Biológica: Manual Técnico.

69. Giugliani, R. Newborn screening for lysosomal diseases: current status and potential

interface with population medical genetics in Latin America. *J. Inherit. Metab. Dis.* **35,** 871–877 (2012).

70. Mechtler, T. P. *et al.* Neonatal screening for lysosomal storage disorders: feasibility and incidence from a nationwide study in Austria. *Lancet (London, England)* **379,** 335–41 (2012).

71. Chien, Y.-H. *et al.* Early Detection of Pompe Disease by Newborn Screening Is Feasible: Results From the Taiwan Screening Program. *Pediatrics* **122,** e39–e45 (2008).

72. Liao, H.-C. *et al.* Detecting multiple lysosomal storage diseases by tandem mass spectrometry--a national newborn screening program in Taiwan. *Clin. Chim. Acta.* **431,** 80–6 (2014).

73. Lin, S.-P. *et al.* A pilot newborn screening program for Mucopolysaccharidosis type I in Taiwan. *Orphanet J. Rare Dis.* **8,** 147 (2013).

74. Spada, M. *et al.* High Incidence of Later-Onset Fabry Disease Revealed by Newborn Screening*. *Am. J. Hum. Genet.* **79,** 31–40 (2006).

75. Paciotti, S. *et al.* First pilot newborn screening for four lysosomal storage diseases in an Italian region: identification and analysis of a putative causative mutation in the GBA gene. *Clin. Chim. Acta.* **413,** 1827–31 (2012).

76. Wittmann, J. *et al.* Newborn screening for lysosomal storage disorders in hungary. *JIMD Rep.* **6,** 117–25 (2012).

77. Scott, C. R. *et al.* Identification of Infants at Risk for Developing Fabry, Pompe, or Mucopolysaccharidosis-I from Newborn Blood Spots by Tandem Mass Spectrometry. *J. Pediatr.* **163,** 498–503 (2013).

78. Hopkins, P. V *et al.* Lysosomal storage disorder screening implementation: findings from the first six months of full population pilot testing in Missouri. *J. Pediatr.* **166,** 172–7 (2015).

79. Elliott, S. *et al.* Pilot study of newborn screening for six lysosomal storage diseases using Tandem Mass Spectrometry. *Mol. Genet. Metab.* **118,** 304–309 (2016).

80. Matern, D. *et al.* Newborn screening for lysosomal storage disorders. *Semin. Perinatol.* **39,** 206–216 (2015).

81. Heese, B. A. Current strategies in the management of lysosomal storage diseases. *Semin. Pediatr. Neurol.* **15,** 119–26 (2008).

82. Acosta, A. X., Abé-Sandes, K., Giugliani, R. & Bittles, A. H. Delivering genetic education and genetic counseling for rare diseases in rural Brazil. *J. Genet. Couns.* **22,** 830–4 (2013).

83. Parenti, G. Treating lysosomal storage diseases with pharmacological chaperones: from concept to clinics. *EMBO Mol. Med.* **1,** 268–79 (2009).

84. Wasserstein, M. P. & Mcgovern, M. M. Genetic basis of the lipid storage disorders. *Futur. Lipidol* **3,** 189–201 (2008).

85. Amartino, H. *et al.* Identification of 17 novel mutations in 40 Argentinean unrelated families with mucopolysaccharidosis type II (Hunter syndrome). *Mol. Genet. Metab.*

Reports **1,** 401–406 (2014).

86. Barranger, J. A. & Cabrera-Salazar, M. A. *Lysosomal storage disorders*. (Springer, 2007).

87. Curiati, M. A. *et al.* The Challenge of Diagnosis and Indication for Treatment in Fabry Disease. *J. Inborn Errors Metab. Screen.* **5,** 232640981668573 (2017).

88. Stenson et al. The Human Gene Mutation Database (HGMD®): 2003 Update. *Hum Mutat* 21:577-581 (2003). Available at: http://www.hgmd.cf.ac.uk/ac/all.php. (Accessed: 21st May 2018)

89. Frankish, A. & Harrow, J. in *Methods in molecular biology (Clifton, N.J.)* **1167,** 129–155 (2014).

90. Lualdi, S. *et al.* Characterization of iduronate-2-sulfatase gene-pseudogene recombinations in eight patients with Mucopolysaccharidosis type II revealed by a rapid PCR-based method. *Hum. Mutat.* **25,** 491–7 (2005).

91. Tayebi, N. *et al.* Reciprocal and nonreciprocal recombination at the glucocerebrosidase gene region: implications for complexity in Gaucher disease. *Am. J. Hum. Genet.* **72,** 519–34 (2003).

92. Filocamo, M. *et al.* A simple non-isotopic method to show pitfalls during mutation analysis of the glucocerebrosidase gene. *J. Med. Genet.* **38,** E34 (2001).

93. Horowitz, M. *et al.* The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics* **4,** 87–96 (1989).

94. REINER, O., WIGDERSON, M. & HOROWITZ, M. Structural Analysis of the Human Glucocerebrosidase Genes. *DNA* **7,** 107–116 (1988).

95. Martínez-Arias, R. *et al.* Sequence variability of a human pseudogene. *Genome Res.* **11,** 1071–85 (2001).

96. HONG, C. M., OHASHI, T., YU, X. J., WEILER, S. & BARRANGER, J. A. Sequence of Two Alleles Responsible for Gaucher Disease. *DNA Cell Biol.* **9,** 233–241 (1990).

97. LATHAM, T. E., THEOPHILUS, B. D. M., GRABOWSKI, G. A. & SMITH, F. I. Heterogeneity of Mutations in the Acid β-Glucosidase Gene of Gaucher Disease Patients. *DNA Cell Biol.* **10,** 15–21 (1991).

98. Cormand, B., Díaz, A., Grinberg, D., Chabás, A. & Vilageliu, L. A new gene-pseudogene fusion allele due to a recombination in intron 2 of the glucocerebrosidase gene causes Gaucher disease. *Blood Cells. Mol. Dis.* **26,** 409–16 (2000).

99. Strasberg, P. M. *et al.* Homozygous presence of the crossover (fusion gene) mutation identified in a type II Gaucher disease fetus: is this analogous to the Gaucher knock-out mouse model? *Biochem. Med. Metab. Biol.* **53,** 16–21 (1994).

100. Sidransky, E. *et al.* The clinical, molecular, and pathological characterisation of a family with two cases of lethal perinatal type 2 Gaucher disease. *J. Med. Genet.* **33,** 132–6 (1996).

101. Finn, L. S., Zhang, M., Chen, S. H. & Scott, C. R. Severe type II Gaucher disease with ichthyosis, arthrogryposis and neuronal apoptosis: molecular and pathological

analyses. *Am. J. Med. Genet.* **91,** 222–6 (2000).

102. Hruska, K. S., LaMarca, M. E., Scott, C. R. & Sidransky, E. Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum. Mutat.* **29,** 567–583 (2008).

103. Torralba, M. A. *et al.* High prevalence of the 55-bp deletion (c.1263del55) in exon 9 of the glucocerebrosidase gene causing misdiagnosis (for homozygous N370S (c.1226A > G) mutation) in Spanish Gaucher disease patients. *Blood Cells. Mol. Dis.* **29,** 35–40

104. Zimran, A. *et al.* A glucocerebrosidase fusion gene in Gaucher disease. Implications for the molecular anatomy, pathogenesis, and diagnosis of this disorder. *J. Clin. Invest.* **85,** 219–22 (1990).

105. Beutler, E., Gelbart, T. & West, C. Identification of Six New Gaucher Disease Mutations. *Genomics* **15,** 203–205 (1993).

106. Schiffmann, R. Therapeutic approaches for neuronopathic lysosomal storage disorders. *J. Inherit. Metab. Dis.* **33,** 373–379 (2010).

107. Ouesleti, S. *et al.* Update of the spectrum of mucopolysaccharidoses type III in Tunisia: identification of three novel mutations and in silico structural analysis of the missense mutations. *World J. Pediatr.* **13,** 374–380 (2017).

108. Atçeken, N. *et al.* Evaluation and identification of IDUA gene mutations in Turkishpatients with mucopolysaccharidosis type I. *Turkish J. Med. Sci.* **46,** 404–8 (2016).

109. Zhang, H. *et al.* Analysis of the IDS gene in 38 patients with Hunter syndrome: the c.879G>A (p.Gln293Gln) synonymous variation in a female create exonic splicing. *PLoS One* **6,** e22951 (2011).

110. Karageorgos, L. *et al.* Mutational analysis of 105 mucopolysaccharidosis type VI patients. *Hum. Mutat.* **28,** 897–903 (2007).

111. Chiong, M. A. D. *et al.* Clinical, biochemical and molecular characteristics of Filipino patients with mucopolysaccharidosis type II - Hunter syndrome. *Orphanet J. Rare Dis.* **12,** 7 (2017).

112. Vrijenhoek, T. *et al.* Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects. *Eur. J. Hum. Genet.* **23,** 1270–1270 (2015).

113. Lualdi, S. *et al.* Identification of nine new IDS alleles in mucopolysaccharidosis II. Quantitative evaluation by real-time RT-PCR of mRNAs sensitive to nonsense-mediated and nonstop decay mechanisms. *Biochim. Biophys. Acta* **1762,** 478–84 (2006).

114. Caciotti, A. *et al.* Optimizing the molecular diagnosis of GALNS: novel methods to define and characterize Morquio-A syndrome-associated mutations. *Hum. Mutat.* **36,** 357–68 (2015).

115. Mishori-Dery, A. *et al.* RFLPs for linkage analysis in families with glycogen storage disease type III. *J. Inherit. Metab. Dis.* **18,** 207–10 (1995).

116. Schulz, A., Kohlschütter, A., Mink, J., Simonati, A. & Williams, R. NCL diseases — clinical perspectives. *Biochim. Biophys. Acta - Mol. Basis Dis.* **1832,** 1801–1806 (2013).

117. Williams, J. C. Restriction Fragment Length Polymorphisms (RFLPs) in the Diagnosis of Metabolic Disease. *Scand. J. Clin. Lab. Invest.* **48,** 74–74 (1988).

118. Lewis, B. D., Nelson, P. V., Robertson, E. F. & Morris, C. P. Mutation analysis of 28 Gaucher disease patients: The Australasian experience. *Am. J. Med. Genet.* **49,** 218–223 (1994).

119. Kroos, M. A., Van der Kraan, M., Van Diggelen, O. P., Kleijer, W. J. & Reuser, A. J. J. Two extremes of the clinical spectrum of glycogen storage disease type II in one family: A matter of genotype. *Hum. Mutat.* **9,** 17–22 (1997).

120. Ekstein, J. *et al.* Mutation Frequencies for Glycogen Storage Disease Ia in the Ashkenazi Jewish Population. *Am. J. Med. Genet.* **129,** 162–164 (2004).

121. PARVARI, R., HERSHKOVITZ, E., CARMI, R. & MOSES, S. PRENATAL DIAGNOSIS OF GLYCOGEN STORAGE DISEASE TYPE 1a BY SINGLE STRANDED CONFORMATION POLYMORPHISM (SSCP). *Prenat. Diagn.* **16,** 862–865 (1996).

122. Maddox, L. O., L1, P., Bennett, A., Descartes, M. & Thompson, J. N. COMPARISON OF SSCP ANALYSIS AND CFLP ANALYSIS FOR MUTATION DETECTION IN THE HUMAN IDURONATE 2-SULFATASE GENE. *Biochem. aria Mol. Biol. Int. Pages* **43,** 1163–1171 (1997).

123. Marcolongo, P. *et al.* Structure and mutation analysis of the glycogen storage disease type 1b gene. *FEBS Lett.* **436,** 247–250 (1998).

124. Santer, R. *et al.* Molecular analysis in glycogen storage disease 1non-A: dHPLC detection of the highly prevalent exon 8 mutations of theG6PT1 gene in German patients. *Hum. Mutat.* **16,** 177–177 (2000).

125. Gucciardi, A. *et al.* A column-switching HPLC-MS/MS method for mucopolysaccharidosis type I analysis in a multiplex assay for the simultaneous newborn screening of six lysosomal storage disorders. *Biomed. Chromatogr.* **28,** 1131–1139 (2014).

126. Hermans, M. M. P., van Leenen, D., Kroos, M. A. & Reuser, A. J. J. Mutation Detection in Glycogen Storage Disease Type II by RT-PCR and Automated Sequencing. *Biochem. Biophys. Res. Commun.* **241,** 414–418 (1997).

127. Ezgu, F. *et al.* Rapid Molecular Diagnosis of Genetic Diseases by High Resolution Melting Analysis: Fabry and Glycogen Storage 1A Diseases. *Genet. Test. Mol. Biomarkers* **18,** 3–7 (2014).

128. Galvis, J., González, J., Uribe, A. & Velasco, H. Deep Genotyping of the IDS Gene in Colombian Patients with Hunter Syndrome. *JIMD Rep.* **19,** 101–9 (2015).

129. Jurecka, A. *et al.* Mucopolysaccharidosis type II in females and response to enzyme replacement therapy. *Am. J. Med. Genet. Part A* **158A,** 450–454 (2012).

130. Breen, C. *et al.* Maternal mosaicism for IDUA deletion clarifies recurrence risk in MPS I. *Hum. genome Var.* **3,** 16031 (2016).

131. Schirinzi, A. *et al.* Identification of GLA gene deletions in Fabry patients by Multiplex Ligation-dependent Probe Amplification (MLPA). *Mol. Genet. Metab.* **94,** 382–385 (2008).

132. Yoshimitsu, M. *et al.* Identification of novel mutations in the α-galactosidase A gene in patients with Fabry disease: Pitfalls of mutation analyses in patients with low α-galactosidase A activity. *J. Cardiol.* **57,** 345–353 (2011).

133. Pollard, L. M., Jones, J. R. & Wood, T. C. Molecular characterization of 355 mucopolysaccharidosis patients reveals 104 novel mutations. *J. Inherit. Metab. Dis.* **36,** 179–87 (2013).

134. Jahnova, H. *et al.* Observational, retrospective study of a large cohort of patients with Niemann-Pick disease type C in the Czech Republic: a surprisingly stable diagnostic rate spanning almost 40 years. *Orphanet J. Rare Dis.* **9,** 140 (2014).

135. Bauer, P. *et al.* Genetic screening for Niemann-Pick disease type C in adults with neurological and psychiatric symptoms: findings from the ZOOM study. *Hum. Mol. Genet.* **22,** 4349–4356 (2013).

136. Brusius-Facchin, A. C. *et al.* Severe phenotype in MPS II patients associated with a large deletion including contiguous genes. *Am. J. Med. Genet. A* **158A,** 1055–9 (2012).

137. Zanetti, A. *et al.* A Hunter Patient with a Severe Phenotype Reveals Two Large Deletions and Two Duplications Extending 1.2 Mb Distally to IDS Locus. *JIMD Rep.* **17,** 13–21 (2014).

138. Ittiwut, C. *et al.* Novel Mutations, Including a Large Deletion in the ARSB Gene, Causing Mucopolysaccharidosis Type VI. *Genet. Test. Mol. Biomarkers* **21,** 58–62 (2017).

139. Mardis, E. R. The impact of next-generation sequencing technology on genetics. *Trends Genet.* **24,** 133–141 (2008).

140. Lionel, A. C. *et al.* Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. *Genet. Med.* **20,** 435–443 (2018).

141. Rehm, H. L. *et al.* ACMG clinical laboratory standards for next-generation sequencing. *Genet. Med.* **15,** 733–747 (2013).

142. Strande, N. T. *et al.* Evaluating the Clinical Validity of Gene-Disease Associations: An Evidence-Based Framework Developed by the Clinical Genome Resource. *Am. J. Hum. Genet.* **100,** 895–906 (2017).

143. Santani, A. *et al.* Development and Validation of Targeted Next-Generation Sequencing Panels for Detection of Germline Variants in Inherited Diseases. *Arch Pathol Lab Med* **1415858,** (2017).

144. Rothberg, J. M. *et al.* An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475,** 348–352 (2011).

145. Varuzza, L. Introdução à análise de dados de sequenciadores de nova geração Versão 2.0.1. (2013).

146. Zeng, F., Jiang, R. & Chen, T. PyroHMMsnp: an SNP caller for Ion Torrent and 454 sequencing data. *Nucleic Acids Res.* **41,** e136–e136 (2013).

147. Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **38,** 1767–71 (2010).

148. Roy, S. *et al.* Standards and Guidelines for Validating Next- Generation Sequencing Bioinformatics Pipelines A Joint Recommendation of the Association for Molecular Pathology and the College of American Pathologists From the Next Generation Sequencing Bioinformatics Pipeline Validation Working Group of the Clinical Practice Committee,* The Association for Molecular Pathology. (2018). doi:10.1016/j.jmoldx.2017.11.003

149. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17,** 405–423 (2015).

150. Richards, C. S. *et al.* ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet. Med.* **10,** 294–300 (2008).

151. Gibbs, R. A. *et al.* A global reference for human genetic variation. *Nature* **526,** 68–74 (2015).

152. Karczewski, K. J. *et al.* The ExAC browser: displaying reference data information from over 60 000 exomes. *Nucleic Acids Res.* **45,** D840–D845 (2017).

153. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29,** 308–11 (2001).

154. Naslavsky, M. S. *et al.* Exomic variants of an elderly cohort of Brazilians in the ABraOM database. *Hum. Mutat.* **38,** 751–763 (2017).

155. Kumar, P., Henikoff, S. & Ng, P. C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat. Protoc.* **4,** 1073–1081 (2009).

156. Sim, N.-L. *et al.* SIFT web server: predicting effects of amino acid substitutions on proteins. *Nucleic Acids Res.* **40,** W452–W457 (2012).

157. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7,** 248–249 (2010).

158. Schwarz, J. M., Cooper, D. N., Schuelke, M. & Seelow, D. MutationTaster2: mutation prediction for the deep-sequencing age. *Nat. Methods* **11,** 361–362 (2014).

159. Desmet, F.-O. *et al.* Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res.* **37,** e67 (2009).

160. Hu, J. & Ng, P. C. SIFT Indel: predictions for the functional effects of amino acid insertions/deletions in proteins. *PLoS One* **8,** e77940 (2013).

161. Bandrés-Ciga, S. *et al.* Analysis of the genetic variability in Parkinson's disease from Southern Spain. *Neurobiol. Aging* **37,** 210.e1-210.e5 (2016).

162. Asselta, R. *et al.* Glucocerebrosidase mutations in primary parkinsonism.

Parkinsonism Relat. Disord. **20,** 1215–1220 (2014).

163. Török, R. *et al.* An assessment of the frequency of mutations in the GBA and VPS35 genes in Hungarian patients with sporadic Parkinson's disease. *Neurosci. Lett.* **610,** 135–8 (2016).

164. Han, F. *et al.* Mutations in the glucocerebrosidase gene are common in patients with Parkinson's disease from Eastern Canada. *Int. J. Neurosci.* **126,** 415–21 (2016).

165. Yu, Z. *et al.* Mutations in the glucocerebrosidase gene are responsible for Chinese patients with Parkinson's disease. *J. Hum. Genet.* **60,** 85–90 (2015).

166. Pulkes, T. *et al.* Glucocerebrosidase mutations in Thai patients with Parkinson's disease. *Parkinsonism Relat. Disord.* **20,** 986–991 (2014).

167. Mitsui, J. *et al.* Mutations for Gaucher disease confer high susceptibility to Parkinson disease. *Arch. Neurol.* **66,** 571–6 (2009).

168. Mata, I. F. *et al.* GBA Variants are associated with a distinct pattern of cognitive deficits in Parkinson's disease. *Mov. Disord.* **31,** 95–102 (2016).

169. Zampieri, S., Cattarossi, S., Bembi, B. & Dardis, A. GBA Analysis in Next-Generation Era: Pitfalls, Challenges, and Possible Solutions. *J. Mol. Diagnostics* **19,** 733–741 (2017).

170. Mason-Suares, H., Landry, L. & S. Lebo, M. Detecting Copy Number Variation via Next Generation Technology. *Curr. Genet. Med. Rep.* **4,** 74–85 (2016).

171. Liu, Q., Zhang, P., Wang, D., Gu, W. & Wang, K. Interrogating the "unsequenceable" genomic trinucleotide repeat disorders by long-read sequencing. *Genome Med.* **9,** 65 (2017).

172. Baudhuin, L. M. *et al.* Confirming Variants in Next-Generation Sequencing Panel Testing by Sanger Sequencing. *J. Mol. Diagnostics* **17,** 456–461 (2015).

173. Nelson, A. C. *et al.* Criteria for Clinical Reporting of Variants from a Broad Target Capture NGS Assay without Sanger Verification. (2015).

174. Mu, W., Lu, H.-M., Chen, J., Li, S. & Elliott, A. M. Sanger Confirmation Is Required to Achieve Optimal Sensitivity and Specificity in Next-Generation Sequencing Panel Testing. *J. Mol. Diagnostics* **18,** 923–932 (2016).

175. Lacbawan, F. L. *et al.* Verification of Performance Specifications of a Molecular Test: Cystic Fibrosis Carrier Testing Using the Luminex Liquid Bead Array. *Arch. Pathol. Lab. Med.* **136,** 14–19 (2012).

176. Jennings, L., Van Deerlin, V. M., Gulley, M. L. & College of American Pathologists Molecular Pathology Resource Committee. Recommended principles and practices for validating clinical molecular pathology tests. *Arch. Pathol. Lab. Med.* **133,** 743–55 (2009).

177. Fernández-Marmiesse, A. *et al.* Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. *Orphanet J. Rare Dis.* **9,** 59 (2014).

178. Lévesque, S. *et al.* Diagnosis of late-onset Pompe disease and other muscle disorders by next-generation sequencing. *Orphanet J. Rare Dis.* **11,** 8 (2016).

179. Yubero, D. *et al.* Targeted Next Generation Sequencing in Patients with Inborn Errors of Metabolism. *PLoS One* **11,** e0156359 (2016).

180. Wood, T. *et al.* Development of a next generation sequencing panel for lysosomal storage disorders. *Mol. Genet. Metab.* **108,** S100 (2013).

181. Coutinho, M. F. *et al.* Development of a Next-Generation Sequencing (NGS) Gene Panel for Lysosomal Storage Diseases. *13th Int. Symp. SPDM (Sociedade Port. Doenças Metabólicas), 16-17 March 2017* (2017).

182. Matthijs, G., Dierking, A. & Schmidtke, J. New EuroGentest/ESHG guidelines and a new clinical utility gene card format for NGS-based testing. *Eur. J. Hum. Genet.* **24,** 1–1 (2016).

183. Di Fruscio, G. *et al.* Lysoplex: An efficient toolkit to detect DNA sequence variations in the autophagy-lysosomal pathway. *Autophagy* **11,** 928–38 (2015).

184. Ng, S. B., Nickerson, D. A., Bamshad, M. J. & Shendure, J. Massively parallel sequencing and rare disease. *Hum. Mol. Genet.* **19,** R119–R124 (2010).

185. Lelieveld, S. H., Spielmann, M., Mundlos, S., Veltman, J. A. & Gilissen, C. Comparison of Exome and Genome Sequencing Technologies for the Complete Capture of Protein-Coding Regions. *Hum. Mutat.* **36,** 815–22 (2015).

186. Meienberg, J., Bruggmann, R., Oexle, K. & Matyas, G. Clinical sequencing: is WGS the better WES? *Hum. Genet.* **135,** 359–362 (2016).

187. Bonafé, L. *et al.* Brief Report: Peripheral Osteolysis in Adults Linked to ASAH1 (Acid Ceramidase) Mutations: A New Presentation of Farber's Disease. *Arthritis Rheumatol. (Hoboken, N.J.)* **68,** 2323–7 (2016).

188. Kondo, H. *et al.* Mutation in VPS33A affects metabolism of glycosaminoglycans: a new type of mucopolysaccharidosis with severe systemic symptoms. *Hum. Mol. Genet.* **26,** 173–183 (2017).

189. Yu, T. W. *et al.* Using Whole-Exome Sequencing to Identify Inherited Causes of Autism. *Neuron* **77,** 259–273 (2013).

190. Prada, C. E. *et al.* Clinical utility of whole-exome sequencing in rare diseases: Galactosialidosis. *Eur. J. Med. Genet.* **57,** 339–344 (2014).

191. Zeng, Q. *et al.* Molecular defects identified by whole exome sequencing in a child with atypical mucopolysaccharidosis IIIB. *J. Pediatr. Endocrinol. Metab.* **30,** 463–469 (2017).

192. Nikkel, S. M. *et al.* Whole-exome sequencing expands the phenotype of Hunter syndrome. *Clin. Genet.* **86,** 172–176 (2014).

193. Kaissi, A. Al *et al.* The constellation of skeletal deformities in a family with mixed types of mucopolysaccharidoses: Case report. *Medicine (Baltimore).* **95,** e4561 (2016).

194. Li, J., Xie, H. & Jiang, Y. Mucopolysaccharidosis IIIB and mild skeletal anomalies: coexistence of NAGLU and CYP26B1 missense variations in the same patient in a Chinese family. *BMC Med. Genet.* **19,** 51 (2018).

195. *PORTARIA No-3.253, DE 29 DE DEZEMBRO DE 2016PORTARIA No-3.253, DE 29*

DE DEZEMBRO DE 2016. (2016).

196. Wong, L.-J. C. Next generation molecular diagnosis of mitochondrial disorders. *Mitochondrion* **13,** 379–387 (2013).

197. Bassaganyas, L. *et al.* Whole exome and whole genome sequencing with dried blood spot DNA without whole genome amplification. *Hum. Mutat.* **39,** 167–171 (2018).

198. Lefterova, M. I. *et al.* Next-Generation Molecular Testing of Newborn Dried Blood Spots for Cystic Fibrosis. *J. Mol. Diagnostics* **18,** 267–282 (2016).

199. Boemer, F. *et al.* A next-generation newborn screening pilot study: NGS on dried blood spots detects causal mutations in patients with inherited metabolic diseases. *Sci. Rep.* **7,** 17641 (2017).

200. Hollegaard, M. V. *et al.* Archived neonatal dried blood spot samples can be used for accurate whole genome and exome-targeted next-generation sequencing. *Mol. Genet. Metab.* **110,** 65–72 (2013).

201. Saavedra-Matiz, C. A. *et al.* Cost-effective and scalable DNA extraction method from dried blood spots. *Clin. Chem.* **59,** 1045–51 (2013).

202. Qian, J. *et al.* Applying targeted next generation sequencing to dried blood spot specimens from suspicious cases identified by tandem mass spectrometry-based newborn screening. *J. Pediatr. Endocrinol. Metab.* **30,** 979–988 (2017).

203. Camargo Neto, E. *et al.* Neonatal screening for four lysosomal storage diseases with a digital microfluidics platform: Initial results in Brazil. *Genet. Mol. Biol.* 0 (2018). doi:10.1590/1678-4685-GMB-2017-0227

ANEXO I – Artigo 5: Investigation of newborns with abnormal results in a newborn screening program for four lysosomal storage diseases in Brazil

---

Publicado na revista *MOLECULAR GENETICS AND METABOLISM*, 2017

Molecular Genetics and Metabolism Reports 12 (2017) 92–97

Contents lists available at ScienceDirect

Molecular Genetics and Metabolism Reports

journal homepage: www.elsevier.com/locate/ymgmr

Investigation of newborns with abnormal results in a newborn screening program for four lysosomal storage diseases in Brazil

Heydy Bravo[a], Eurico Camargo Neto[b], Jaqueline Schulte[b], Jamile Pereira[b], Claudio Sampaio Filho[c], Fernanda Bittencourt[d], Fernanda Sebastião[d], Fernanda Bender[d], Ana Paula Scholz de Magalhães[d], Régis Guidobono[d], Franciele Barbosa Trapp[d], Kristiane Michelin-Tirelli[d], Carolina F.M. Souza[d], Diana Rojas Málaga[a], Gabriela Pasqualim[a,f], Ana Carolina Brusius-Facchin[d], Roberto Giugliani[a,d,e,f,*]

[a] Post-Graduate Program of Genetics and Molecular Biology, UFRGS, Av. Bento Gonçalves 9500, 91501-970 Porto Alegre, RS, Brazil
[b] CTN Diagnósticos, Av. Ipiranga 5000, Porto Alegre, RS 90610-000, Brazil
[c] Intercientifica, Av. Ouro Fino, São José dos Campos, SP 12233-401, Brazil
[d] Medical Genetics Service, HCPA, Rua Ramiro Barcelos 2350, Porto Alegre, RS 90035-903, Brazil
[e] Post-Graduate Program in Medical Sciences, UFRGS, Rua Ramiro Barcelos 2400, 90035-003 Porto Alegre, RS, Brazil
[f] Gene Therapy Center, Experimental Research Center, HCPA, Rua Ramiro Barcelos 2350, Porto Alegre, RS 90035-903, Brazil

ARTICLE INFO

Keywords:
Lysosomal storage diseases
Newborn screening
Confirmatory diagnosis
Carrier
Pseudodeficiency
Brazil

ABSTRACT

Lysosomal storage diseases (LSDs) are genetic disorders, clinically heterogeneous, mainly caused by defects in genes encoding lysosomal enzymes that degrade macromolecules. Several LSDs already have specific therapies that may improve clinical outcomes, especially if introduced early in life. With this aim, screening methods have been established and newborn screening (NBS) for some LSDs has been developed. Such programs should include additional procedures for the confirmation (or not) of the cases that had an abnormal result in the initial screening. We present here the methods and results of the additional investigation performed in four babies with positive initial screening results in a program of NBS for LSDs performed by a private laboratory in over 10,000 newborns in Brazil. The suspicion in these cases was of Mucopolysaccharidosis I - MPS I (in two babies), Pompe disease and Gaucher disease (one baby each). One case of pseudodeficiency for MPS I, 1 carrier for MPS I, 1 case of pseudodeficiency for Pompe disease and 1 carrier for Gaucher disease were identified. This report illustrates the challenges that may be encountered by NBS programs for LSDs, and the need of a comprehensive protocol for the rapid and precise investigation of the babies who have an abnormal screening result.

1. Introduction

Lysosomal storage diseases (LSDs) are genetic disorders with an estimated overall prevalence of 1 in 7,700 live births [1]. They are mainly caused by monogenic defects in genes encoding lysosomal enzymes that degrade macromolecules such as glycolipids, glycoproteins and mucopolysaccharides. These defects produce an abnormal and progressive lysosomal accumulation of specific substrates, leading to structural changes and deterioration of the cellular function. LSDs are clinically heterogeneous, being usually undetectable at birth, and characterized by progressive manifestations that may include different organs and systems in the body [2]. Treatment for LSDs, already available for several of them, consists of enzyme replacement, transplantation of hematopoietic stem cells, substrate synthesis inhibition, pharmacological chaperones and some other strategies [2,3]. The specific treatment, when introduced early, may prevent irreversible pathological changes or significantly minimize disease manifestations [4,5].

These facts have motivated the development of screening methods to be used in large scale, enabling strategies such as newborn screening (NBS). Once NBS programs for LSDs are established, additional procedures for confirmatory diagnosis should be available as a mandatory part of these programs, to rule out false positives and to enable the prompt start of therapy whenever indicated in true positive cases.

Recently, NBS for LSDs was introduced by a newborn screening laboratory, the CTN (*Centro de Triagem Neonatal*), based in Porto Alegre, Brazil. The program was a pilot project to evaluate the use of a digital microfluidic (DMF) platform to measure simultaneously the activities of

* Corresponding author at: Medical Genetics Service, HCPA, Rua Ramiro Barcelos 2350, Porto Alegre, RS, Brazil.
E-mail address: rgiugliani@hcpa.edu.br (R. Giugliani).

http://dx.doi.org/10.1016/j.ymgmr.2017.06.006
Received 4 May 2017; Received in revised form 15 June 2017; Accepted 16 June 2017
Available online 04 July 2017

α-L-iduronidase (IDUA), acid α-glucosidase (GAA), acid β-glucosidase (GBA) and α-galactosidase (GLA) to screen for MPS I, Pompe disease, Gaucher disease and Fabry disease, respectively [Neto EC, personal communication]. The procedures for the first-tier screening were performed as described previously by Sista et al. [6,7], and are already being used in newborn screening programs for LSDs [8]. Cut off values were estimated as the activity 30% below the mean enzyme activity obtained with the analysis of DBS samples from 1,000 unaffected babies samples. These cutoffs were validated with the blind analysis of samples obtained from previously confirmed cases of MPS I, Gaucher, Fabry and Pompe diseases [Neto EC, personal communication].

Here, we present the results of the additional investigation performed in the cases that presented initial abnormal results in the above screening program. This investigation was based on biochemical and molecular genetics approaches. We also discuss the challenges encountered in the interpretation of these results.

2. Materials and methods

2.1. Samples

The cases with initial abnormal results in the program of NBS for LSDs were referred from the NBS laboratory (CTN) to the Medical Genetics Service of Hospital de Clínicas de Porto Alegre (HCPA). Both institutions are located in Porto Alegre, Rio Grande do Sul State, Brazil.

Dried blood spots (DBS), whole blood and urine samples were collected from the cases that had abnormal results in the initial screening for one of the four LSDs tested, for further investigation at the reference center. Blood samples were also collected from the parents in three of the cases for related analyses.

The biochemical and genetic investigations were performed at the Laboratory of Inborn Errors of Metabolism and at the Laboratory of Molecular Genetics, respectively, of the Medical Genetics Service (SGM) of HCPA. SGM/HCPA is a reference center for rare diseases in Brazil, and a WHO Collaborating Center for the Development of Medical Genetic Services in Latin America since 2004 [9].

2.2. Enzyme activity analyses

Enzyme activities of α-L-iduronidase (IDUA; EC 3.2.1.76), acid α-glucosidase (GAA; EC 3.2.1.20) and acid β-glucosidase (GBA; EC 3.2.1.45) were measured in leukocytes by fluorometric assays following procedures previously described [10–12]. Likewise, enzyme activities in DBS and plasma were measured by fluorometric assays in accordance with previous reports [10,13].

Chitotriosidase was measured in plasma by a fluorometric assay as reported previously [14].

2.3. Urinary glycosaminoglycans (GAGs) analysis

Urinary GAGs were analyzed by standard quantitative and qualitative methods, the dimethylmethylene blue (DBM) colorimetric assay and the monodimensional electrophoresis, respectively [15–17].

2.4. Gene analysis

2.4.1. Analysis of IDUA gene (OMIM *252800) for MPS I

Genomic DNA was isolated from peripheral blood sample in EDTA for case 1 and from blood impregnated in filter paper for case 4. The 14 exons and flanking regions of the *IDUA* gene were amplified by PCR and subsequently sequenced [18]. Identified variants were interpreted based on information found in the Human Gene Mutation Database (HGMD), dbSNP, ClinVar, ExAc and literature review [19–22]. New variants were analyzed *in silico* to predict pathogenicity using softwares such as Poly-Phen2 and SIFT [23,24].

2.4.2. Analysis of GAA gene (OMIM *606800) for Pompe disease

Genomic DNA was isolated from peripheral blood cells samples and used for sequencing in the Ion Torrent Personal Genome Machine (Thermo Scientific™), using a customized panel (Ion AmpliSeq™ Thermo Scientific™) that included the *GAA* gene. Analysis of data used the platform of bioinformatics from Ion Torrent Suite and Ion Reporter (Thermo Scientific™) version 5.0. All procedures were performed in accordance of the manufacturer's recommendations.

Sanger sequencing using ABI 3500 Genetic Analyzer (Applied Biosystems) was also used for the analysis of intron 1, exon 12 and 15 of *GAA* gene of proband's parents, as previously described [25]. Identified variants were interpreted using different databases such as Human Gene Mutation Database (HGMD), dbSNP, ClinVar, ExAc, Pompe Disease Mutation Database (Erasmus MC: Pompe Center), and literature review [19–22,26].

2.4.3. Analysis of GBA gene (OMIM *606463) for Gaucher disease

Genomic DNA was isolated from peripheral blood samples and then sequenced in the Ion Torrent Personal Genome Machine (Thermo Scientific™), using a customized panel (Ion AmpliSeq™ Thermo Scientific™) that included the *GBA* gene. Then, data were analyzed at the platform of bioinformatics from Ion Torrent Suite and Ion Reporter (Thermo Scientific™) version 5.0. All the above procedures followed the manufacturer's recommendations. Analysis was complemented by Sanger sequencing of exon 10 of the *GBA* gene to evaluate the presence of a pseudodeficiency allele [27]. Identified variants were interpreted using different databases such as Human Gene Mutation Database (HGMD), dbSNP, ClinVar, ExAc and literature review [19–22].

3. Results

Four cases, that screened positive among the first 10,567 babies tested in the program of NBS for LSDs, were further investigated. Data of the analyses performed for diagnostic confirmation and the results observed for each case are shown in Table 1. Description of each case is presented below.

3.1. Case 1: suspicion of MPS I

A female baby was referred for further investigation, after resulting positive for a NBS for MPS I, which revealed a low IDUA activity (0.8 μmol/L/h; cut off: > 5.0) measured on DBS.

Urinary GAGs were analyzed and showed a normal GAGs quantitation for the age and a normal GAGs pattern at the qualitative analysis. IDUA activity was measured in DBS, plasma and leukocytes samples. IDUA activity was reported as undetectable in DBS. Measurement in plasma showed a normal enzyme activity and the analysis in leukocytes revealed an IDUA activity below the normal range (11 nmol/h/mg protein, with normal reference range from 27 to 171).

After considering all the biochemical results, it was not possible to reach a conclusion about the MPS I diagnosis. Therefore, molecular analysis of the *IDUA* gene was performed, with the identification of the variant c.251G > C [p.(Gly84Ala)] and the variant NM_000203.4(IDUA):c.246C > G (p.His82Gln). The variant p.(Gly84Ala) was a recently reported variant, predicted as possibly pathogenic by *in silico* analysis and located at the same codon where two pathogenic variants were already described [18]. The variant p.His82Gln was previously described as benign and possibly leading to pseudodeficiency, resulting to low *in vitro* enzyme activity in normal subjects [28–30].

Thus, putting together the results of normal urinary GAGs, low IDUA activity in leukocytes (but higher than that usually observed in affected cases for MPS I) and a genotype with a possibly pathogenic variant and a variant associated with pseudodeficiency, the conclusion was that the baby presented pseudodeficiency for MPS I.

Table 1
Confirmatory investigation of cases screened positive in a program of NBS for LSDs in Brazil.

	Case 1	Case 2	Case 3	Case 4
	MPS I?	Pompe?	Gaucher?	MPS I?
Enzyme analysis	IDUA	GAA	GBA	IDUA
DBS-fluorometry	Undetectable	NP	2.8 nmol/h/mL (2.2–17)	NP
Plasma-fluorometry	11 nmol/h/mL (6.6–34)	NP	NP	NP
Leukocytes-fluorometry	11 nmol/h/mg protein (27–171)	1.00 nmol/h/mg protein (1.00–7.60) Father: 1.9 Mother: 2.70	5.6 nmol/h/mg protein (10–45) Father: 8.1 Mother: 22.0	27 nmol/h/mg protein (27–171)
Urinary GAGs				
Quantitation (DMB - colorimetry)	197 μg/mg creatinine (133–460)	NP	NP	272 μg/mg creatinine (133–460)
Electrophoresis (qualitative)	Normal GAG pattern	NP	NP	Normal GAG pattern
Gene analysis	*IDUA*	*GAA*	*GBA*	*IDUA*
Mutation 1	c.251G > C	c.-32-13T > G	c.1226A > G	c.1205G > A
Effect	p.(Gly84Ala)	Splice site variant	p.Asn409Ser (N370S)	p.Trp402Ter
Significance	Predicted pathogenic	Pathogenic variant	Pathogenic variant	Pathogenic variant
Mutation 2	c.246C > G	c.[1726G > A; 2065G > A]	No pathogenic variant identified	No pathogenic variant identified
Effect	p.His82Gln	p.[Gly576Ser; Glu689Lys]		
Significance	Pseudodeficiency allele	Pseudodeficiency allele Father: c.-32-13T > G Mother: p.[Gly576Ser; Glu689Lys]		Father: c.1205G > A Mother: No pathogenic variant

Numbers in parenthesis, in enzyme analysis and urinary GAGs, are reference values. IDUA: α-L iduronidase; GAA: acid α-glucosidase; GBA: acid β-glucosidase; MPS I: mucopolysaccharidosis type 1. DBS: dried blood spot; GAGs: glycosaminoglycans. NP: not performed.

3.2. Case 2: suspicion of Pompe disease

A male baby, clinically normal, was referred for further investigation after presenting a low GAA activity (4.3 μmol/L/h; cut off: > 10) in a NBS for Pompe disease.

For confirmatory diagnosis, GAA activity was measured in leukocytes and resulted in slightly low (0.94 nmol/h/mg protein, with normal reference range from 1.00 to 7.60) in an initial measurement and at the lower limit of the reference range (1.0 nmol/h/mg protein) when the analysis was repeated.

Given the slightly low enzyme activity (although higher than that usually observed in patients with Pompe disease), a conclusion about the tentative Pompe diagnosis was not possible. Then, *GAA* gene sequencing was performed to elucidate the case. It was detected a known pathogenic variant in heterozygosis, the NM_000152.4(GAA):c.-32-13T > G in one chromosome, and in the other chromosome a previously reported pseudodeficiency allele [31,32] that consists of two variants, the NM_000152.4(GAA):c.1726G > A (p.Gly576Ser) and the NM_000152.3(GAA):c.2065G > A (p.Glu689Lys). Variants found by NGS were confirmed using Sanger sequencing.

Additionally, the parents of the infant were also evaluated by enzymatic and molecular analyses. The enzyme assays revealed a normal GAA activity in leukocytes for both parents. The molecular analysis showed that the father was carrier of the variant c.-32-13T > G and the mother was carrier for the two variants, c.1726G > A (p.Gly576Ser) and c.2065G > A (p.Glu689Lys).

Hence, based on all the above results in the infant and the information provided for the analysis in the parents, the case was defined as pseudodeficiency for Pompe disease.

3.3. Case 3: suspicion of Gaucher disease

A male newborn, referred for further investigation after a result in the NBS for Gaucher disease that showed a low GBA activity (6.1 μmol/L/h; cut off: > 7) in a DBS sample.

In the additional investigation, GBA activity in DBS exhibited a normal activity. The enzyme assay performed in leukocytes resulted in a low GBA activity (5.6 nmol/h/mg protein, with normal reference range from 10 to 45). Chitotriosidase was not helpful, as it was evaluated in DBS (activity undetectable, with reference range from 0 to 44 nmol/h/mL) and in plasma (activity 0.1 nmol/h/ml, with normal reference values ranging from 8.8 to 132). As biochemical results were not conclusive, *GBA* gene sequencing was performed, and the variant NM_001005741.2(GBA):c.1226A > G (p.Asn409Ser) was identified in heterozygosis. This is a well-known pathogenic variant also described as p.N370S. Additionally, it was discarded the possibility of pseudodeficiency after identifying a normal sequence for exon 10 of *GBA* gene that is the usual location of complex recombination between the *GBA* gene and the pseudogene.

The parents were also evaluated. Analysis of GBA activity in leukocytes resulted in a low activity for the father only, being normal for the mother. This sample was unsuitable for molecular analysis, which was not performed in the parents as they did not return for blood collection.

Then, gathering all the above information, the conclusion was that this baby was as a carrier for Gaucher disease.

3.4. Case 4: suspicion of MPS I

A female newborn was referred for further investigation after being screened positive for a NBS for MPS I. The screening resulted in a low IDUA activity (2.4 μmol/L/h; cut off: > 5.0).

Evaluation of this case started with the urinary GAGs analysis that resulted normal in the quantitative and qualitative analyses. Then, enzyme activity was measured in leukocytes and revealed an IDUA activity at the lower limit of the reference range (27 nmol/h/mg protein, with reference range from 27 to 171). Given this borderline result of the enzyme activity and the normal urinary excretion of GAGs,

biochemical results were considered inconclusive.

Molecular analysis with sequencing of the *IDUA* gene was then performed in the baby, with the identification of a known pathogenic variant in heterozygosis, the NM_000203.4(IDUA):c.1205G > A (p.Trp402Ter). Targeted gene analysis was also performed in both parents, by sequencing of the affected exon. It demonstrated the presence of this variant in heterozygosis at the father's DNA and absent in the mother's sample.

Based on the enzymatic assay and the gene analysis results, together to normal excretion of GAGs in urine, the conclusion was that the baby is a carrier for MPS I.

4. Discussion

We report the investigation performed in the four presumptive cases for LSDs identified in a pilot study of NBS for 4 LSDs (MPS I, Fabry, Gaucher, and Pompe diseases) carried out in a NBS laboratory in Brazil. Two of the cases had suspicion of MPS I, one had suspicion of Gaucher disease and one had suspicion of Pompe disease. The investigation included biochemical and molecular analyses performed in the babies and in their parents. No affected subject for any of the diseases was diagnosed. However, we did not classify these cases as false positives, as they were identified as having pseudodeficiency (one case of suspected MPS I and one case of suspected Pompe disease) or as carriers (one case of suspected MPS I and one case of suspected Gaucher disease).

The first baby had a suspicion of MPS I. MPS I, caused by IDUA deficiency that fail to degrade the glycosaminoglycans heparan and dermatan sulfate, is diagnosed by measuring mainly a reduced IDUA activity in leukocytes or in other nucleated cell and by either one or both increased excretion of GAGs in urine and a pattern of heparan and dermatan sulfate excretion at the electrophoresis [33]. Biochemical investigation showed normal GAG excretion, suggesting an absence of functional impact of an apparent IDUA deficiency on GAGs degradation. Normal GAG excretion with low IDUA activity suggests the possibility of pseudodeficiency, and molecular analysis is recommended to elucidate the diagnosis. Despite the presence of a possibly pathogenic variant p.(Gly84Ala), the presence of a pseudodeficiency allele p.His82Gln allowed normal degradation of GAGs. Pseudodeficiency condition was found in other NBS programs for MPS I, with an estimated frequency of 0.01% to 0.02% of the total screened samples in each study [8,34]. These NBS programs, carried out mainly in U.S.A. (Missouri, Illinois and New York), reported pseudodeficiency cases among the screened positive samples for MPS I and the number of confirmed pseudodeficiency cases was higher than the true affected cases. Although NBS programs of other countries such as Taiwan and Italy did not report pseudodeficiency cases for MPS I [35,36], the possibility to find this condition in the evaluation of suspected MPS I should be clearly taken in consideration. Therefore, this case was identified as pseudodeficiency for MPS I, without pathogenic consequences, allowing the prediction of a normal child.

Pseudodeficiency has been already described as a possible confounder in the interpretation of enzymatic assay results for some LSDs [37], including Pompe disease. Diagnosis of Pompe disease is established by a decreased GAA activity in leukocytes or fibroblast and a genotype demonstrating pathogenic variants of the *GAA* gene in homozygosis or in compound heterozygosis [38]. Because enzyme assay has limitations to discriminate pseudodeficiency and carrier status of affected or normal cases, gene analysis is required to establish the diagnosis. The genotyping of the baby with suspected Pompe disease allowed the identification of a combination of a previously reported pseudodeficiency allele with a known pathogenic mutation, both in heterozygosis, which explain the slight reduction of the GAA activity. Previous *in vitro* studies have shown that the two variants of the pseudodeficiency allele, when combined, reduce the GAA activity by approximately 80% in comparison to the expression of wild-type cDNA [31] and are highly frequent in Asian populations [32]. Likewise, the c.-32-13T > G, a splice site variant of intron 1, has been reported as the most frequent pathogenic variant in adult onset Caucasian patients [39] and may reduce the GAA activity to a range of 3% to 20% of the normal when presented in compound heterozygous state, combined with other deleterious *GAA* gene variants [40,41]. Since this variant was observed mostly in juvenile and adult form of Pompe disease, it is considered of mild effect. Combination of a pseudodeficiency allele and a pathogenic variant may exhibit different levels of reduction of the GAA activity as observed in the case investigated in this study and contrasted by other study where the described case showed an important decrease of GAA activity, which may be accounted for the effect of a nonsense mutation considered more deleterious p.[Gly576Ser; Glu689Lys]/p.Trp746Ter [31]. Other newborn screening studies for Pompe disease have also reported similar cases of carriers with an additional pseudodeficiency allele that were part of the false-positive cases found in that screening program [32,42,43]. Thus, caution has been already recommended in the interpretation of enzyme activity results in cases when pseudodeficiency alleles are present. The diagnosis of this case was established as pseudodeficiency for Pompe disease, allowing the prediction of a normal clinical course for the proband.

One baby had a suspicion of Gaucher disease, which is caused by a deficient GBA activity, leading to glucocerebroside accumulation in cells of monocyte or macrophage lineage. Its diagnosis is usually established after demonstrating enzyme deficiency in leukocytes or fibroblasts [44]. The case showed a low enzyme activity in leukocytes but not so reduced as observed in affected cases [45]. When enzyme activity results show an overlap of the values found in carriers and in non-carriers, *GBA* gene analysis should be performed [44]. Chitotriosidase activity could provide important information if elevated, which would suggest Gaucher disease. When it is very low, as in the present case, results are not as informative as it could be caused by a common mutation that affects its activity [46,47]. To elucidate the case, molecular analysis of the *GBA* gene was performed, being identified the most common disease-causing variant (N370S), that has been associated to Gaucher disease type 1 [48]. Carriers for Gaucher disease were identified in other NBS programs, such as those performed in Washington, Illinois and New York in the U.S.A., Hungary and Taiwan, with a frequency estimated in the range of 0.002% to 0.02% of total screened samples [49–51]. Genotypes included different variants, but the p.Asn409Ser (p.N370S) was observed in all these NBS studies and reported as the most common allele among the identified alleles [34]. Therefore, in our study, as the pathogenic variant was found in a heterozygous state, the baby was only a carrier and consequently there should be no risk to developing clinical disease.

Our last case was, again, one with a suspicion of MPS I. The measurement of IDUA activity in leukocytes was inconclusive, with an enzyme activity in the lower limit of the reference range. The molecular analysis of the *IDUA* gene elucidated the diagnosis demonstrating a common pathogenic variant (p.Trp402Ter) in heterozygous state. This variant in homozygous state has been associated with the severe phenotype of MPS I [52]. A Brazilian study showed that this variant accounted for 38% of the alleles in patients with MPS I [53]. Other NBS programs also found carriers for MPS I with an estimated frequency of 0.001% to 0.005% of the total screened samples, including all cases reported as confirmed carriers [8,34,35,43,50]. Although, not all these studies reported the genotype identified, the reported variants were different to the one found in our study. Being a carrier for MPS I, this baby is not at risk of developing clinical disease.

The investigation performed in these cases illustrates the possible strategies for confirmatory diagnosis in asymptomatic subjects from NBS programs for LSDs and the challenges that may be faced during its interpretation. Previous studies on NBS for LSDs discuss briefly on the additional procedures used for the investigation of suspected cases, with variable strategies according to the laboratory. Some perform enzymatic and molecular analyses simultaneously, while others use only the molecular analysis. Among the challenges during

interpretation, the presence of pseudodeficiencies or carrier status represents situations difficult to diagnose by biochemical methods, which, however, are important to identify the functional status of the patient.

Molecular analysis seems to be critical for the understanding of each case, but may also show some difficulties in the interpretation when new gene variants of unknown significance are identified, that will require further prediction exercises and functional studies to elucidate its effect and validate its significance.

Therefore, all these aspects should be considered in the process of diagnostic confirmation, especially when the cases are identified in mass screening programs of clinically normal subjects, as it is the case of NBS.

Finally, it is worthy to mention the absolute need of having comprehensive diagnostic protocols in place when a NBS for LSDs is performed. In the investigation of babies screened positive, the integration of the different pieces of the screening team, (screening lab, biochemical diagnosis lab, molecular genetics lab and clinical group) is very important to establish the correct diagnosis of each case.

5. Conclusions

Biochemical and molecular procedures for confirmatory investigation of newborns who had abnormal results in the initial test in NBS programs for LSDs should be an essential part of the program, and should be performed, whenever possible, in reference centers with high expertise in the diagnosis of these diseases. This allows a rapid and precise investigation of the babies who have an abnormal screening result, reducing parental anxiety in false-positives and allowing prompt initiation of therapy in the cases with confirmed disease.

Author contributions

HB and RG conceived the investigation for confirmatory diagnosis, wrote the first draft and analyzed the data; ECN supervised the NBS for LSDs program; JS and JP performed the NBS analyses; CSF provided expert advice on NBS; FB and FS performed the enzyme analysis for confirmatory diagnosis; RRG performed the urine GAGs analysis; KM-T supervised the enzyme and GAGs analyses; ACB-F, GP, DRM and FBT performed the molecular analyses for confirmatory diagnosis; RG supervised the whole procedures of the investigation for confirmatory diagnosis; all authors revised and approved the final version of this manuscript.

Conflicts of interest

The authors declare no conflict of interest to report in relation to this manuscript.

Acknowledgments

The work described in this paper was supported by the MPS Brazil Network (FIPE/HCPA 03-066), the LSD Brazil Network (FIPE/HCPA 05-215) and INAGEMP (CNPq grant 573993/2008–04). HB received a PhD scholarship from CNPq, Brazil (grant 140125/2014-0). The authors are grateful to Dr. Marina Siebert and Dr. Ida Schwartz, who kindly performed Sanger sequencing of selected regions of the *GBA* gene in the case 3 DNA sample.

References

[1] P.J. Meikle, J.J. Hopwood, A.E. Clague, W.F. Carey, Prevalence of lysosomal storage disorders, JAMA 281 (3) (1999 Jan 20) 249–254.
[2] G.H. Hoffman, J. Zschocke, W.L. Nyhan, Inherited Metabolic Diseases: A Clinical Approach, Springer-Verlag Berlin-Heidelberg, 2010 (401 pp).
[3] F.M. Platt, R.H. Lachmann, Treating lysosomal storage disorders: current practice and future prospects, Biochim. Biophys. Acta 1793 (4) (2009) 737–745.
[4] J.J. McGill, A.C. Inwood, D.J. Coman, M.L. Lipke, D. de Lore, S.J. Swiedler, J.J. Hopwood, Enzyme replacement therapy for mucopolysaccharidosis VI from 8 weeks of age–a sibling control study, Clin. Genet. 77 (5) (2010 May) 492–498, http://dx.doi.org/10.1111/j.1399-0004.2009.01324.x (Epub 2009 Nov 23).
[5] O. Gabrielli, L.A. Clarke, A. Ficcadenti, L. Santoro, L. Zampini, N. Volpi, G.V. Coppa, 12 year follow up of enzyme-replacement therapy in two siblings with attenuated mucopolysaccharidosis I: the important role of early treatment, BMC Med. Genet. 17 (2016 Mar 10) 19, http://dx.doi.org/10.1186/s12881-016-0284-4.
[6] R.S. Sista, A.E. Eckhardt, T. Wang, C. Graham, J.L. Rouse, S.M. Norton, V. Srinivasan, M.G. Pollack, A.A. Tolun, D. Bali, D.S. Millington, V.K. Pamula, Digital microfluidic platform for multiplexing enzyme assays: implications for lysosomal storage disease screening in newborns, Clin. Chem. 57 (10) (2011 Oct) 1444–1451, http://dx.doi.org/10.1373/clinchem.2011.163139 (Epub 2011 Aug 22).
[7] R.S. Sista, T. Wang, N. Wu, C. Graham, A. Eckhardt, T. Winger, V. Srinivasan, D. Bali, D.S. Millington, V.K. Pamula, Multiplex newborn screening for Pompe, Fabry, Hunter, Gaucher, and Hurler diseases using a digital microfluidic platform, Clin. Chim. Acta 424 (2013 Sep 23) 12–18, http://dx.doi.org/10.1016/j.cca.2013.05.001 (Epub 2013 May 7).
[8] P.V. Hopkins, C. Campbell, T. Klug, S. Rogers, J. Raburn-Miller, J. Kiesling, Lysosomal storage disorder screening implementation: findings from the first six months of full population pilot testing in Missouri, J. Pediatr. 166 (1) (2015 Jan) 172–177, http://dx.doi.org/10.1016/j.jpeds.2014.09.023 (Epub 2014 Oct 18).
[9] R. Giugliani, F.P. Vairo, M. Riegel, C.F. de Souza, I.V. Schwartz, S.D. Pena, Rare disease landscape in Brazil: report of a successful experience in inborn errors of metabolism, Orphanet J. Rare Dis. 11 (1) (2016 Jun 10) 76, http://dx.doi.org/10.1186/s13023-016-0458-3.
[10] J.J. Hopwood, V. Muller, A. Smithson, N. Baggett, A fluorometric assay using 4-methylumbelliferyl alpha-L-iduronide for the estimation of alpha-L-iduronidase activity and the detection of Hurler and Scheie syndromes, Clin. Chim. Acta 92 (2) (1979 Mar 1) 257–265.
[11] Y. Li, C.R. Scott, N.A. Chamoles, A. Ghavami, B.M. Pinto, F. Turecek, M.H. Gelb, Direct multiplex assay of lysosomal enzymes in dried blood spots for newborn screening, Clin. Chem. 50 (10) (2004 Oct) 1785–1796 (Epub 2004 Aug 3).
[12] S.P. Peters, P. Coyle, R.H. Glew, Differentiation of beta-glucocerebrosidase from beta-glucosidase in human tissues using sodium taurocholate, Arch. Biochem. Biophys. 175 (2) (1976 Aug) 569–582.
[13] G. Civallero, K. Michelin, J. de Mari, M. Viapiana, M. Burin, J.C. Coelho, R. Giugliani, Twelve different enzyme assays on dried-blood filter paper samples for detection of patients with selected inherited lysosomal storage diseases, Clin. Chim. Acta 372 (1–2) (2006 Oct) 98–102 (Epub 2006 May 18).
[14] C.E. Hollak, S. van Weely, M.H. van Oers, J.M. Aerts, Marked elevation of plasma chitotriosidase activity. A novel hallmark of Gaucher disease, J. Clin. Invest. 93 (3) (1994 Mar) 1288–1292.
[15] J.G. de Jong, R.A. Wevers, Sambeek R. Liebrand-van, Measuring urinary glycosaminoglycans in the presence of protein: an improved screening procedure for mucopolysaccharidoses based on dimethylmethylene blue, Clin. Chem. 38 (6) (1992 Jun) 803–807.
[16] C.A. Pennock, A review and selection of simple laboratory methods used for the study of glycosaminoglycan and the diagnosis of the mucopolysaccharidoses, J. Clin. Pathol. 29 (1976) 111–123.
[17] R. Cappelletti, M. Del Rosso, V.P. Chiarrugi, A new electrophoretic method for complete separation of all known animal glycosaminoglycans in a monodimensional run, Anal. Biochem. 99 (1979) 311–315.
[18] G. Pasqualim, E. Poletto, U. Matte, R. Giugliani, Mutational profile of Brazilian MPS I patients, J. Inborn Errors Metab. Screen. 4 (2016) 1–82, http://dx.doi.org/10.1177/2326409816653634 (Suppl.)p.16..
[19] The Human Gene Mutation Database (HGMD), Available online: http://www.hgmd.cf.ac.uk/ac/index.php (accessed on Sept–Nov 2015; May–Jul, 2016).
[20] dbSNP, Available online: https://www.ncbi.nlm.nih.gov/SNP/ (accessed on Sept–Nov 2015; May–Jul, 2016).
[21] ClinVar, Available online: https://www.ncbi.nlm.nih.gov/clinvar/ (accessed on Sept–Nov 2015; May–July, 2016).
[22] Exome Aggregation Consortium (ExAC). http://exac.broadinstitute.org/ (Accessed: Sept–Nov 2015; May–Jul, 2016).
[23] Poly-Phen2 Tool, Available online: http://genetics.bwh.harvard.edu/pph2/ (accessed on Sept–Nov 2015).
[24] SIFT, Available online: http://sift.bii.a-star.edu.sg/ (accessed on Sept–Nov 2015).
[25] S.M. Oba-Shinjo, R. da Silva, F.G. Andrade, R.E. Palmer, R.J. Pomponio, K.M. Ciociola, M. S. Carvalho, P.S. Gutierrez, G. Porta, C.D. Marrone, V. Munoz, A.K. Grzesiuk, J.C. Llerena Jr., C.R. Berditchevsky, C. Sobreira, D. Horovitz, T.P. Hatem, E.R. Frota, R. Pecchini, J.A. Kouyoumdjian, L. Werneck, V.M. Amado, J.S. Camelo Jr., R.J. Mattaliano, S.K. Marie, Pompe disease in a Brazilian series: clinical and molecular analyses with identification of nine new mutations, J. Neurol. 256 (11) (2009 Nov) 1881–1890, http://dx.doi.org/10.1007/s00415-009-5219-y (Epub 2009 Jul 9).
[26] Pompe Disease Mutation Database (Erasmus MC: Pompe Center), Available online: https://www.erasmusmc.nl/klinische_genetica/research/lijnen/pompe_center/?lang=en (accessed on May, 2016).
[27] M. Siebert, H. Bock, K. Michelin-Tirelli, J.C. Coelho, R. Giugliani, M.L. Saraiva-Pereira, Novel mutations in the glucocerebrosidase gene of brazilian patients with Gaucher disease, JIMD Rep. 9 (2013) 7–16, http://dx.doi.org/10.1007/8904_2012_174 (Epub 2012 Oct 9).
[28] G. Yogalingam, X.H. Guo, V.J. Muller, D.A. Brooks, P.R. Clements, E.D. Kakkis, J.J. Hopwood, Identification and molecular characterization of alpha-L-iduronidase mutations present in mucopolysaccharidosis type I patients undergoing enzyme

replacement therapy, Hum. Mutat. 24 (3) (2004 Sep) 199–207.
[29] ClinVar. NM_000203.4, Available at: https://www.ncbi.nlm.nih.gov/clinvar/variation/92637/ (Accessed on Sept–Nov, 2015, Dec 2016, April 2017).
[30] L.M. Pollard, S.R. Braddock, K.M. Christensen, D.J. Boylan, L.D. Smith, B.A. Heese, A.M. Atherton, C.E. Lawson, M.E. Strenk, M. Willing, L. Manwaring, T.C. Wood, Three apparent pseudo-deficiency alleles in the *IDUA* gene identified by newborn screening, 63rd Annual Meeting of the American Society of Human Genetics, Boston, MA, October 22–26, 2013 Available at: http://www.ashg.org/2013meeting/pdf/46025_Platform_bookmark%20for%20Web%20Final%20from%20AGS.pdf (Accessed April 2017).
[31] M.A. Kroos, R.A. Mullaart, L. Van Vliet, R.J. Pomponio, H. Amartino, E.H. Kolodny, G.M. Pastores, R.A. Wevers, A.T. Van der Ploeg, D.J. Halley, A.J. Reuser, p.[G576S; E689K]: pathogenic combination or polymorphism in Pompe disease? Eur. J. Hum. Genet. 16 (8) (2008 Aug) 875–879, http://dx.doi.org/10.1038/ejhg.2008.34 (Epub 2008 Feb 27).
[32] S. Kumamoto, T. Katafuchi, K. Nakamura, F. Endo, E. Oda, T. Okuyama, M.A. Kroos, A.J. Reuser, T. Okumiya, High frequency of acid alpha-glucosidase pseudodeficiency complicates newborn screening for glycogen storage disease type II in the Japanese population, Mol. Genet. Metab. 97 (3) (2009 Jul) 190–195, http://dx.doi.org/10.1016/j.ymgme.2009.03.004 (Epub 2009 Mar 18).
[33] L.A. Clarke, Mucopolysaccharidosis Type I. 2002 Oct 31 [updated 2016 Feb 11], in: R.A. Pagon, M.P. Adam, H.H. Ardinger, et al. (Eds.), GeneReviews®, University of Washington, Seattle, Seattle (WA), 1993–2016(Internet). (Cited 2016 Dec). Available from: https://www.ncbi.nlm.nih.gov/books/NBK1162/.
[34] Burton BK, Kishnani PS, Wasserstein M. Newborn screening for lysosomal diseases: recent progress and unanswered questions. Satellite Symposium Presented at: 13th Annual WORLDSymposium, Feb 13, 2017, San Diego, CA.
[35] S.P. Lin, H.Y. Lin, T.J. Wang, C.Y. Chang, C.H. Lin, S.F. Huang, C.C. Tsai, H.L. Liu, J. Keutzer, C.K. Chuang, A pilot newborn screening program for Mucopolysaccharidosis type I in Taiwan, Orphanet J. Rare Dis. 8 (2013 Sep 22) 147, http://dx.doi.org/10.1186/1750-1172-8-147.
[36] S. Paciotti, E. Persichetti, S. Pagliardini, M. Deganuto, C. Rosano, C. Balducci, M. Codini, M. Filocamo, A.R. Menghini, V. Pagliardini, S. Pasqui, B. Bembi, A. Dardis, T. Beccari, First pilot newborn screening for four lysosomal storage diseases in an Italian region: identification and analysis of a putative causative mutation in the *GBA* gene, Clin. Chim. Acta 413 (23–24) (2012 Nov 20) 1827–1831, http://dx.doi.org/10.1016/j.cca.2012.07.011 (Epub 2012 Jul 20).
[37] A. Mehta, B. Winchester, Chapter 3, laboratory diagnosis, Lysosomal Storage Disorders: A Practical Guide, Wiley-Blackwell, UK, 2012, p. 22.
[38] N. Leslie, B.T. Tinkle, Glycogen Storage Disease Type II (Pompe Disease) 2007 Aug 31 [updated 2013 May 9], in: R.A. Pagon, M.P. Adam, H.H. Ardinger, et al. (Eds.), GeneReviews®, University of Washington, Seattle, Seattle (WA), 1993–2016(Internet). (cited 2016 Dec). Available from: https://www.ncbi.nlm.nih.gov/books/NBK1261/.
[39] Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD, 1966-2017. MIM Number: #232300: created 1986 (Updated 2012), (Cited 2016 Dec 26). World Wide Web URL: https://omim.org/entry/232300.
[40] Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD, 1966-2017. MIM Number: * 606800: created 2002 (Updated 2011), (Cited 2016 Dec 26). World Wide Web URL: https://omim.org/entry/606800#0006.
[41] M.A. Kroos, R.J. Pomponio, M.L. Hagemans, J.L. Keulemans, M. Phipps, M. DeRiso, R.E. Palmer, M.G. Ausems, N.A. Van der Beek, O.P. Van Diggelen, D.J. Halley, A.T. Van der Ploeg, A.J. Reuser, Broad spectrum of Pompe disease in patients with the same c.-32-13T- > G haplotype, Neurology 68 (2) (2007 Jan 9) 110–115.
[42] P. Labrousse, Y.H. Chien, R.J. Pomponio, J. Keutzer, N.C. Lee, V.R. Akmaev, T. Scholl, W.L. Hwu, Genetic heterozygosity and pseudodeficiency in the Pompe disease newborn screening pilot program, Mol. Genet. Metab. 99 (4) (2010 Apr) 379–383, http://dx.doi.org/10.1016/j.ymgme.2009.12.014 (Epub 2009 Dec 28).
[43] C.R. Scott, S. Elliott, N. Buroker, L.I. Thomas, J. Keutzer, M. Glass, M.H. Gelb, F. Turecek, Identification of infants at risk for developing Fabry, Pompe, or mucopolysaccharidosis-I from newborn blood spots by tandem mass spectrometry, J. Pediatr. 163 (2) (2013 Aug) 498–503, http://dx.doi.org/10.1016/j.jpeds.2013.01.031 (Epub 2013 Mar 1).
[44] G.M. Pastores, D.A. Hughes, Gaucher Disease. 2000 Jul 27 [updated 2015 Feb 26], in: R.A. Pagon, M.P. Adam, H.H. Ardinger, et al. (Eds.), GeneReviews®, University of Washington, Seattle, Seattle (WA), 1993–2016(Internet). (Cited 2016 Dec) Available from: https://www.ncbi.nlm.nih.gov/books/NBK1269/.
[45] K. Michelin, A. Wajner, F.T. de Souza, A.S. de Mello, M.G. Burin, M.L. Pereira, R.F. Pires, R. Giugliani, J.C. Coelho, Application of a comprehensive protocol for the identification of Gaucher disease in Brazil, Am. J. Med. Genet. A 136 (1) (2005 Jul 1) 58–62.
[46] R.G. Boot, G.H. Renkema, M. Verhoek, A. Strijland, J. Bliek, T.M. de Meulemeester, M.M. Mannens, J.M. Aerts, The human chitotriosidase gene. Nature of inherited enzyme deficiency, J. Biol. Chem. 273 (40) (1998 Oct 2) 25680–25685.
[47] M.E. Grace, M. Balwani, I. Nazarenko, A. Prakash-Cheng, R.J. Desnick, Type 1 Gaucher disease: null and hypomorphic novel chitotriosidase mutations-implications for diagnosis and therapeutic monitoring, Hum. Mutat. 28 (9) (2007 Sep) 866–873.
[48] Online Mendelian Inheritance in Man, OMIM®. Johns Hopkins University, Baltimore, MD, 1966-2017. MIM Number: * 606463: created 2001 (Updated 2017). [Cited 2016 Dec]. World Wide Web URL: https://www.omim.org/entry/606463#0003.
[49] J. Wittmann, E. Karg, S. Turi, E. Legnini, G. Wittmann, A.K. Giese, J. Lukas, U. Gölnitz, M. Klingenhäger, O. Bodamer, A. Mühl, A. Rolfs, Newborn screening for lysosomal storage disorders in Hungary, JIMD Rep. 6 (2012) 117–125, http://dx.doi.org/10.1007/8904_2012_130 (Epub 2012 Mar 21).
[50] S. Elliott, N. Buroker, J.J. Cournoyer, A.M. Potier, J.D. Trometer, C. Elbin, M.J. Schermer, J. Kantola, A. Boyce, F. Turecek, M.H. Gelb, C.R. Scott, Pilot study of newborn screening for six lysosomal storage diseases using Tandem Mass Spectrometry, Mol. Genet. Metab. 118 (4) (2016 Aug) 304–309, http://dx.doi.org/10.1016/j.ymgme.2016.05.015 (Epub 2016 May 20).
[51] H.C. Liao, C.C. Chiang, D.M. Niu, C.H. Wang, S.M. Kao, F.J. Tsai, Y.H. Huang, H.C. Liu, C.K. Huang, H.J. Gao, C.F. Yang, M.J. Chan, W.D. Lin, Y.J. Chen, Detecting multiple lysosomal storage diseases by tandem mass spectrometry–a national newborn screening program in Taiwan, Clin. Chim. Acta 431 (2014 Apr 20) 80–86, http://dx.doi.org/10.1016/j.cca.2014.01.030 (Epub 2014 Feb 7).
[52] H.S. Scott, T. Litjens, J.J. Hopwood, C.P. Morris, A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype, Hum. Mutat. 1 (2) (1992) 103–108.
[53] U. Matte, G. Yogalingam, D. Brooks, S. Leistner, I. Schwartz, L. Lima, et al., Identification and characterization of 13 new mutations in mucopolysaccharidosis type I patients, Mol. Genet. Metab. 78 (1) (2003 Jan) 37–43.

ANEXO II – Artigo 6: Detección de la mutación E8SJM en el gen LIPA, por PCR en tiempo real, para la investigación de la enfermedad por almacenamiento de ésteres de colesterol

---

Publicado na revista *ANALES DE LA FACULTAD DE MEDICINA- REVISTA DE INVESTIGACIÓN UNMSM*: 2018; 79(1)

Detección de la mutación E8SJM en el gen *LIPA*, por PCR en tiempo real, para la investigación de la enfermedad por almacenamiento de ésteres de colesterol

Detection of the E8SJM mutation in the *LIPA* gene, by real-time PCR, for the investigation of cholesteryl ester storage disease

Diana Rojas Málaga[1], Ursula Matte[1,3], Carlos Thadeu Schmidt Cerski[2], Roberto Giugliani[3,4]

[1] Centro de Terapia Génica, Hospital de Clínicas de Porto Alegre, RS, Brasil.
[2] Servicio de Patología, Hospital de Clínicas de Porto Alegre, RS, Brasil.
[3] Departamento de Genética, UFRGS, Porto Alegre, RS, Brasil
[4] Servicio de Genética Médica, Hospital de Clínicas de Porto Alegre, RS, Brasil.

Correspondencia
Diana Rojas Málaga
dmalaga@hcpa.edu.br
Telef. +55 51 3359-7212
Direcc. Servicio de Genética Médica, Hospital de Clínicas de Porto Alegre.
Rua Ramiro Barcelos 2350, Porto Alegre RS 90035-903 Brasil.

Recibido: 25 de mayo 2017
Aceptado: 2 de septiembre 2017.

Conflicto de intereses: No existe conflicto de interés en el presente trabajo.

Financiamiento: Este trabajo fue financiado por FIPE/HCPA (proyecto 13-0189)

Consentimiento: Todos los experimentos presentados en esta publicación están incluidos en el proyecto de investigación aprobado por sus aspectos éticos y metodológicos por el Comité de Ètica en investigación del Hospital de Clínicas de Porto Alegre (proyecto 13-0189).

Citar como: Rojas Málaga D, Matte U, Cerski CTS, Giugliani R. Detección de la mutación E8SJM en el gen LIPA, por PCR en tiempo real, para la investigación de la enfermedad por almacenamiento de ésteres de colesterol. An Fac med. 2017;78(4):409-13
DOI: http://dx.doi.org/10.15381/anales.v78i4.14262

An Fac med. 2017;78(4):409-13 / http://dx.doi.org/10.15381/anales.v78i4.14262

Resumen

Introducción: La enfermedad por Almacenamiento de Ésteres de Colesterol (CESD; *Cholesteryl Ester Storage Disease*) es una enfermedad de depósito lisosomal, su presentación es bastante variable y su diagnóstico constituye un desafío. Además, existe un número de anomalías observadas en los pacientes con CESD que se sobreponen a diagnósticos más comunes, siendo probable que sea subdiagnosticada. La mayoría de pacientes relatados hasta el momento son portadores de la mutación E8SJM en el gen *LIPA*. En este sentido, el auxilio en el diagnóstico es fundamental pues existen opciones de terapia en desarrollo. **Diseño:** Estudio observacional. **Objetivo:** Estandarizar la técnica de PCR en tiempo real para la detección de la mutación más frecuente, E8SJM, en muestras de sangre periférica y de biopsia hepática para el auxilio diagnóstico de CESD y futuros estudios de prevalencia de la mutación. **Institución:** Centro de Terapia Génica del Hospital de Clínicas de Porto Alegre (HCPA), Brasil. **Material biológico:** Muestras de ADN extraídas de sangre periférica y tejido hepático parafinado. **Principales medidas de resultados:** Presencia/Ausencia de la mutación E8SJM. **Resultados:** Se estandarizó la reacción de PCR en tiempo real, la mutación fue detectada correctamente y posteriormente validada por secuenciación de Sanger. La mutación fue analizada en 137 muestras y encontrada en apenas una paciente que ingresó al Servicio de Genética Médica del HCPA con diagnóstico clínico y bioquímico de CESD/Wolman. **Conclusiones:** La técnica de PCR en tiempo real es ideal para la detección rápida y en gran escala de la mutación frecuente asociada a CESD.

Palabras clave: Enfermedades por Almacenamiento Lisosomal; Esterol Esterasa; Enfermedad de Acumulación de Colesterol Éster; Enfermedad de Wolman; Técnicas de Diagnóstico Molecular.

Abstract

Introduction: Cholesteryl Ester Storage Disease (CESD) is a lysosomal storage disorder, its presentation is highly variable and its diagnosis challenging. In addition, there are several abnormalities observed in patients with CESD who overlap with more common diagnoses and are likely to be underdiagnosed. Most patients reported to date are carriers of the E8SJM mutation in the *LIPA* gene. In this sense, diagnostic assistance is essential because there are options for therapy in development, as well as mutation prevalence studies. **Design:** Observational research. **Objective:** To standardize the real-time PCR technique for the detection of the most frequent mutation, E8SJM, in peripheral blood and liver biopsy specimens for the diagnosis of CESD and future mutation prevalence studies. **Institution:** Center of Gene Therapy of the Hospital de Clinicas de Porto Alegre (HCPA), Brazil. **Biological material:** DNA samples extracted from peripheral blood and paraffin-embedded liver tissue. **Main outcome measures:** Presence / Absence of E8SJM mutation. **Results:** The PCR reaction was standardized in real time; the mutation was correctly detected and validated by Sanger sequencing. The mutation was analyzed in 137 samples and found in only one patient who entered the Medical Genetics Service of the HCPA with clinical and biochemical diagnosis of CESD/Wolman. **Conclusions:** The real-time PCR technique is ideal for rapid and large-scale detection of the frequent CESD-associated mutation.

Keywords: Lysosomal Storage Diseases; Sterol Esterase; Cholesterol Ester Storage Disease; Wolman Disease; Molecular Diagnostic Techniques.

INTRODUCCIÓN

La enfermedad por Almacenamiento de Ésteres de Colesterol (CESD; OMIM 278000) es una enfermedad rara, de herencia autosómica recesiva, que pertenece al grupo de enfermedades por depósito lisosomal. La CESD es causada por mutaciones en el gen ***LIPA*** (10q23.2-23.3) que afectan la actividad enzimática de la enzima lipasa ácida lisosómica (LAL, EC 3.1.1.13) [1]. Este gen, compuesto por 10 exones, codifica una proteína de 399 aminoácidos. Hasta el momento fueron descritas más de 30 mutaciones (de sentido erróneo, sin sentido, en sitios de ***splicing***, pequeñas y grandes deleciones y rearreglos complejos [2] (Figura 1).

La enzima LAL es responsable por la hidrólisis intracelular de triglicéridos y de ésteres de colesterol derivados de las lipoproteínas plasmáticas [3]. Su deficiencia lleva a la acumulación progresiva de estas moléculas en varios órganos y subsecuentemente causa enfermedad hepática, niveles elevados de transaminasas y de esteres de colesterol LDL en el suero. Ocurre hepatomegalia, causado por esteatosis hepática y fibrosis, que pueden llevar a cirrosis micronodular y muerte [4].

La presentación clínica de la CESD es bastante variable. Mientras que algunos pacientes presentan enfermedad hepática en la infancia, otros no son diagnosticados hasta que las complicaciones de manifiestan en la edad adulta [5].

El diagnóstico de la CESD es un desafío. En ocasiones la hepatomegalia es el signo principal, y a veces, el único signo clínico. En el caso de que se realice una biopsia hepática, se observa la combinación de acumulación de macrófagos con evidencia microscópica de almacenamiento lipídico intralisosomal [5]. Bioquímicamente, la enfermedad es reconocida por una marcada reducción en la actividad de la lipasa ácida lisosomal (con actividad residual debajo de 10% de lo normal) [6,7]. Molecularmente, la enfermedad es confirmada por análisis del gen ***LIPA*** y la detección de variantes patogénicas. La principal técnica utilizada es la del PCR convencional donde cada uno de los 10 exones es amplificado y secuenciado utilizando el método de Sanger [4].

Un factor complicador es el hecho de que algunas características clínicas, radiológicas y bioquímicas de la CESD son comunes en pacientes que presentan otras enfermedades hepáticas, incluyendo la enfermedad hepática grasa no alcohólica (***NAFLD, nonalcoholic fatty liver disease***), la esteatosis hepática no alcohólica (NASH, ***nonalcoholoic steatohepatitis***) y la cirrosis criptogénica, entre otras [8-11]. Por lo tanto, es probable que los pacientes con CESD sean erróneamente diagnosticados, o no sean diagnosticados. La prevalencia estimada de CESD en la población caucásica e hispánica es de ~0.8 cada 100 000 (~1 en 130 000; 95% CI: ~1 en 90 000 a 1 en 170 000) [12,13].

La grande mayoría de pacientes con CESD, aproximadamente 60% (95% IC: 51%-69%), son portadores de la mutación E8SJM (c.894G>A) en por lo menos un alelo [4,13,14]. Esta mutación introduce un sitio de splicing alternativo, resultando en la deleción del exon 8 en el mRNA y originando una proteína 24 aminoácidos más corta. Esta observación justifica la utilización de la detección de esta mutación como un procedimiento inicial en el proceso de genotipaje y hasta mismo como herramienta auxiliar en el proceso diagnóstico. El desarrollo de la terapia de reposición enzimática para la CESD hace que el diagnóstico necesite ser rápido, específico, siendo de extrema importancia, una vez que la terapia parece cambiar substancialmente la historia natural de la enfermedad [15,16].

Sin embargo, antes de que sea aplicable, es necesario la estandarización de las técnicas implicadas, desde el procesamiento de las muestras hasta el análisis e interpretación de los resultados.

MÉTODOS

Material Biológico

Fueron utilizados: 1) noventa y seis muestras de sangre periférica (EDTA) de individuos sanos como controles normales, 2) una muestra positiva para la mutación E8SJM, extraída de sangre periférica y 3) cuarenta muestras de tejido hepático parafinado proveniente de pacientes sometidos a trasplante hepática en el Hospital de Clínicas de Porto Alegre, con diagnóstico prévio de NASH o de cirrosis criptogénica. 4) Además, fue incluida en el análisis una muestra de ADN extraída de sangre periférica de una paciente que entró al Servicio de Genética Médica del HCPA (SGM-HCPA) con diagnóstico clínico y bioquímico de CESD/Wolman. La paciente, de 3 años de edad, presentaba hepatomegalia, esteatosis hepática y en los exámenes laboratoriales un discreto aumento de aminotransferasas y aumento moderado de colesterol y triglicéridos. Los resultados histológicos y de microscopia electrónica sugerían enfermedad de Wolman o de Depósito de ésteres de colesterol (CESD). Bioquímicamente se detectó una actividad enzimática de la lipasa ácida en leucocitos disminuida: 1,1nmol/h/mg prot (valor de referencia: 112-378).

MÉTODOS

La extracción de ADN a partir de sangre periférico fue realizada con el kit comercial Easy-DNA® (Invitrogen, USA), de acuerdo a las indicaciones del fabricante. La extracción de DNA a partir de tejido hepático parafinado fue realizada utilizando el método de Coura et al (2005). Los ADN obtenidos fueron cuantificados y la pureza fue analizada en el espectrofotómetro NanoDrop® ND-1000 (Thermo Scientific, USA).

La identificación de la mutación frecuente, E8SJM, fue realizada por PCR en tiempo real utilizando el sistema de genotipaje *Taqman® SNP*, utilizando sondas y cebadores personalizados: Cebadores: 5'-CTG GAA CTT CTG TGC AAA ACA TGT-3' y 5'-CCC CAA ATG CAC TCC TGG AA-3'; Sonda Taqman 1 para la secuencia normal: VIC/5'-TGG AGC CAG GTA GGC-3'/NFQ; Sonda Taqman 2 para la mutación E8SJM: FAM/5'-TGG AGC CAA GTA GGC-3'/NFQ) y el StepOnePlus™ Real-Time PCR System (Applied Biosystems, USA). La sonda Taqman de la secuencia normal y de la mutación E8SJM fueron marcadas con los fluorocromos VIC (verde) y FAM (azul), respectivamente.

La reacción de PCR en tiempo real fue realizada siguiendo las instrucciones del fabricante, con algunas modificaciones,

Tabla 1. Resumen de algunos resultados de la extracción de DNA y los genotipos determinados por ensayo TaqMan.

Denaturación inicial	95°C	10 minutos	1 ciclo
Denaturación	95°C	15 segundos	40 ciclos
Alineamiento	58°C	30 segundos	
Extensión	60°C	30 segundos	
Extensión final	58°C	1 minuto	1 ciclo

utilizando 100ng de ADN por reacción, cada muestra fue analizada en triplicado. El programa de amplificación incluye lo siguiente (ver tabla 1).

La colecta de datos es realizada a medida que la reacción de PCR ocurre, donde es detectada la señal fluorescente producida proporcionalmente durante la amplificación del ADN blanco. Los productos de PCR fueron analizados utilizando el gráfico de multicomponentes proporcionado por el StepOnePlus™ software (v2.3). El análisis de los datos se traduce en la evaluación de las curvas de amplificación (Figura 2A), en las que se representa la fluorescencia detectada versus el número de ciclos de PCR. Los resultados fueron expresados en ausencia/presencia de la mutación.

Para la confirmación del alelo positivo para la mutación E8SJM en el PCR en Tiempo Real, la secuenciación se realizó por el método de Sanger del exon 8 utilizando los ***primers forward*** 5'-TTAGTGCTTTGAAGGGCAAAA-3' y ***reverse*** 3'-TCTATTTGGAAAGGGTTTGCAT-5', diseñados utilizando el software ***online*** Primer3 v.0.4.0. No se utilizó ningún paquete estadístico pues el presente estudio es un análisis cualitativo de ausencia/presencia de mutación.

Tabla 2. Resumen de algunos resultados de la extracción de DNA y los genotipos determinados por ensayo TaqMan.

Sangre periférico			Tejido hepático parafinado		
Código	Conc. DNA (ng/μL)	Genotipo	Código	Conc. DNA (ng/μL)	Genotipo
1	200.5	G/G	11	51.8	G/G
2	180.2	G/G	12	71.0	G/G
3	130	G/G	13	101	G/G
4	145.1	G/G	14	61.8	G/G
5	269	G/G	15	71.3	G/G
6	150.6	G/G	16	52.8	G/G
7	324	G/G	17	52.7	G/G
8	186.1	G/G	18	46.9	G/G
9	192	G/G	19	44.8	G/G
10	210.1	G/G	20	46.4	G/G
Control	99.1	G/A			
Paciente con sospecha clínica de CESD	109	G/A			

Consideraciones Éticas

La presente investigación fue aprobada por sus aspectos éticos y metodológicos por el Comité de Ética en Investigación (CEP) del Hospital de Clínicas de Porto Alegre (proyecto 13-0189).

RESULTADOS

La extracción de ADN a partir de sangre periférica fue realizada con éxito. Se obtuvo un promedio de 205 ng/uL con razón de pureza (DO 260/280) de 1.85. Originalmente el protocolo propuesto por Coura et al (2005) fue utilizado para la extracción de ADN a partir de tejido de carcinoma colorrectal parafinado, esta técnica se caracteriza por ser una alternativa simple y de bajo costo. Esta técnica resultó ser reproducible en nuestra investigación, permitiéndonos extraer ADN con cantidad y calidad suficiente para la realización de los análisis moleculares. Para el ADN extraído de tejido parafinado se obtuvo una concentración promedio de 54ng/uL con razón de pureza (DO 260/280) de 1.88. El detalle de las concentraciones de algunas muestras de DNA analizadas en este estudio se encuentra en la tabla 2.

La Figura 2A muestra el gráfico de la cinética de amplificación por PCR en tiempo real del DNA extraído a partir de sangre periférica. Durante los ciclos iniciales de la reacción de PCR no se presentó ningún cambio significativo de la intensidad de la fluorescencia emitida. La referencia pasiva (línea roja) permaneció constante a lo largo del proceso de PCR. Se observó que el número de ciclos necesarios para comenzar a obtener señal de fluorescencia detectable fue en promedio 26 (valor Ct). En el control positivo, se observa la amplificación del alelo normal (verde) y del alelo mutado (azul), correspondiendo a un heterocigoto para la mutación E8SJM. Para el ADN extraído del tejido hepático parafinado, el valor Ct fue en promedio 28. No fue detectada la mutación E8SJM en ninguna de las muestras de ADN extraídas de tejido parafinado de pacientes con diagnóstico de NASH o cirrosis criptogénica. La mutación tampoco fue detectada en ninguna muestra de los

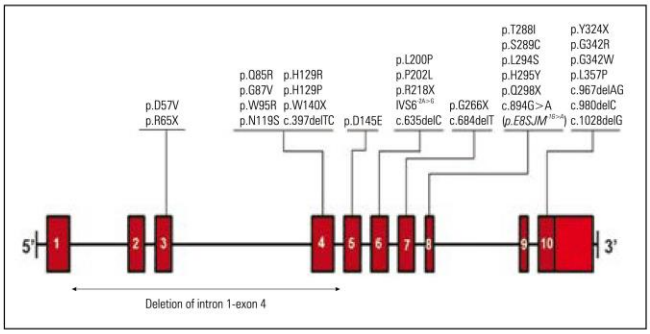

Figura 1. Mutaciones descritas en el gen LIPA[2].

lada de LDL, hiperlipidemia combinada, deficiencia de HDL y dislipidemia mixta; de los 1375 pacientes analizados, 6 individuos presentaron la mutación E8SJM, siendo 3 de estos, heterocigotos compuestos y así afectados con CESD [23]. La identificación de portadores y afectados en estas poblaciones, implica que la misma está enriquecida con portadores de mutaciones en el gen ***LIPA*** y subraya la importancia de las pruebas genéticas para las enfermedades raras en las cuales las manifestaciones clínicas pueden superponerse al fenotipo de trastornos más prevalentes. A pesar de la gran semejanza en las manifestaciones clínicas

controles normales. Esta mutación fue detectada en la paciente de 3 años que entró al SGM-HCPA con sospecha clínica de CESD (Figura 2A).

Debido a que la secuenciación por el método de Sanger ser considerado el "gold estándar" para el diagnóstico molecular de mutaciones, fue realizado la secuenciación del exon 8 para la confirmación del resultado obtenido por el PCR en tiempo real, confirmándose el estado en heterocigosis de la mutación E8SJM (Figura 2B).

DISCUSIÓN

La enfermedad por Almacenamiento de Ésteres de Colesterol (CESD) es una enfermedad genética rara. Aunque hasta la fecha se han descrito aproximadamente 155 casos en el mundo, es probable que la enfermedad sea más frecuente y las formas menos graves estén siendo subdiagnosticadas [4,17,18,19].

La dificultad de su diagnóstico radica en la heterogeneidad de la enfermedad y en la ausencia de síntomas específicos, los cuales se pueden sobreponer con las de otras enfermedades [20,21]. Un estudio reveló la presencia de casos de CESD en grupos de pacientes con hipercolesterolemia, de 276 individuos analizados, 5 presentaron una variante patogénica en el gen ***LIPA***, existiendo 2 heterocigotos para E8SJM [22]. Un segundo trabajo, con un objetivo similar, tamizó la mutación E8SJM en pacientes con elevación ais-

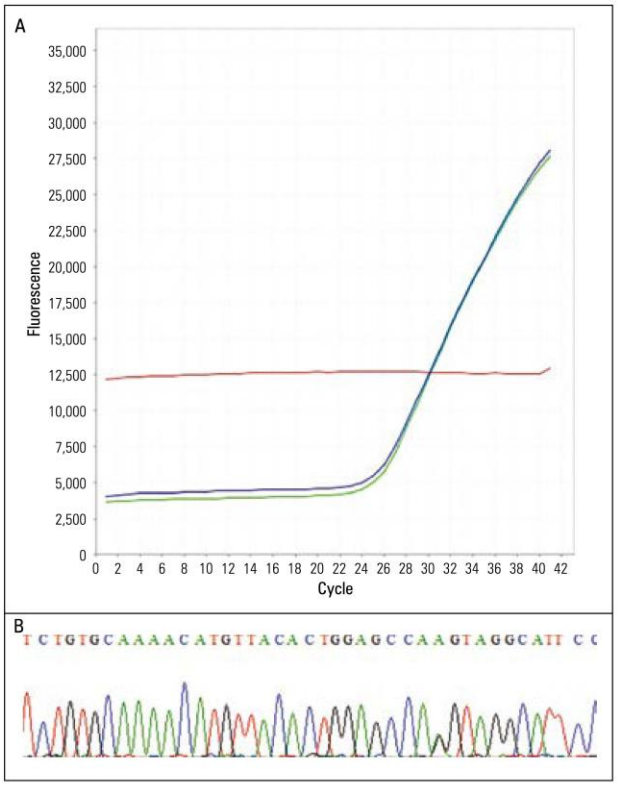

Figura 2. Detección por PCR en tiempo real (Taqman) de los alelos selvaje (verde) y mutado (azul) E8SJM. (A) Heterocigoto para la mutación E8SJM (B) Electroferograma de la secuencia en *forward* correspondiendo al exon 8 del gen *LIPA* mostrando la mutación frecuente p.E8SJM (c.894GA) en heterocigosis (flecha roja).

entre estas enfermedades hepáticas NAFLD, NASH, cirrosis criptogénica y CESD, no se ha publicado en la literatura algún trabajo donde haya sido estudiado una cohorte de pacientes, inicialmente diagnosticados con estas enfermedades con el fin de excluir CESD, confirmando así la hipótesis diagnóstica inicial.

Diferentes estudios han utilizado distintas fuentes para obtener el ADN para la detección de mutaciones en el gen ***LIPA*** (leucocitos extraídos de sangre periférica, sangre impregnada en papel filtro, fibroblastos, saliva, entre otros), así como diferentes metodologías para la detección de las variantes en el gen (MS-PCR, secuenciación por Sanger y RT-PCR) [4,12]. Existen otras técnicas moleculares modernas, entre ellos la secuenciación de nueva generación utilizando paneles de genes o secuenciación del exoma, que presentan varias ventajas como la rapidez de resultados y la posibilidad de analizar varios en paralelo; sin embargo, el costo de estos análisis hace inviable su utilización como herramienta de tamizaje [24]. En este trabajo proponemos el PCR en tiempo real como una herramienta para el tamizaje a gran escala de la mutación E8SJM (presente en 60% de los pacientes), asociada a la CESD [13]. En este escenario, nuestra investigación trae como novedad la utilización de ADN extraído de tejido hepático parafinado usando un protocolo casero de bajo costo, seguido de PCR en tiempo real para la detección de la mutación.

La técnica de PCR en tiempo real, con la utilización de las sondas ***Taqman***, fue estandarizada y permitió determinar rápidamente la presencia o ausencia de la alteración génica más común. A pesar de que la mutación E8SJM no fue detectada en ninguna de las muestras analizadas, la utilidad clínica de este ensayo fue demostrada con el análisis molecular de la paciente con diagnóstico clínico y bioquímico de CESD/Wolman. La detección de la mutación por PCR en tiempo real fue realizada de manera exitosa, encontrándose en estado de heterocigosis, lo cual significa que la paciente posee una copia del gen mutado y confirmado por secuenciación de Sanger (Figura 2B).

El PCR en tiempo real representa una técnica simple, rápida y confiable para un tamizaje inicial en el diagnóstico de CESD, pudiendo incluso ser utilizada en estudios de prevalencia de CESD. Contar con esta herramienta para el auxilio de tamizaje a gran escala de esta mutación es fundamental desde que no existen datos de prevalencia en Latinoamérica.

REFERENCIAS BIBLIOGRÁFICAS

1. Anderson RA, Bryson G M, Parks J S. Lysosomal acid lipase mutations that determine phenotype in Wolman and cholesterol ester storage disease. Molecular Genetics and Metabolism.1999;68(3):333–345. https://doi.org/10.1006/mgme.1999.2904.
2. Marin-Valencia I, Pascual J. Cap 36 Wolman Disease. Rosenberg's Molecular and Genetic Basis of Neurological and Psychiatric. Elsevier. Fifth Edition. 2015: 403- 410.
3. Human Genome Mutation Database (HGMD), consultado 23/11/2016. Disponible en: http://www.hgmd.cf.ac.uk/ac/index.php.
4. Bernstein DL, Hülkova H, Bialer MG, Desnick RJ. Cholesteryl ester storage disease: Review of the findings in 135 reported patients with an underdiagnosed disease. Journal of Hepatology. 2013; 58(6):1230–1243. https://doi.org/10.1016/j.jhep.2013.02.014.
5. Aguisanda F, Thorne N, Zheng W. Targeting Wolman Disease and Cholesteryl Ester Storage Disease: Disease Pathogenesis and Therapeutic Development. Current Chemical Genomics and Translational Medicine. 2017; 11: 1–18. https://doi.org/ 10.2174/2213988501711010001.
6. Weiler C, Freudenberg F, Müller-Höcker J. Cholesterol ester storage disease-A rare disease or a rare diagnosis?. Pathologe. 2009; 30(1): 65-69. https://doi.org/ 10.1007/s00292-009-1124-5.
7. Boldrini, Devito R, Biselli R, Filocamo M, Bosman C. Wolman disease and cholesteryl ester storage disease diagnosed by histological and ultrastructural examination of intestinal and liver biopsy. Pathology Research and Practice. 2004;200(3):231-240. https://doi.org/10.1016/j.prp.2003.11.001.
8. Hulkova H, Elleder M. Distinctive histopathological features that support a diagnosis of cholesterol ester storage disease in liver biopsy specimens. Histopathology.2012;60 (7): 1107-13. https://doi.org/10.1111/j.1365-2559.2011.04164.x.
9. Gasche C, Aslanidis C, Kain R, Exner M, Helbich T, Dejaco C, Schmitz G, Ferenci P. A novel variant of lysosomal acid lipase in cholesteryl ester storage disease associated with mild phenotype and improvement on lovastatin. J Hepatol. 1997; 27(4):744-50. https://doi.org/10.1016/S0168-8278(97)80092-X.
10. Chatrath H, Keilin S, Attar B. Cholesterol ester storage disease (CESD) diagnosed in an asymptomatic adult. Digestive diseases and sciences. 2009; 54 (1): 168-73. https://doi.org/10.1007/s10620-008-0310-2.
11. Hamilton J, Jones I, Srivastava R, Galloway P. A new method for the measurement of lysosomal acid lipase in dried blood spots using the inhibitor Lalistat 2. Clin. Chim. Acta. 2012; 413:1207–1210. https://doi.org/ 10.1016/j.cca.2012.03.019.
12. Muntoni S, Wiebusch H, Jansen-Rust M, Rust S, Seedorf U, Schulte H, et al. Prevalence of Cholesteryl Ester Storage Disease. Arteriosclerosis, Thrombosis, and Vascular Biology. 2007;27(8): 1866–1868. https://doi.org/10.1161/ATVBAHA.107.146639.
13. Scott S, Liu B, Nazarenko I, Martis S, Kozlitina J, Yang Y, et al. Frequency of the Cholesteryl Ester Storage Disease Common LIPA E8SJM Mutation (c.894G>A) in Various Racial and Ethnic Groups. Hepatology. 2013; 58(3): 958-965. https://doi.org/10.1002/hep.26327.
14. Fasano T, Pisciotta L, Bocchi L, Guardamagna O, Assandro P, Rabacchi C, et al. Calandra S. Lysosomal lipase deficiency: Molecular characterization of eleven patients with Wolman or cholesteryl ester storage disease. Molecular Genetics and Metabolism. 2012; 105(3): 450–456. https://doi.org/10.1016/j.ymgme.2011.12.008.
15. Balwani M, Breen C, Enns GM, Deegan PB, Honzík T, Jones S, et al. Clinical effect and safety profile of recombinant human lysosomal acid lipase in patients with cholesteryl ester storage disease. Hepatology. 2013; 58(3): 950–957. https://doi.org/10.1002/hep.26289.
16. Burton BK, Balwani M, Feillet F, Barić I, Burrow TA, Camarena C, et al. A Phase 3 Trial of Sebelipase Alfa in Lysosomal Acid Lipase Deficiency. New England Journal of Medicine. 2015; 373(11): 1010–1020. https://doi.org/10.1056/NEJMoa1501365.
17. Kamenet E, Mollaki V, Hamilton J, Zakharova E & Drogari E. Molecular analysis of 15 homozygous LAL-D patients from Greece using DBS screening. Atherosclerosis. 2017; 263: e100-e101. https://doi.org/10.1016/j.atherosclerosis.2017.06.328
18. Carbajal-Rodríguez L. Deficiencia de lipasa ácida lisosomal. Reporte de dos casos. Revista Mexicana de Pediatria. 2016; 83 (2): 49-54
19. Reynders J, Burton BK, del Angel G. Novel LIPA mutations resulting in lysosomal acid lipase deficiency. Molecular Genetics and Metabolism. 2017; 20 (1-2): S114–S115. https://doi.org/10.1016/j.ymgme.2016.11.294
20. Bernstein D, Desnick R. Cholesteryl ester storage disease (CESD): An under-recognized and treatable LSD with liver dysfunction/failure and dyslipidemia. Molecular Genetics and Metabolism. 2013; 108(2): S24. https://doi.org/10.1016/j.ymgme.2012.11.036
21. Basu P, Nair T, Farhat S, Foustin S, Ang L, Anwarullah A, et al. Unrevealing a novel association of cholesterol ester storage disease (CESD) & non-alcoholic fatty liver disease(NAFLD)-a similar clinical spectrum with different etiology a prospective clinical study. Journal of Hepatology. 2012; 56(2): S508. https://doi.org/10.1016/S0168-8278(12)61299-9
22. Sjouke B, Defesche JC, de Randamie JSE, Wiegman A, Fouchier SW, Hovingh GK. Sequencing for LIPA mutations in patients with a clinical diagnosis of familial hypercholesterolemia. Atherosclerosis. 2016; 251: 263-265. http://dx.doi.org/10.1016/j.atherosclerosis.2016.07.008
23. Kane JP, Stock E, Pullinger CR, Movsesyan I, Malloy MJ, Quinn A. Identifying Cases of Cholesteryl Ester Storage Disease in a Tertiary Lipid Clinic. Journal of Clinical Lipidology. 2013; 7(3): 249. http://dx.doi.org/10.1016/j.jacl.2013.03.034
24. Stitziel NO, Fouchier SW, Sjouke B, Peloso GM, Moscoso AM, Auer PL, et al. Blood Institute GO Exome Sequencing Project. Exome sequencing and directed clinical phenotyping diagnose cholesterol ester storage disease presenting as autosomal recessive hypercholesterolemia. Arteriosclerosis, thrombosis, and vascular biology. 2013; 33(12): 2909-14. https://doi.org/10.1161/ATVBAHA.113.302426

ANEXO III– Artigo 7: Recent advances in molecular testing to improve early diagnosis in children with mucopolysaccharidosis

---

Aceito para publicação na revista EXPERT REVIEW OF MOLECULAR DIAGNOSTIC, 2018.

EXPERT REVIEW OF MOLECULAR DIAGNOSTICS
2018, VOL. 18, NO. 10, 855–866
https://doi.org/10.1080/14737159.2018.1523722

Taylor & Francis
Taylor & Francis Group

REVIEW

Recent advances in molecular testing to improve early diagnosis in children with mucopolysaccharidoses

Ana Carolina Brusius-Facchin[a], Diana Rojas Malaga[a,b], Sandra Leistner-Segal[a,c] and Roberto Giugliani [a,b,c,d]

[a]Medical Genetics Service, HCPA, Porto Alegre, RS, Brazil; [b]Postgraduate Program of Genetics and Molecular Biology, UFRGS, Porto Alegre, RS, Brazil; [c]Postgraduate Program in Medical Science, UFRGS, Porto Alegre, RS, Brazil; [d]Department of Genetics, UFRGS, Porto Alegre, RS, Brazil

ABSTRACT

Introduction: The mucopolysaccharidoses (MPS) are a group of lysosomal storage disorders with high phenotypic and genotypic heterogeneity, making precise diagnosis challenging. Although enzyme activity assay is considered the gold standard for the diagnosis of these disorders, molecular testing can greatly refine this task. New methods for rapid detection of variants are useful to reduce the 'diagnostic odyssey' faced by patients and their family, to lead to appropriate genetic counseling and to select the most appropriate therapy for each case.

Areas covered: We review and discuss the advantages, disadvantages and limitations of the modern technologies in the field of molecular diagnosis of MPS, presenting our own experience.

Expert commentary: While current molecular genetics testing for MPS mostly relies on PCR and Sanger sequencing, promising alternative techniques have emerged over the last few years, and its application into routine clinical practice is gaining momentum.

ARTICLE HISTORY
Received 25 April 2018
Accepted 11 September 2018

KEYWORDS
Molecular genetic testing; lysosomal storage disorders; mucopolysaccharidoses; next generation sequencing; mutation identification

1. Introduction

Lysosomal Storage Disorders (LSD) are a heterogeneous group of at least 50 rare genetic disorders caused by progressive accumulation of specific substrates, generally due to deficiency of a lysosomal enzyme, and less frequently to defects in non-enzymatic lysosomal protein, transmembrane protein and trafficking, among others [1].

The mucopolysaccharidoses (MPS) represent the second most prevalent group of LSD, after the sphingolipidoses [2]. Including the recently described MPSPS (mucopolysaccharidosis-plus syndrome) [3,4], it consists of 12 different types and subtypes of monogenic disorders (Table 1), with a combined birth prevalence ranging from 2.2 per 100 000 in British Columbia to 5.2 per 100 000 live births in the United Arab Emirates [5–11]. However, preliminary data from newborn screening programs indicate that the prevalence could be considerably higher [12]. All MPS have autosomal recessive inheritance, with the exception of Hunter syndrome (MPS II), which is X-linked recessive.

These disorders are caused by mutations in genes that encode lysosomal hydrolases, which are responsible for the catabolism of glycosaminoglycans (GAGs) – dermatan sulfate, heparan sulfate, keratan sulfate, chondroitin sulfate, and hyaluronan – resulting in its accumulation in tissues of affected individuals [1]. Depending on the specific enzyme deficiency involved, the GAGs catabolism may be blocked in a particular step. There are 11 known enzyme deficiencies, each one resulting in a different subtype of MPS; for MPSPS no lysosomal enzyme deficiency has been reported (Table 1). The accumulation of partially degraded GAGs causes progressive multisystem disorders with variable penetrance and severity, which are not evident at birth in most cases, but with features becoming apparent usually in the first years of life, depending on the MPS type/subtype [13,14]. All of this makes MPS early identification challenging, causing a significant delay between symptoms onset and precise diagnosis, especially in patients with attenuated phenotype where the final diagnosis could take even decades or remain undiagnosed [15–18].

2. Mutation spectrum

The wide spectrum seen in the MPS is mainly caused by a high allelic heterogeneity in the MPS-related genes and thus differences cannot always be correlated with residual enzyme activity [19]. Precise molecular testing can potentially predict the disease severity and future outcomes. For example, some mutations have been associated with milder phenotypes: MPS I (p.Ser633Trp [20]), MPS IIIA (p.Arg206Pro [21], p.Ser347Phe/p.Asp444Gly [22] and p.Glu369Lys/p.Pro128Leu [23]), MPS IIIC (p.Gly262Arg and p.Ser539Cys [24,25]); and others with severe phenotype: MPS I (p.Trp402Ter [26] and p.Gln70Ter), MPS II (p.Ser333Leu and IDS total gene deletions [27,28]), MPS IIIA (p.Arg433Gln [23]).

To date, more than 2100 mutations have been reported in all 12 genes related to MPS, with the majority of individuals showing private mutations (~70%) [29,30]. This broad mutational spectrum is composed mostly by missense/nonsense variants (65%), followed by small deletion/insertions/indels (~19%), splicing defects (~8%), complex rearrangements

CONTACT Roberto Giugliani rgiugliani@hcpa.edu.br Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos 2350, 90035-903 Porto Alegre, RS, Brazil

Table 1. The mucopolysaccharidoses.

Type	Eponym	Deficient enzyme	GAGs stored	Gene	Chr.	OMIM
I	Hurler (IH)	α-L-iduronidase	DS, HS	*IDUA*	4p16	607,014
	Scheie (IS)					607,016
	Hurler-Scheie (IH/S)					607,015
II	Hunter	Iduronate sulfatase	DS, HS	*IDS*	Xq28	309,900
III	Sanfilippo A	Heparan N-sulfatase	HS	*SGSH*	17q25	252,900
	Sanfilippo B	α-N-acetylglucosaminidase		*NAGLU*	17q21	252,920
	Sanfilippo C	Acetyl CoA: α-glucosaminide acetyltransferase		*HGSNAT*	8p11	252,930
	Sanfilippo D	N-acetylglucosamine 6-sulfatase		*GNS*	12q14	252,940
IV	Morquio A	Galactosamine-6-sulfate sulfatase	KS, CHS	*GALNS*	16q24	253,000
	Morquio B	β-galactosidase	KS	*GLB1*	3p21	253,010
VI	Maroteaux-lamy	Arylsulfatase B	DS, CHS	*ARSB*	5q14	253,200
VII	Sly	β-glucuronidase	DS, CHS, HS	*GUSB*	17q21	253,220
IX	Natowicz	Hyaluronidase	H	*HYAL1*	3p21	601,492
MPSPS	Mucopolysaccharidosis-plus syndrome	Not an enzyme deficiency	HS	*VPS33A*	12q24	617,303

Chr, chromosome, HS: heparan sulfate, D.S: dermatan sulfate, K.S: keratan sulfate, CHS: chondroitin-6-sulfate, H: Hyaluronan.

(4%), gross deletion/insertion/indels (3.4%) and defects in regulatory regions (0.2%) (http://www.hgmd.cf.ac.uk/ac) (Table 2).

In MPS II, approximately 9% of all disease-causing mutations correspond to complex rearrangements involving the iduronate-2-sulfatase gene (*IDS*) and its homologous pseudogene (*IDS-2, IDSP1*), located in inverted orientation 25kb telomeric to the functional gene [31–33]. The presence of the pseudogene likely renders the *IDS* gene more prone to recombination [34]. *IDS-2* is homologous to exon 2 and 3, and introns 2, 3, and 7 of the transcribed *IDS*, with exon 3 being 100% identical and the other sequences with 96% identity [35].

The frequency of mutations in different genes within and between populations varies due several factors, such as selective pressure, population isolation, uniparental isodisomy (UPD), and mosaicism. The most important ones that have been described for MPS, are founder effect, that is associated to population isolation, UPD, a rare condition that could occur an autosomal recessive disease, and mosaicism.

An example of founder effect is a mutation in the *ARSB* gene (p.His178Leu) that causes MPS type VI disease and is found in high frequency in the Northeast of Brazil. The incidence of MPS VI varies from 1 in 43,000 to 1 in 1,505,000 births [36]. However, in the county of Monte Santo, located in the countryside of the state of Bahia, the frequency of the disease is 1: 5,000 inhabitants [37]. Similarly, results of *ARSB* gene analysis in patients from Poland, Belarus, Lithuania, and Estonia showed a recurrent mutation (p.Arg152Trp), suggesting a founder effect also in that region [38]. Another example occurs in MPS type IVA patients from the Northeast region of Brazil, mainly from the state of Paraiba. There is a high frequency of consanguinity in this region, originally colonized by a few Portuguese families [39,40]. The high frequency of the same mutation (p.Ser341Arg) (with the same haplotype) and high rate of parental consanguinity provides a strong evidence for the hypothesis of a founder effect [41].

Other mutations with high frequency within certain populations were described for MPS IIIA (*SGSH*): p.Ser66Trp (Italy) [42]; MPS IIIB (*NAGLU*): p.Arg234Cys (Portugal) [43] and p.Tyr140Cys, p. His414Arg and p.Arg626Ter (Greek patients) [44]; MPS IIIC (*HGSNAT*): c.372-2A>G (Spain) [45], c.234+1G>A (Spain and Morocco) [45], p.Arg344Cys and p.Ser518Phe (Netherlands) [25], c.525dupT (Portugal) [46,47], c.852-1G>A (South of Italy) [48]; MPS IVA (*GALNS*): p.Ile113Phe and p.Thrp312Ser (Northern Ireland) [49,50] and p.Gly301Cys (Colombia) [51]; MPS IVB (*GLB1*): p. Trp273Leu (European regions) [52]; MPS VI (*ARSB*): p.His147Pro, p.His111Pro and p.Ser334Ile in Colombia; and probably in MPSPS (*VPS33A*): p.Arg498Trp (Russia/Yakutia Republic), which suggests a regional founder effect, though further experiments are necessary to prove this hypothesis [3].

Table 2. Systematic review of the mutation characteristics.

			Mutation type (%)					
Disorder	Gene	Mutations reported*	M/N	S	R	SD/SI/SID	GD/GI/GID	CR
MPS I	*IDUA*	257	57.5	15	0.5	23	3	1
MPS II	*IDS*	628	50	9	0	29	9	3
MPS IIIA	*SGSH*	145	76	2	0	19	3	0
MPS IIIB	*NAGLU*	166	69	4	0	23	4	0
MPS IIIC	*HGSNAT*	68	56	20.5	0	16	6	1.5
MPC IIID	*GNS*	25	28	16	0	40	8	8
MPS IVA	*GALNS*	333	74.5	9.5	0	12	3	1
MPS IVB	*GLB1*	215	77	7.5	0	14	0.5	1
MPS VI	*ARSB*	197	75	6	0	16	3	0
MPS VII	*GUSB*	64	83	8	2	6	1	0
MPS IX	*HYAL1*	3	34	0	0	33	0	33
MPSPS	*VPS33A*	1	100	0	0	0	0	0
		Total mean	65.00	8.10	0.20	19.30	3.40	4.00

M/N: missense/nonsense, S: splicing, R: regulatory, SD: small deletions, SI: small insertions, SID: Small indels, GD: gross deletions, GI: gross insertions, GID: Gross indels, CR: complex rearrangements.
*HGMD professional 2017.3 (consulted January 2018).

UPD has been described only in one patient with MPS type IVA who inherited the two copies of a mutation from one parent [53], which is not usually observed in recessive diseases. This reinforces that parental analysis should be mandatory for an accurate genetic counseling regarding recurrence risk.

Somatic mosaicism for a disease-causing variant in *IDUA* (whole gene deletion) was reported in the mother of a single family [54].

3. MPS diagnosis

The diagnosis of MPS is based mainly in biochemical and molecular testing. Usually the first step in the diagnosis pathway is the urine screening with quantitation and qualitative evaluation (by electrophoresis or thin-layer chromatography) of urinary GAGs [55,56]. These results, together with the clinical findings, can help guiding the choice of the confirmatory enzyme assay to be performed. Assays to measure activity of specific enzymes in leukocytes or fibroblasts are considered the gold standard for a definitive diagnosis. For some enzymes, assays in plasma are also available. In some cases, it is possible to perform the enzyme assay in dried blood spots (DBS) [10]. This option is instrumental in regions where collecting and shipping whole blood or other tissue samples is impractical. The present recommendation is that positive results in DBS should be confirmed in leucocytes or fibroblasts. When this is not possible, enzyme assays should be performed at least twice in DBS (in two independent samples) and/or should have the results confirmed by genotyping [10]. Although the finding of a specific enzyme deficiency in leucocytes or fibroblasts confirms the diagnosis, molecular analysis of the respective gene is recommended, whenever possible (Figure 1).

Molecular genetics has several roles in the MPS diagnosis: (1) If screening (urine and DBS) and enzyme analysis results are inconclusive, molecular analysis is required to discriminate pseudodeficiency (when enzyme activity is deficient only *in*

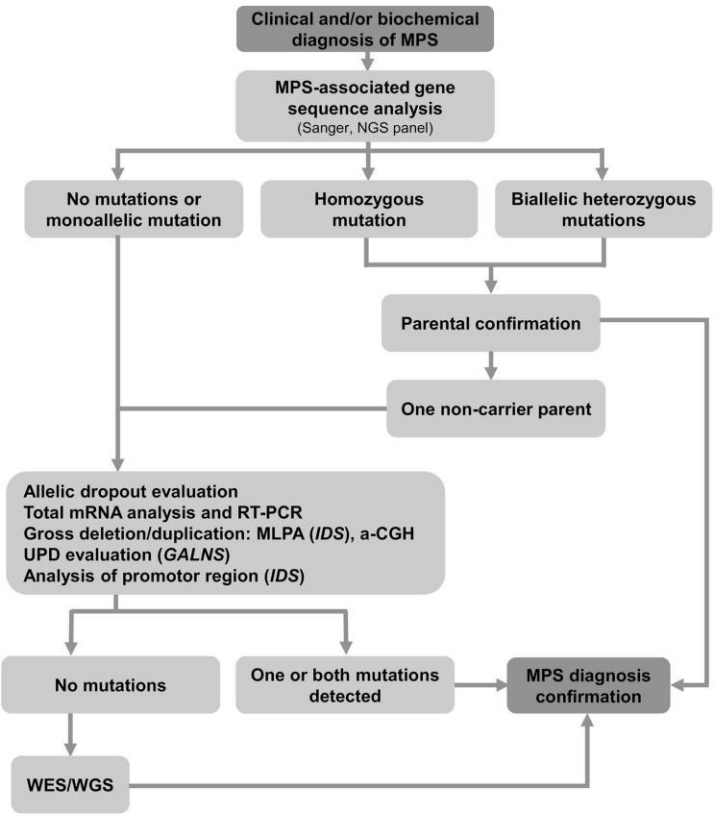

Figure 1. Flowchart of molecular investigation for MPS. MPS, mucopolysaccharidosis; MLPA, multiplex ligation-dependent probe amplification; NGS, next-generation sequencing; mRNA, RNA messenger; RT-PCR, real-time PCR; a-CGH, microarray-based comparative genomic hybridization; UPD, uniparental disomy; WES, whole exome sequencing; WGS, whole genome sequencing.

vitro, with no clinical consequences) and carrier status from affected or normal cases [57]; (2) Results of enzyme assays in DBS, can be combined with molecular analysis to reach a final diagnosis [10]; (3) phenotype prediction, which may be important for management decisions [58]; (4) identification of suitability of the patient to a mutation-specific therapy, as stop codon read through [59–61]; and (5) prenatal diagnosis, alternatively or in addition to biochemical diagnosis [62,63].

4. MPS newborn screening

Over the years, some MPS (I, II, IVA, and VI) have been candidates to be evaluated for inclusion in NBS programs around the world. As knowledge on the prevalence increases, screening methods and treatment options become increasingly available and several pilot programs of large scale NBS have been completed. For MPS I, pilot studies, in Taiwan, Italy, Korea, Austria, and Hungary, have demonstrated the necessity of including this disorder in the list of routine recommended NBS programs, since disease-specific treatments as ERT and HSCT are available and early introduction leads to better outcomes [64]. Besides these studies, some countries are currently evaluating MPS I as candidate for routine NBS programs (Netherlands and UK). In the US, MPS I, as other three LSDs (Pompe, Gaucher and Fabry, was recently added to this list and are underway.

Currently, DNA sequencing is used only as a second-tier test for certain diseases or as a confirmatory test for a positive newborn screen. Notably, mutation analysis increases sensitivity, reducing the false-positive rates of primary assay, and clarifying ambiguous results, as obtained in carriers and pseudodeficiency cases [57].

IDUA pseudodeficiency alleles (as p.Ala79Thr, p.His82Gln, p.Asp223Asn, and p.Val322Glu) can result in decreased enzyme activities when found in homozygosis or in compound heterozygosis with another pseudodeficiency allele, a pathogenic variant or a variant of unknown significance [65]. Pseudodefiency alleles have also been identified in MPS VII (p.Asp152Asn [66]) and MPS VI (p.Thr212Ile [67]).

5. Molecular genetics analysis in the MPS

A range of different molecular approaches for investigations of the disease-causing mutations, family screening, and genetic counseling are available [30], each one with its own indications and limitations (Table 3). Molecular analysis is especially required for the diagnosis of MPSPS, which results from non-enzymatic lysosomal protein deficiency. The interpretation of molecular genetics results and the diagnostic conclusion should take into consideration clinical and biochemical data.

Biochemical and genetic testing may be performed to confirm and classify the MPS disorders. Each type of MPS has a known mutation that is specific to the diseases type. In some MPS, differentiation between subtypes could be tricky. For MPS II, for example, there are two recognized clinical entities (mild and severe) and mutation analysis is often helpful in distinguishing between the two of them. For MPS III and IV, although little clinical difference exists between subtypes, they can be easily distinguished by enzymatic assay of the different enzymes involved in heparan and keratan sulfate degradation, respectively.

5.1. Sanger sequencing

Sanger sequencing remains as the main method for identification of genetic variations in monogenic disorders when the gene involved is known. The sequence analysis based on Sanger´s method can detect point mutations and small insertions and deletions and is still considered the gold standard for DNA sequencing, including several studies reported for MPS [68–72]. However, due to the high level of allelic heterogeneity and the fact that this methodology can only analyze one DNA segment/exon at a time, it became a labor intensive, time-consuming and expensive process. Also, caution must be given when interpreting sequencing results. In autosomal recessive disorders it is important to determine if the pathogenic mutations found are in different alleles (*trans*), since it can affect diagnostic conclusions and especially genetic counseling and prenatal diagnosis (Figure 2(a)). If both mutations are present in the same allele (*cis*), the subject might be an unaffected carrier due to expression of the normal allele or he/she still could be an affected patient but due to the presence of an unidentified mutation.

Currently, Sanger sequencing is usually the molecular genetics method chosen to investigate subjects in a family with a known mutation in a specific MPS gene, to identify carriers or for prenatal diagnosis. In some centers, Sanger sequencing is also applied to fill any potential gaps that NGS-based approach may leave [73].

For MPS II, the presence of a pseudogene can complicate molecular genetics analysis as it can be simultaneously amplified in the PCR reaction leading to an amplification bias and erroneous genotyping. Nevertheless, this can be overcome through the design of gene-specific primers chosen to avoid the amplification of IDS-2 [74].

5.2. Other types of analysis

In general, sequence analysis has the potential of detecting pathogenic variants in 80–99% of probands (for MPS II and MPS IVB, respectively) [75]. For the detection of variants, other than point mutations and small indels, methods based on gene targeted deletion/duplication analysis: quantitative polymerase chain reaction (qPCR), long-range PCR, MLPA (Multiplex Ligation-dependent Probe Amplification), gene-targeted microarray designed to detect single-exon, and multi-exon deletions or duplications can be used to complement the molecular strategy methods for the analysis of complex rearrangements include: sequencing, SNP (Single Nucleotide Polymorphism) analysis, gene-targeted deletion/duplication analysis, and chromosomal microarray (CMA) [75].

Analyses of the promoter region of the IDS gene should be considered in patients with previous biochemical diagnosis of MPS II and without detecting any pathogenic variant after sequencing of the whole IDS coding region and exon-intron boundaries [76].

Table 3. General indication and limitations of the different molecular techniques recommended for MPS diagnosis.

Test method	Indication	Limitation
Sanger sequencing		To detect DNA variants located in coding regions of the associated gene and small segments of immediately adjacent intronic regions, including point mutations, small insertion/deletions and splice site variants.
Changes in copy number variations (CNV) are missed.		
RT-PCR	Ideally suited for screening of a specific point mutations.	Requires optimization for each locus analyzed Novel, unreported variants are not detected.
HRM	Determination of DNA methylation status and genotyping.	Changes in CNV are missed.
MLPA	To establish the presence of CNV.	Necessary to confirm results with other techniques (Sanger sequencing, NGS technologies, and a-CGH techniques among other techniques) or with different probes than the first probemix, due to interference from polymorphisms.
CGH and a-CGH	To detect gross deletions and duplications, unbalance rearrangements and deletion boundary which are not detected by sequence analysis.	Will not detect balance rearrangements and low level mosaic imbalances may not be detected as well.
NGS	Detection of point mutations and small insertions and deletions in many different genes in a cost and time efficient manner.	Contribution of pseudogenes sequences (e.g. MPS II) or other highly homologous sequences. Sanger sequencing is still necessary for variant confirmation, and to meet coverage standards. Translocation or inversions, repeat expansions, are not detected.

5.2.1. Real-time PCR (RT-PCR)

Real-time PCR is based on simultaneous PCR amplification and detection or quantification of a fluorescently labeled target sequence, and is being used for SNP genotyping (TaqMan®, a hydrolysis probe), and for studies of Copy Number Variation (CNV) and gene expression, aiming at the differential diagnosis, correct evaluation of patient´s allelic composition, and deletion detection [77,78].

For diagnostic purposes, RT-PCR genotyping represents a simple and reliable technology for high-throughput genetic screening. In this scenario, TaqMan® probes are a valuable tool for the screening of common mutations, as the ones describes with founder effects in certain geographic or ethnic populations (Figure 2(b)) [79]. TaqMan® SNP genotyping assays consist of pre-optimized unlabeled PCR primer pairs and two TaqMan probes for allelic discrimination, one with FAM dye label and one with VIC dye label attached to the 5′ and nonfluorescent quenchers (NFQ) on the 3′ end. Each TaqMan probe anneals specifically to a complementary sequence, if present, between the forward and reverse primer sites. When the probe is intact, the proximity of the quencher dye to the reporter dye suppresses the report fluorescence. The exonuclease activity of DNA polymerase cleaves only probes hybridized to the target. Cleavage separates the reporter dye from the quencher dye, increasing fluorescence by the reporter. The increase in fluorescence occurs only if the amplified target sequence is complementary to the probe. Fluorescence detected is directly proportional to the fluorophore released and the amount of DNA template present in the PCR and indicates which alleles are in the sample. In Figure 2(b), RT-PCR results are presented as a multicomponent plot, showing a homozygous patient for c.533A>T (p.His178Leu) in ARSB (MPS VI), a mutation with founder effect associated to population isolation in the Northeast of Brazil.

A recent study, using intragenic polymorphisms to identify some association between any particular mutation and specific haplotype in MPS IVA Brazilian patients, was performed by Real time PCR using the TaqMan® methodology with probes designed and customized for analysis of six different intragenic polymorphic sites. Fourteen different haplotypes were identified and one of them associated with p.Ser341Arg, suggesting a founder effect of this mutation in Brazilian population [41].

5.2.2. Multiplex ligation-dependent probe amplification (MLPA)

MLPA is a high-throughput multiplex PCR technique that is able to detect abnormal copy numbers, such as deletions and duplications, in genomic sequences [80]. Reverse Transcriptase MLPA (RT-MLPA) and Methylation-Specific MLPA (MS-MLPA) are variations of this technique, intended for mRNA and methylation profiling, respectively [81,82]. These methods are based on the hybridization of several probes with different lengths to the target sequence, which are then ligated and amplified by a single primer pair. Subsequently, amplicons are separated by capillary electrophoresis and analyzed by comparison of peek height patterns between unknown and reference samples.

One important limitation of MLPA is the effect of genetic variants on the hybridization of probes. Mismatches with the target sequences can lead to reduced or absent probe signal, falsely indicating a deletion. Therefore, it is necessary to confirm results either using another method or a different set of probes, particularly when the possible deletion includes only one probe.

MLPA has been extensively used for diagnosis of MPS II, once partial deletions or complete IDS gene deletion are estimated to be the cause of disease in 20% of all cases [83,84] (Figure 2(c), and unraveling novel mechanism underlying diagnosis of MPS as the reported somatic mosaic disease-causing variant in *IDUA* (MPS I) [54]. In contrast of PCR followed by sequencing, MLPA can also be used for detecting female carriers in whom deletion is masked by the amplification of the normal X chromosome. As a sensible and available tool, MLPA can be used in prenatal diagnosis and genetic counseling [83].

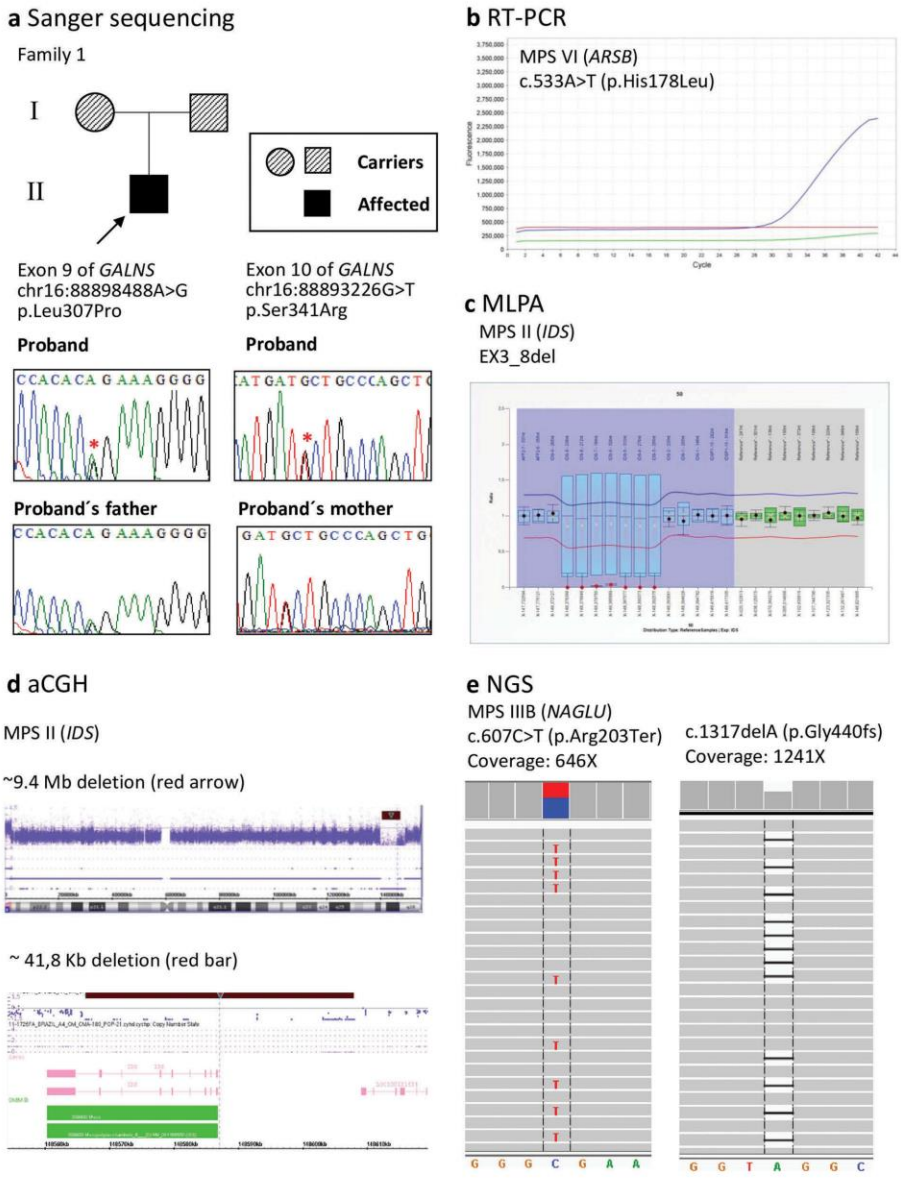

Figure 2. Molecular test using for MPS diagnosis confirmation. (a) Sanger sequencing, showing a compound heterozygous mutations of *GALNS* (MPS IVA) were identified in exon 9 and 10. The proband's parents were asymptomatic heterozygous carriers for the compound mutations; (b) Real-time PCR, *ARSB* homozygous patient for c.533A>T (p.His178Leu) pathogenic variant associated with MPS VI; (c) MLPA; (d) a-CGH. The ideogram of X chromosome shows the location of each probe. Top panel: a ~ 9.4 Mb deletion at Xq27.1-Xq28 (red Arrow) in a MPS II patient. Bottom panel: a high-resolution view of the Xq28 deletion region in MPS II patient. A ~ 41.8 Kb deletion is shown (red bar); (e) Next-generation sequencing, *NAGLU* compound heterozygous for pathogenic variant c.607C>T (p.Arg203Ter) and c.1317del A (p.Gly440fs) with 646X and 1241X, determined by a customized NGS gene panel for MPS III.

phenotypes. In this scenario, a broad gene panel approach could be very helpful, allowing to expand the phenotypic spectrum of these disorders. For example, the involvement of the central nervous system is predominant in MPS I, II, III, and VII, which justified the inclusion of the associated genes in a more neurological panel; while MPS IV, VI, and VII could be included in a skeletal dysplasia panel.

Ambivalent amplification of *IDS* and its pseudogene *IDS-2*, due to its high homology, could result in an ambivalent mapping of the homologous regions reads in the analysis of NGS, complicating to the detection of genuine mutations. Primers/probes may bind to both *IDS* and *IDS-2*, resulting in false positive and making the molecular analysis unreliable. In this scenario, besides some target enrichment strategies [107], bioinformatic data analysis (as bioinformatic filtering) is crucial to overcome the pseudogene presence [74].

WGS is still not part of routine clinical diagnosis and is often used in research context or to answer very specific questions in the diagnosis of complex alterations. WES, on the other hand, is standard diagnostic procedure in many places and, as costs keep decreasing, it may become the primary choice for molecular diagnosis in many situations: (1) Utility in finding novel genes associated with rare conditions, as the newly discovered MPS type (MPSPS) [3,4]; (2) expanding the recognized phenotypic spectrum of well-known disease [108,109]; (3) to elucidate complex phenotype as reported by the group of Kaissi et al, when the genetic confirmation in MPS gene involved in a pair of siblings adds to the obscure nature of the disease [110]; and (4) as first-tier diagnostic tool for MPS, with subsequent traditional biochemistry test (GAG dosage and enzyme assay) to confirm molecular diagnosis, in an inversion of the traditional diagnostic algorithm, which may be a trend for the future as cost of sequencing continues to go down and as fewer laboratories continue to perform sophisticated enzyme assays [111]. Ethical aspects are one of the main challenges of WES due incidental findings, as the identification of pathogenic mutations in genes not related to the main investigation [112].

In the case of MPS, when the specific enzyme deficiency is already identified and the gene to be investigated is known, the decision about the use of Sanger sequencing or NGS depends on the availability and costs involved, which may vary largely from region to region. It should be considered that usually NGS may have gaps in the coverage that require additional Sanger sequencing of selected regions to complete the analysis.

Also, diagnostic reports of variants of unknown significance (especially when previously unreported) should preferably be accompanied of *in silico* analysis of possible effects on splicing or protein activity or localization, including signal peptide alterations. Alternative splicing can be confirmed by cDNA sequencing. Anyway, in the case of MPS, the significance of all molecular genetics findings could be checked by biochemical assays, as the pathogenic variations should have a functional impact to cause a MPS.

Establishing an early and reliable diagnosis is crucial to ensure the appropriate care management and disease-specific treatment when available, such as enzyme replacement therapy (ERT), for MPS type I (laronidase), II (idursulfase), IVA (elosulfase alfa), VI (galsulfase), and VII (vestronidase alfa). Early treatment is essential for a better outcome of these current therapies that have the potential of altering the clinical natural history and contributing to reduce the morbidity of these disorders since no cure exist for any of them [113–115]. In the case of MPS I, hematopoietic stem cell transplantation (HSCT) is indicated in early-diagnosed children in whom a severe (Hurler) phenotype is predicted [116–119], although it can be ineffective in several others (MPS II, III) [64]. Also, new therapies are being developed for specific types of mutations, such as stop-codon read through for MPS I patients. Preclinical studies have demonstrated that stop-codon read-through induced by gentamicin results in enzyme activity that reduced substrate storage. This approach had more effect for some specific mutations, such as p.Trp402Ter and p.Gln70Ter, associated with Hurler phenotype [61].

6. Expert commentary

The MPSs are a group of lysosomal diseases with biochemical defects well characterized and identifiable by specific biochemical methods, which provide a reliable insight into the functional deficiency presented by the patient. The recent identification of a new MPS type (MPSPS) is an exception to this rule. As all MPS-related genes are known, this simplifies the task of identifying the causative mutations, as they could be searched by standard Sanger sequencing. However, the fact that in many instances at least one of the mutations is not found, requires the use of alternative molecular genetics methods to further explore the topography of the genetic material, justifying the need of specialized laboratories to deal with the difficult cases. Although the diagnosis is usually confirmed by biochemical methods (with the exception of the cases of pseudodeficiency and of some cases with borderline biochemical results), to find the causative mutations is important for phenotype prediction, selection of the most appropriate therapy, carrier identification and prenatal diagnosis. The recent introduction of massive parallel sequencing, or next-generation sequencing, allows genotyping with gene panels, which include the MPS genes. This could replace Sanger sequencing when costs are competitive and operational conditions favorable, and can also be helpful when the biochemical investigation is unavailable or incomplete (for example, when specific enzyme assay is not performed or performed only in DBS). Whole exome sequencing and whole genome sequencing do not have a major role in the MPS diagnosis, but they could help to identify patients with unusual phenotypes. When a variant of unknown significance is found, its potential pathogenicity should be tested with specific prediction software, and results correlated with functional biochemical assays. Careful consideration will also need to be given to variant of unknow significance (VUS) identified through NGS. The report of these findings is highly dependent on the laboratory; while some laboratories always report VUS, others will not disclose such information. A depth analysis of this type of variant in the context of phenotype should be taken by bioinformaticians and clinicians in order to decide the most appropriate clinical management in each individual case. A variant will remain as a VUS until a functional evaluation can provide a more definitive assessment of impact or

other evidences can be gathered to reclassify it as pathogenic/benign.

Although not formally required for patient diagnosis in the MPS, identification of the molecular genetics variations is now a standard practice which provides useful information with practical implications, including in many instances phenotype prediction and choice of the most appropriate therapy. As NGS becomes more affordable for families and operationally friendly for health care professionals, an inversion of the traditional investigation flow (clinical -> biochemical -> molecular genetics) may occur, with the genotyping preceding the more laborious, time-consuming and in some cases more expensive biochemical procedures. Anyway, even with the identification of patient's genotype, the biochemical confirmation of the functional defect will remain a need in the MPS field.

7. Conclusion

The growing possibilities of approaches for the diagnosis of MPS combined with the advent of new therapeutic strategies focuses on the importance of defined molecular diagnostic parameters. We summarized here the technical advances in molecular testing procedures and when these new approaches are used, decreasing the percentage of patients without molecular characterization, which is important for phenotype prediction, genetic counseling, detection of carriers and prenatal/preimplantation diagnosis.

8. Five-year view

It is likely that the advent of massive parallel sequencing, or next-generation sequencing, which has undoubtedly and greatly accelerated biological discoveries, will continue to be a major tool for the characterization of the MPS genes. Although NGS technology is becoming more available and affordable, several laboratories will continue to use other molecular genetics approaches for MPS diagnosis, especially when the two pathogenic mutations are not found with the sequencing methods.

It is important to consider the benefits and limitations of each NGS approach when deciding on the best testing strategy. In our opinion, along the next 5 years targeted-panel NGS will continue to be the preferred molecular genetic testing because it provides a reliable clinical application while eliminating unexpected ethical dilemmas. WGS/WES will be important tools in cases of atypical MPS phenotypes and to characterize a 'novel' gene associated to a new type of MPS, as the case of MPSPS.

We are convinced that the interaction between biochemical and molecular approaches is essential for the diagnosis of MPS, since biochemical tests play an important role in directing molecular genetic analysis and driving the use of ancillary methods until the picture is fully defined.

We believe that within the next few years several NBS programs for MPS will kick-off in which enzyme assays by tandem mass spectrometry (MS/MS) will be the first-tier followed by molecular analyses as second-tier. The molecular analyses will enable prognosis prediction and targeted treatment (e.g. read-through for nonsense mutations and chaperone therapy for mutations leading to loss of protein stability). Even though MPS are marked by high allelic heterogeneity, molecular analyses also aid identification of pseudodeficinecy variants and combined with biomaker assessments will be the gold-stantard approach for the omics era.

Key issues

- MPS are a group of LSDs with high phenotypic and genotypic heterogeneity, which makes the precise diagnosis challenging.
- In most cases, biochemical testing is sufficient for MPS diagnosis and for differentiation between subtypes. Molecular testing is mandatory for MPSPS diagnostic and for distinguishing MPS subtypes (e.g MPS I).
- Several molecular methods for rapid detection of variants are useful to reduce the 'diagnostic odyssey,' as well for additional information which could benefit the patient and their family.
- Sequence analysis (by Sanger or NGS) of MPS-associated gene offered a high diagnostic yield, since 80% (*IDS*) to 99% (*GLB1*) are reported to be point mutations and small indels. For the other variant, other molecular analysis approaches are available.
- Molecular genetic testing is important since it provides additional information for supporting treatment choice, predict disease severity in some cases, differentiate pseudodeficiency from real deficiency, prenatal diagnosis, etc.
- Newborn Screening of MPS I is underway in some countries. For other MPS, several pilot studies are underway.
- MPS disorder panel could be the best option if clinical and biochemical findings point toward a particular MPS type/subtype, increasing diagnostic yield. However, targeted gene testing is not always advantageous, mainly in milder or atypical phenotypes.
- WES/WGS are useful tools, especially for milder or atypical MPS phenotypes.

Funding

This article was not funded.

Declaration of interest

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

ORCID

Roberto Giugliani http://orcid.org/0000-0001-9655-3686

References

1. Filocamo M, Morrone A. Lysosomal storage disorders: molecular basis and laboratory testing. Hum Genomics. 2011;5(3):156–169.
2. Kingma SDK, Bodamer OA, Wijburg FA. Epidemiology and diagnosis of lysosomal storage disorders; challenges of screening. Best Pract Res Clin Endocrinol Metab. 2015;29(2):145–157.
3. Kondo H, Maksimova N, Otomo T, et al. Mutation in VPS33A affects metabolism of glycosaminoglycans: a new type of mucopolysaccharidosis with severe systemic symptoms. Hum Mol Genet. 2017;26(1):173–183.
4. Dursun A, Yalnizoglu D, Gerdan OF, et al. A probable new syndrome with the storage disease phenotype caused by the VPS33A gene mutation. Clin Dysmorphol. 2017;26(1):1–12.
5. Applegarth DA, Toone JR, Lowry RB, Incidence of inborn errors of metabolism in British Columbia, 1969-1996. Pediatrics. 2000;105(1):e10. Accessed 2018 Jun 6 http://www.ncbi.nlm.nih.gov/pubmed/10617747.
6. Al-Jasmi FA, Tawfig N, Berniah A, et al. Prevalence and novel mutations of lysosomal storage disorders in United Arab Emirates. In: JIMD reports. Vol 10; 2012:1–9. .
7. Poupětová H, Ledvinová J, Berná L, et al. The birth prevalence of lysosomal storage disorders in the Czech Republic: comparison with data in different populations. J Inherit Metab Dis. 2010;33(4):387–396.
8. Meikle PJ, Hopwood JJ, Clague AE, et al., Prevalence of lysosomal storage disorders. JAMA. 1999;281(3):249–254. Accessed 2018 Jun 11 http://www.ncbi.nlm.nih.gov/pubmed/9918480.
9. Poorthuis BJ, Wevers RA, Kleijer WJ, et al. The frequency of lysosomal storage diseases in The Netherlands. Hum Genet. 105(1–2):151–156. http://www.ncbi.nlm.nih.gov/pubmed/10480370. Accessed 2018 Jun 11
10. Giugliani R, Federhen A, Michelin-Tirelli K, et al. Relative frequency and estimated minimal frequency of lysosomal storage diseases in Brazil: report from a reference laboratory. Genet Mol Biol. 2017;40(1):31–39.
11. Pinto R, Caseiro C, Lemos M, et al. Prevalence of lysosomal storage diseases in Portugal. Eur J Hum Genet. 2004;12(2):87–92.
12. Navarrete-Martínez JI, Limón-Rojas AE, de Jesús Gaytán-García M, et al. Newborn screening for six lysosomal storage disorders in a cohort of Mexican patients: three-year findings from a screening program in a closed Mexican health system. Mol Genet Metab. 2017;121(1):16–21.
13. Neufeld EF, Muenzer J. The mucopolysaccharidoses. In: David Valle, et al. eds. The online metabolic and molecular bases of inherited disease. New York, NY: McGraw-Hill Medical; 2014. https://ommbid.mhmedical.com/content.aspx?bookid=971&sectionid=62642135&jumpsectionID=62642144#1102893981. Accessed 2018 Jun 26
14. Muenzer J. The mucopolysaccharidoses: a heterogeneous group of disorders with variable pediatric presentations. J Pediatr. 2004;144(5):S27–S34.
15. Wilcox WR. Lysosomal storage disorders: the need for better pediatric recognition and comprehensive care. J Pediatr. 2004;144(5):S3–S14.
16. Vieira T, Schwartz I, Muñoz V, et al. Mucopolysaccharidoses in Brazil: what happens from birth to biochemical diagnosis?. Am J Med Genet Part A. 2008;146A(13):1741–1747.
17. Wijburg FA, Węgrzyn G, Burton BK, et al. Mucopolysaccharidosis type III (Sanfilippo syndrome) and misdiagnosis of idiopathic developmental delay, attention deficit/hyperactivity disorder or autism spectrum disorder. Acta Paediatr. 2013;102(5):462–470.
18. Vijay S, Wraith J. Clinical presentation and follow-up of patients with the attenuated phenotype of mucopolysaccharidosis type I. Acta Paediatr. 2005;94(7):872–877.
19. Clarke LA. The mucopolysaccharidoses: a success of molecular medicine. Expert Rev Mol Med. 2008;10:e1.
20. Peña-Gomar IM, Rosas-Trigueros JL, Reyes-López CA. C.1898c>G/p.ser633trp IDUA mutation cause an intermediate-to-severe phenotype in mucopolysaccharidosis type I patient: clinical and computational approach. Mol Genet Metab. 2016;117(2):S92.
21. Gabrielli O, Coppa GV, Bruni S, et al. An adult Sanfilippo type A patient with homozygous mutation R206P in the sulfamidase gene. Am J Med Genet A. 2005;133A(1):85–89.
22. Miyazaki T, Masuda N, Waragai M, et al. An adult Japanese Sanfilippo A patient with novel compound heterozygous S347F and D444G mutations in the sulphamidase gene. J Neurol Neurosurg Psychiatry. 2002;73(6):777–778.
23. Di Natale P, Villani GRD, Di Domenico C, et al., Analysis of Sanfilippo A gene mutations in a large pedigree. Clin Genet. 2003;63(4):314–318. Accessed 2018 Jun 26 http://www.ncbi.nlm.nih.gov/pubmed/12702166.
24. Berger-Plantinga EG, Vanneste JAL, Groener JEM, et al. Adult-onset dementia and retinitis pigmentosa due to mucopolysaccharidosis III-C in two sisters. J Neurol. 2004;251(4):479–481.
25. Ruijter GJG, Valstar MJ, van de Kamp JM, et al. Clinical and genetic spectrum of Sanfilippo type C (MPS IIIC) disease in The Netherlands. Mol Genet Metab. 2008;93(2):104–111.
26. Scott HS, Litjens T, Hopwood JJ, et al. A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype. Hum Mutat. 1992;1(2):103–108.
27. Brusius-Facchin AC, Schwartz IVD, Zimmer C, et al. Mucopolysaccharidosis type II: identification of 30 novel mutations among Latin American patients. Mol Genet Metab. 2014;111(2):133–138.
28. Brusius-Facchin AC, De Souza CFM, Schwartz IVD, et al. Severe phenotype in MPS II patients associated with a large deletion including contiguous genes. Am J Med Genet A. 2012;158A(5):1055–1059.
29. Amartino H, Ceci R, Masllorens F, et al. Identification of 17 novel mutations in 40 Argentinean unrelated families with mucopolysaccharidosis type II (Hunter syndrome). Mol Genet Metab Rep. 2014;1:401–406.
30. Lehman TJA, Miller N, Norquist B, et al. Diagnosis of the mucopolysaccharidoses. Rheumatology. 2011;50(suppl 5):v41–v48.
31. Scarpa M. Mucopolysaccharidosis Type II. Seattle: University of Washington; 1993. [Accessed 2018 Jun 13]. http://www.ncbi.nlm.nih.gov/pubmed/20301451.
32. Lualdi S, Regis S, Di Rocco M, et al. Characterization of iduronate-2-sulfatase gene-pseudogene recombinations in eight patients with mucopolysaccharidosis type II revealed by a rapid PCR-based method. Hum Mutat. 2005;25(5):491–497.
33. Timms KM, Lu F, Shen Y, et al. 30 kb of DNA sequence reveals two new genes and a regional duplication distal to the human Iduronate-2-sulfate sulfatase locus. https://genome.cshlp.org/content/5/1/71.full.pdf. Accessed June 13, 2018.
34. Bondeson M-L, Dahl N, Malmgren H, et al. Inversion of the IDS gene resulting from recombination with IDS-related sequences in a common cause of the Hunter syndrome. Hum Mol Genet. 1995;4(4):615–621.
35. Bondeson ML, Malmgren H, Dahl N, et al., Presence of an IDS-related locus (IDS2) in Xq28 complicates the mutational analysis of Hunter syndrome. Eur J Hum Genet. 1995;3(4):219–227. Accessed 2018 Jun 14 http://www.ncbi.nlm.nih.gov/pubmed/8528670.
36. Valayannopoulos V, Nicely H, Harmatz P, et al. Mucopolysaccharidosis VI. Orphanet J Rare Dis. 2010;5(1):5.
37. Costa-Motta FM, Acosta AX, Abé-Sandes K, et al. Genetic studies in a cluster of mucopolysaccharidosis type vi patients in Northeast Brazil. Mol Genet Metab. 2011;104(4):603–607.
38. Jurecka A, Piotrowska E, Cimbalistiene L, et al. Molecular analysis of mucopolysaccharidosis type VI in Poland, Belarus, Lithuania and Estonia. Mol Genet Metab. 2012;105(2):237–243.
39. Santos S, Pequeno AADS, Galvão CRC, et al. The causes of physical disability in municipalities of the northeast of Brazil and an estimate of costs of specialized services. Cien Saude Colet. 2014;19(2):559–568. Accessed 2018 Jun 26 http://www.ncbi.nlm.nih.gov/pubmed/24863832.

40. Weller M, Tanieri M, Pereira JC, et al. Consanguineous unions and the burden of disability: a population-based study in communities of Northeastern Brazil. Am J Hum Biol. 2012;24(6):835–840.
41. Bochernitsan AN, Brusius-Facchin AC, Couto RR, et al. Spectrum of GALNS mutations and haplotype study in Brazilian patients with mucopolysaccharidosis type IVA. Meta Gene. 2018;16:77–84.
42. Di Natale P, Balzano N, Esposito S, et al. Identification of molecular defects in Italian Sanfilippo A patients including 13 novel mutations. Hum Mutat. 1998;11(4):313–320.
43. Mangas M, Nogueira C, Prata M, et al. Molecular analysis of mucopolysaccharidosis type IIIB in Portugal: evidence of a single origin for a common mutation (R234C) in the Iberian Peninsula. Clin Genet. 2008;73(3):251–256.
44. Beesley CE, Young EP, Vellodi A, et al., Mutational analysis of Sanfilippo syndrome type A (MPS IIIA): identification of 13 novel mutations. J Med Genet. 2000;37(9):704–707. Accessed 2018 Jun 14 http://www.ncbi.nlm.nih.gov/pubmed/11182930.
45. Canals I, Elalaoui S, Pineda M, et al. Molecular analysis of Sanfilippo syndrome type C in Spain: seven novel HGSNAT mutations and characterization of the mutant alleles. Clin Genet. 2011;80(4):367–374.
46. Coutinho M, Lacerda L, Prata M, et al. Molecular characterization of Portuguese patients with mucopolysaccharidosis IIIC: two novel mutations in the HGSNAT gene. Clin Genet. 2008;74(2):194–195.
47. Feldhammer M, Durand S, Mrázová L, et al. Sanfilippo syndrome type C: mutation spectrum in the heparan sulfate acetyl-CoA: α-glucosaminide N-acetyltransferase (*HGSNAT*) gene. Hum Mutat. 2009;30(6):918–925.
48. Fedele AO, Filocamo M, Di Rocco M, et al. Mutational analysis of the HGSNAT gene in Italian patients with mucopolysaccharidosis IIIC (Sanfilippo C syndrome). Hum Mutat. 2007;28(5):523. Mutation in brief #959. Online.
49. Yamada N, Fukuda S, Tomatsu S, et al. Molecular heterogeneity in mucopolysaccharidosis IVA in Australia and Northern Ireland: nine novel mutations including T312S, a common allele that confers a mild phenotype. Hum Mutat. 1998;11(3):202–208.
50. Tomatsu S, Fukuda S, Cooper A, et al. Mucopolysaccharidosis IVA: identification of a common missense mutation I113F in the N-acetylgalactosamine-6-sulfate sulfatase gene. Am J Hum Genet. 1995;57(3):556–563. Accessed 2018 Jun 14 http://www.ncbi.nlm.nih.gov/pubmed/7668283.
51. Kato Z, Fukuda S, Tomatsu S, et al. A novel common missense mutation G301C in the N-acetylgalactosamine-6-sulfate sulfatase gene in mucopolysaccharidosis IVA. Hum Genet. 1997;101(1):97–101.
52. Paschke E, Milos I, Kreimer-Erlacher H, et al. Mutation analyses in 17 patients with deficiency in acid β-galactosidase: three novel point mutations and high correlation of mutation W273L with Morquio disease type B. Hum Genet. 2001;109(2):159–166.
53. Catarzi S, Giunti L, Papadia F, et al. Morquio A syndrome due to maternal uniparental isodisomy of the telomeric end of chromosome 16. Mol Genet Metab. 2012;105(3):438–442.
54. Breen C, Mercer J, Jones SA, et al. Maternal mosaicism for IDUA deletion clarifies recurrence risk in MPS I. Hum Genome Var. 2016;3(1):16031.
55. Piraud M, Boyer S, Mathieu M, et al., Diagnosis of mucopolysaccharidoses in a clinically selected population by urinary glycosaminoglycan analysis: a study of 2,000 urine samples. Clin Chim Acta. 1993;221(1–2):171–181. Accessed 2018 Jun 26 http://www.ncbi.nlm.nih.gov/pubmed/8149634.
56. Gallegos-Arreola MP, Machorro-Lazo MV, Flores-Martínez SE, et al. Urinary glycosaminoglycan excretion in healthy subjects and in patients with mucopolysaccharidoses. Arch Med Res. 31(5):505–510. http://www.ncbi.nlm.nih.gov/pubmed/11179586. Accessed 2018 Jun 26
57. Bravo H, Neto EC, Schulte J, et al. Investigation of newborns with abnormal results in a newborn screening program for four lysosomal storage diseases in Brazil. Mol Genet Metab Rep. 2017;12:92–97.
58. Pasqualim G, Ribeiro MG, Da Fonseca GGG, et al. p.L18P: a novel IDUA mutation that causes a distinct attenuated phenotype in mucopolysaccharidosis type I patients. Clin Genet. 2015;88(4):376–380.
59. Kamei M, Kasperski K, Fuller M, et al. Aminoglycoside-induced premature stop codon read-through of mucopolysaccharidosis type I patient Q70X and W402X mutations in cultured cells. JIMD Rep. 2014;13:139–147.
60. Bartolomeo R, Polishchuk EV, Volpi N, et al. Pharmacological read-through of nonsense ARSB mutations as a potential therapeutic approach for mucopolysaccharidosis VI. J Inherit Metab Dis. 2013;36(2):363–371.
61. Hein LK, Bawden M, Muller VJ, et al. α-l-Iduronidase premature stop codons and potential read-through in mucopolysaccharidosis type I patients. J Mol Biol. 2004;338(3):453–462.
62. Gaber KR, Ibrahim MM, Farag MK, et al. Prenatal genetic testing, counseling and follow-up of 33 Egyptian pregnant females with history of mucopolysaccharidoses. Egypt J Med Hum Genet. 2015;16(2):159–163.
63. Leistner-Segal S, Brusius-Facchin AC, Gus R, et al. Fast and robust protocol for prenatal diagnosis of mucopolysaccharidosis type II. Clin Biomed Res. 2014;34(4). Accessed 2018 Jun 26 http://seer.ufrgs.br/index.php/hcpa/article/view/49872/33150.
64. Schielen PCJI, Kemper EA, Gelb MH. Newborn screening for lysosomal storage diseases: a concise review of the literature on screening methods, therapeutic possibilities and regional programs. Int J Neonatal Screen. 2017;3:2.
65. Clarke LA, Atherton AM, Burton BK, et al. Mucopolysaccharidosis type I newborn screening: best practices for diagnosis and management. J Pediatr. 2017;182:363–370.
66. Vervoort R, Islam MR, Sly W, et al. A pseudodeficiency allele (D152N) of the human beta-glucuronidase gene. Am J Hum Genet. 1995;57(4):798–804. Accessed 2018 Jun 26 http://www.ncbi.nlm.nih.gov/pubmed/7573038.
67. Sandberg S, Deanching M, Hoganson G, et al. Pseudo-deficiency allele of the N-acetylgalactosamine-4-sulfatase gene identified in a family with Maroteaux–lamy syndrome (mucopolysaccharidosis type VI). Mol Genet Metab. 2008;93(2):34.
68. Ouesleti S, Coutinho MF, Ribeiro I, et al. Update of the spectrum of mucopolysaccharidoses type III in Tunisia: identification of three novel mutations and in silico structural analysis of the missense mutations. World J Pediatr. 2017;13(4):374–380.
69. Atçeken N, Özgül RK, Yücel Yilmaz D, et al. Evaluation and identification of IDUA gene mutations in Turkishpatients with mucopolysaccharidosis type I. Turkish J Med Sci. 2016;46(2):404–408.
70. Zhang H, Li J, Zhang X, et al. Analysis of the IDS gene in 38 patients with Hunter syndrome: the c.879G>A (p.Gln293Gln) synonymous variation in a female create exonic splicing. PLoS One. 2011;6(8):e22951.
71. Karageorgos L, Brooks DA, Pollard A, et al. Mutational analysis of 105 mucopolysaccharidosis type VI patients. Hum Mutat. 2007;28(9):897–903.
72. Chiong MAD, Canson DM, Abacan MAR, et al. Clinical, biochemical and molecular characteristics of Filipino patients with mucopolysaccharidosis type II - Hunter syndrome. Orphanet J Rare Dis. 2017;12(1):7.
73. Vrijenhoek T, Kraaijeveld K, Elferink M, et al. Next-generation sequencing-based genome diagnostics across clinical genetics centers: implementation choices and their effects. Eur J Hum Genet. 2015;23(9):1142–1150.
74. Claes Kathleen BM, De Leeneer K Dealing with pseudogenes in molecular diagnostics in the next-generation sequencing era. In: Methods in molecular biology, Vol. 1167. Clifton, NJ; 2014. 303–315.
75. Adam M, Ardinger H, Pagon R, et al. GeneReviews® - PubMed - NCBI. https://www.ncbi.nlm.nih.gov/pubmed/?term=20301295. Accessed 2018 Jun 26
76. Brusius-Facchin AC, Abrahão L, Schwartz IVD, et al. Extension of the molecular analysis to the promoter region of the iduronate 2-sulfatase gene reveals genomic alterations in mucopolysaccharidosis type II patients with normal coding sequence. Gene. 2013;526(2):150–154.

77. Caciotti A, Tonin R, Rigoldi M, et al. Optimizing the molecular diagnosis of GALNS: novel methods to define and characterize Morquio-A syndrome-associated mutations. Hum Mutat. 2015;36(3):357–368.
78. Lualdi S, Di Rocco M, Corsolini F, et al. Identification of nine new IDS alleles in mucopolysaccharidosis II. Quantitative evaluation by real-time RT-PCR of mRNAs sensitive to nonsense-mediated and nonstop decay mechanisms. Biochim Biophys Acta. 2006;1762(4):478–484.
79. Heid CA, Stevens J, Livak KJ, et al. Real time quantitative PCR. Genome Res. 1996;6(10):986–994.
80. Schouten JP. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. 2002;30(12):57e–57.
81. Eldering E, Spek CA, Aberson HL, et al. Expression profiling via novel multiplex assay allows rapid assessment of gene regulation in defined signalling pathways. Nucleic Acids Res. 2003;31(23):e153. Accessed 2018 Jun 26 http://www.ncbi.nlm.nih.gov/pubmed/14627843.
82. Nygren AOH, Ameziane N, Duarte HMB, et al. Methylation-specific MLPA (MS-MLPA): simultaneous detection of CpG methylation and copy number changes of up to 40 sequences. Nucleic Acids Res. 2005;33(14):e128.
83. Galvis J, González J, Uribe A, et al. Deep genotyping of the IDS gene in Colombian patients with Hunter syndrome. JIMD Rep. 2015;19:101–109.
84. Jurecka A, Krumina Z, Žuber Z, et al. Mucopolysaccharidosis type II in females and response to enzyme replacement therapy. Am J Med Genet Part A. 2012;158A(2):450–454.
85. Wong L-JC, Dimmock D, Geraghty MT, et al. Utility of oligonucleotide array-based comparative genomic hybridization for detection of target gene deletions. Clin Chem. 2008;54(7):1141–1148.
86. Tayeh MK, Chin ELH, Miller VR, et al. Targeted comparative genomic hybridization array for the detection of single- and multiexon gene deletions and duplications. Genet Med. 2009;11(4):232–240.
87. Wang J, Zhan H, Li F-Y, et al. Targeted array CGH as a valuable molecular diagnostic approach: experience in the diagnosis of mitochondrial and metabolic disorders. Mol Genet Metab. 2012;106(2):221–230.
88. Zanetti A, Tomanin R, Rampazzo A, et al. A Hunter patient with a severe phenotype reveals two large deletions and two duplications extending 1.2 Mb distally to IDS locus. JIMD Rep. 2014;17:13–21.
89. Champion KJ, Basehore MJ, Wood T, et al. Identification and characterization of a novel homozygous deletion in the α-N-acetylglucosaminidase gene in a patient with Sanfilippo type B syndrome (mucopolysaccharidosis IIIB). Mol Genet Metab. 2010;100(1):51–56.
90. Ittiwut C, Boonbuamas S, Srichomthong C, et al. Novel mutations, including a large deletion in the ARSB gene, causing mucopolysaccharidosis type VI. Genet Test Mol Biomark. 2017;21(1):58–62.
91. Barba M, Czosnek H, Hadidi A. Historical perspective, development and applications of next-generation sequencing in plant virology. Viruses. 2014;6(1):106–136.
92. Grada A, Weinbrecht K. Next-generation sequencing: methodology and application. J Invest Dermatol. 2013;133(8):e11.
93. Roy S, Coldren C, Karunamurthy A, et al. Standards and guidelines for validating next-generation sequencing bioinformatics pipelines A joint recommendation of the association for molecular pathology and the college of American pathologists from the next generation sequencing bioinformatics pipeline validation working group of the clinical practice committee, The Association for Molecular Pathology (AMP). J Mol Diagn. 2018;20(1):4–27.
94. Saito S, Ohno K, Sekijima M, et al. Database of the clinical phenotypes, genotypes and mutant arylsulfatase B structures in mucopolysaccharidosis type VI. J Hum Genet. 2012;57(4):280–282.
95. Mutation Database for N-acetylgalactosamine-6-sulfatase, GALNS. http://galns.mutdb.org/about. Accessed 2018 Jun 14
96. Saito S, Ohno K, Okuyama T, et al. Structural basis of mucopolysaccharidosis type II and construction of a database of mutant iduronate 2-sulfatases. PLoS One. 2016;11(10):e0163964.
97. Sim N-L, Kumar P, Hu J, et al. SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res. 2012;40(W1):W452–W457.
98. Marloes M, Diego C, Markus K, et al. A method and server for predicting damaging missense mutations. Nat Publ Gr. 2010;7(16):3133–3164.
99. Schwarz JM, Cooper DN, Schuelke M, et al. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014;11(4):361–362.
100. Desmet F-O, Hamroun D, Lalande M, et al. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res. 2009;37(9):e67.
101. Hu J, Ng PC. SIFT Indel: predictions for the functional effects of amino acid insertions/deletions in proteins. PLoS One. 2013;8(10):e77940.
102. Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: revisions 2007. Genet Med. 2008;10(4):294–300.
103. Wei X, Jin F, Ye Y, et al. A novel mutation of IDS gene in a Chinese patient with mucopolysaccharidosis II by next-generation sequencing. Clin Chim Acta. 2011;412(23–24):2340–2342.
104. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. Nat Rev Genet. 2013;14(4):295–300.
105. Fernández-Marmiesse A, Morey M, Pineda M, et al. Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. Orphanet J Rare Dis. 2014;9(1):59.
106. Di Fruscio G, Schulz A, De Cegli R, et al. Lysoplex: an efficient toolkit to detect DNA sequence variations in the autophagy-lysosomal pathway. Autophagy. 2015;11(6):928–938.
107. Valencia CA, Rhodenizer D, Bhide S, et al. Assessment of target enrichment platforms using massively parallel sequencing for the mutation detection for congenital muscular dystrophy. J Mol Diagn. 2012;14(3):233–246.
108. Nikkel SM, Huang L, Lachman R, et al. Whole-exome sequencing expands the phenotype of Hunter syndrome. Clin Genet. 2014;86(2):172–176.
109. Zeng Q, Fan Y, Wang L, et al. Molecular defects identified by whole exome sequencing in a child with atypical mucopolysaccharidosis IIIB. J Pediatr Endocrinol Metab. 2017;30(4):463–469.
110. Kaissi AA, Hofstaetter J, Weigel G, et al. The constellation of skeletal deformities in a family with mixed types of mucopolysaccharidoses: case report. Medicine (Baltimore). 2016;95(32):e4561.
111. Selmer KK, Gilfillan GD, Strømme P, et al. A mild form of mucopolysaccharidosis IIIB diagnosed with targeted next-generation sequencing of linked genomic regions. Eur J Hum Genet. 2012;20(1):58–63.
112. Ashton-Prolla P, Goldim JR, e Vairo FP, et al. Genomic analysis in the clinic: benefits and challenges for health care professionals and patients in Brazil. J Community Genet. 2015;6(3):275–283.
113. Muenzer J. Early initiation of enzyme replacement therapy for the mucopolysaccharidoses. Mol Genet Metab. 2014;111(2):63–72.
114. Tajima G, Sakura N, Kosuga M, et al. Effects of idursulfase enzyme replacement therapy for mucopolysaccharidosis type II when started in early infancy: comparison in two siblings. Mol Genet Metab. 2013;108(3):172–177.
115. Aldenhoven M, Wynn RF, Orchard PJ, et al. Long-term outcome of Hurler syndrome patients after hematopoietic cell transplantation: an international multicenter study. Blood. 2015;125(13):2164–2172.
116. Ghosh A, Miller W, Orchard PJ, et al. Enzyme replacement therapy prior to haematopoietic stem cell transplantation in mucopolysaccharidosis type I: 10 year combined experience of 2 centres. Mol Genet Metab. 2016;117(3):373–377.
117. Poe MD, Chagnon SL, Escolar ML. Early treatment is associated with improved cognition in Hurler syndrome. Ann Neurol. 2014;76(5):747–753.
118. de Ru MH, Boelens JJ, Das AM, et al. Enzyme replacement therapy and/or hematopoietic stem cell transplantation at diagnosis in patients with mucopolysaccharidosis type I: results of a European consensus procedure. Orphanet J Rare Dis. 2011;6(1):55.
119. Wraith JE, Beck M, Lane R, et al. Enzyme replacement therapy in patients who have mucopolysaccharidosis I and are younger than 5 years: results of a multinational study of recombinant human-l-iduronidase (Laronidase). Pediatrics. 2007;120(1):e37–e46.

ANEXO IV – Artigo 8: Phenotype-Oriented NGS Panels for Mucopolysaccharidoses: Validation and Potential Use in the Diagnostic Flowchart

---

Artigo aceito para publicação na revista *GENETICS AND MOLECULAR BIOLOGY*, 2018

Manuscript title: Phenotype-Oriented NGS Panels for Mucopolysaccharidoses: Validation and Potential Use in the Diagnostic Flowchart

Short running title: NGS Panels for Mucopolysaccharidoses

Ana Carolina Brusius-Facchin[1,2], Marina Siebert[3,4], Delva Leão[3], Diana Rojas Malaga[1,2], Gabriela Pasqualim[2,3,5], Franciele Trapp[1],Ursula Matte[2,3,4,5,6], Roberto Giugliani[1,2,5,6] and Sandra Leistner-Segal[1,6]

1. Medical Genetics Service, Hospital de Clínicas de Porto Alegre (HCPA), Brazil
2. Post-Graduation Program in Genetics and Molecular Biology (PPGBM) – Universidade Federal do Rio Grande do Sul (UFRGS), Brazil;
3. Molecular and Protein Analysis Unit (UAMP) – Centro de Pesquisa Experimental, HCPA, Brazil
4. Gene Therapy Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil;
5. Genetics Departament – Universidade Federal do Rio Grande do Sul (UFRGS), Brazil.
6. Post-Graduation Program in Medicine: Medical Sciences (PPGCM) – Universidade Federal do Rio Grande do Sul (UFRGS), Brazil;

Correspondence to:

SandraLeistner-Segal, PhD
Medical Genetics Service/HCPA
Ramiro Barcelos, 2350
Porto Alegre, RS, Brazil
90035-903
Tel: +55 51 33597211
e-mail: ssegal@hcpa.edu.br

Abstract

Mucopolysaccharidosis (MPS) are a group of rare genetic disorders caused by deficiency in the activity of specific lysosomal enzymes required for the degradation of glycosaminoglycans (GAGs). The defect in the activity of these enzymes responsible for the degradation of GAGs will result in its abnormal accumulation inside the lysosomes of most cells, inducing progressive cellular damage and multiple organ failure. DNA samples from 70 patients with biochemical diagnosis of different MPSs genotype confirmed by Sanger sequencing were used to evaluate a Next Generation Sequencing (NGS) protocol. Eleven genes related to MPSs were divided into three different panels according to the clinical phenotype. This strategy lead to identification of several pathogenic mutations distributed across all exons of MPSs related genes. We were able to identify 96% of all gene variants previously identified by Sanger sequencing, showing high sensitivity in detecting different types of mutations. Furthermore, new variants were not identified representing 100% specificity by the NGS protocol. The use of this NGS approach for genotype identification in MPSs is an attractive option for diagnosis of MPS patients. In addition, the MPS diagnosis workflow could be divided in a two-tier approach: NGS as a first-tier followed by biochemical confirmation as a second-tier.

Keywords: Lysosomal storage disease; mucopolysaccharidoses; next Generation Sequencing; target Sequence; mutation detection.

Data has not been published elsewhere and all authors have approved the submission of this manuscript.

Introduction

The mucopolysaccharidoses (MPSs) are a group of rare genetic disorders caused by deficiency in the activity of specific lysosomal enzymes required for the degradation of glycosaminoglycans (GAGs). Thus, a defect in the activity of any of the eleven enzymes responsible for the stepwise degradation of GAGs will result in its abnormal accumulation inside the lysosomes of most cells, inducing progressive cellular damage and multiple organ failure, with consequent reduction of the quality of life and life expectancy (Neufeld and Muenzer, 2001).

The MPS disorders are inherited in an autosomal recessive manner and affect males and females equally. The exception is MPS II, an X-linked recessive disorder that primarily affects males (due to autosomal X-chromosomal translocation and non-random X-chromosome inactivation, rare female patients with MPS II have also been reported) (Neufeld and Muenzer, 2001). The clinical manifestations of the MPSs vary considerably, with a wide spectrum of signs and symptoms in multiple organ systems, being chronic and progressive conditions (Muenzer et al, 2011).

MPS I, II and VII are characterized by many similar clinical features, that include skeletal abnormalities (dysostosis multiplex), coarse facies, corneal opacity, hearing loss, decreased pulmonary function, cardiac disease, umbilical and inguinal hernias, visceromegaly, among other problems. The disease onset and rate of progression varies from attenuated to severe manifestations. In addition to the symptomatic manifestations, patients with severe forms of MPS I, II and VII may have cognitive impairment, typically appearing in childhood. MPS VI patients present similar somatic manifestations as MPS I, II and VII, also with a wide spectrum of severity but with absence of cognitive impairment (Muenzer et al,2009, 2011; Martin et al, 2008;Montaño et al, 2016; Valayannopoulos et al, 2005).

Patients with any forms of MPS III present severe central nervous system degeneration with little or no somatic involvement. This disorder may be recognized by a rapid loss of social skills with aggressive behavior and hyperactivity, hirsutism, and coarse facies. The skeletal pathology is relatively mild and often became apparent only after a diagnosis is established ((Neufeld and Muenzer, 2001; Valstar et al, 2010).

Both forms of MPS IV are characterized by skeletal dysplasia, ligamentous laxity, odontoid hypoplasia and short stature, without cognitive impairment. Patients with severe phenotype may live into their second or third decade, and those with attenuated disease may live much longer (Northover et al, 1996; Tomatsu et al, 2011).

MPS IX is the rarest form of MPS, with only 4 patients diagnosed to date, mainly with joint disease, short stature, bifid uvula, submucosal cleft palate, flat nasal bridge, generalized cutaneous swelling, multiple periarticular soft-tissue masses and popliteal cyst (Triggs-Raine et al, 1999).

The purpose of this work was to validate and establish the sensitivity and specificity of NGS panels to identify genetic mutations in MPS patients previously diagnosed and genotyped. Panels were designed to be phenotype-oriented, considering the feasibility of this approach as a first-tier alternative followed by biochemical confirmation as a second-tier.

Materials and Methods

1.1 Patients

This study was approved by the Hospital de Clinicas de Porto Alegre Research Ethics Committee, which is recognized by the Office for Human Research Protections as an Institutional Review Board (IRB0000921).

Samples from 70 MPS patients (8 MPS I, 12 MPS II, 23 MPS III, 17 MPS IV, 6 MPS VI, 4 MPS VII) and 8 controls were included in this study. All patients had a previous biochemical diagnosis and were already genotyped by Sanger sequencing.

1.2 AmpliSeq gene panels

Ion Torrent semiconductor technology is able to load, onto an Ion 314™, 316™ and 318™ chip, 10 Mb, 100 Mb and 1Gb of sequence per run respectively. Targeted NGS can be achieved by Ion AmpliSeq™ technology (Thermo Fisher Scientific); an ultrahigh-multiplex PCR amplification strategy that uses very low input genomic DNA for a simple and fast library construction of specific human genes or genomic regions (Buermans and den Dunnen,2014).

To set up a fast and comprehensive assay for molecular analysis of MPS based in NGS sequencing using the Ion Torrent Personal Genome Machine™, we designed three customized AmpliSeq™ panels for sequencing eleven genes (*IDUA, IDS, SGSH, NAGLU, HGSNAT, GNS, GALNS, GLB1, ARSB, GUSB* and *HYAL1*) related with MPSs. These panels were validated in 78 samples (70 MPS patients and 8 controls) previously sequenced by Sanger method.

The eleven genes related with MPS disease were divided into three different customized panels according to the similarity of clinical presentations. Panel 1, firstly comprised MPS types I, II, VI and VII and was later redesigned to include MPS IX; panel 2 included all MPS III types and panel 3 included the MPS IV types A and B. AmpliSeq™

primer panels were designed by Ion AmpliSeq™ Designer software (Thermo Fisher Scientific) resulting in a 2-pool design for each panel.

1.3 DNA isolation

The genomic DNA (gDNA) was extracted from fresh peripheral blood samples using Salting out method or automated extraction using iPrep Purelink gDNA Blood kit (Invitrogen™). DNA sample quantity and purity of the nucleic acid samples were assessed using Nanodrop 1000 instrument (Thermo Fisher Scientific). All of the gDNAs had 260/280-nm absorbance ratios between 1.8 and 2.0. Patient's DNAs were adjusted to a final concentration of 10 ng/μL.

1.4 Library preparation

The customized Ion AmpliSeq™ panel was processed using the Ion AmpliSeq Library Kit 2.0 (Thermo Fisher Scientific) according to the manufacturer's recommendations, starting from 10 ng of gDNA per pool. The samples were barcoded with the Ion Express Barcode Kit (Thermo Fisher Scientific) to optimize the patient pooling on the same sequencing chip. The number of samples by chip was determined taking into account the total length of each panel, aiming for at least 100X coverage (number of reads per amplicon).

1.5 Ion Torrent PGM Sequencing

The template preparation was performed using an Ion One Touch 2 System (Thermo Fisher Scientific) and an Ion One Touch ES (Thermo Fisher Scientific) following the latest version of the manufacturer's manuals. The template-positive Ion Sphere Particles (ISP+) were sequenced on Ion Torrent Personal Genome Machine (PGM) (Thermo Fisher Scientific) using the 314 Chip v2 following the Ion PGM Sequencing 200 Kit v2 manual. We used 500-flow runs, which support a template read length of approximately 200 bp.

1.6 Bioinformatics analyses

The raw data was processed with the Torrent Suite Software v5.0 (Thermo Fisher Scientific) using the standard pipeline parameters. Read alignment and variant identification were performed with the Torrent Mapping Program (TMAP) v3.4.1 and Torrent Variant Caller (TVC) v5.0 software. We used hg19 (Human Genome version 19, UCSC) as reference to perform alignment with TMAP and a BED (Browser Extensible Display,

UCSC) file to define our regions of interest on TVC (Merriman et al, 2012). The Integrative Genome Viewer (IGV; Broad Institute) was used for the analysis of depth coverage, sequence quality, and variant visualization (Thorvaldsdóttir et al, 2013).

Coverage Analysis plugin from Torrent Suite software was also used to establish four different types of coverage charts. For each sample, the amplicon coverage summary file was downloaded, sorted according to .bed reference file and the total reads information was used to create a unique spreadsheet for all samples of a given panel. Each amplicon depth values were divided by the median depth of the batch, following by calculation of the $\log_{25}$ ratio to access comparisons between sequenced samples. To evaluate the intra-run performance of each amplicon, the mean of $\log_{25}$ value and the standard deviation were calculated. In order to analyze the inter-run reproducibility, we concatenated the intra-run results. We also used the normalized total reads mean values to explore the depth of each sequenced gene.

Results

In this study, we aimed to evaluate the sensitivity and specificity of three custom Ampliseq panels designed for amplifying the coding regions of MPS related genes. These designed panels allowed the analysis of 81 out of 90 exons, targeting 90-95% of the genes (table 1 a;b).

Libraries of 78 samples, previously sequenced by Sanger sequencing, were re-sequenced, by NGS, in twelve separated runs: six for panel 1 and three for panel 2 and 3.

To ensure adequate depth and coverage for variant identification we sequenced 8-16 barcoded samples on 314-chip. For this sample set, we obtained an average of mapped reads of 76.00 0 for panel 1, 58.000 for panel 2 and 47.000 for panel 3, average of reads on target, depth of coverage and uniformity varied from gene to gene (Table 1c).

We were able to identify 250 variants in the 70 samples. Variants were filtered according: a) minimum coverage of 100X; b) variant detection in both strands and c) variant frequency <1%. After filtration, 65 variants remained, including 48 missense/nonsense mutations and 17 pathogenic indels.

Overall, with our panels we were able to identify 90% of all variants previously identified by Sanger method, 96% in panel 1, 100% in panel 2 and 100% in panel 3, showing a high sensitivity in detecting different types of mutations using this approach (Table 2). Furthermore, no other known pathogenic mutations were detected, in addition to those found by Sanger sequencing, showing 100% specificity of our assays.

A partial deletion, present in two MPS II patients could be inferred from coverage graphics. These events are noticed as decreased coverage of consecutives exons, compared with the median observed for samples sequenced in the same batch (Figure 1). Nevertheless, the exact breakpoints could not be determined, and other technologies, such as array-CGH, should be used for this purpose.

The exon I of some genes (*IDUA*, *ARSB*, *NAGLU*, *HGSNAT* and *GALNS)*, and other four exons (exon 10 of *IDUA* gene, exon 7 & 11 of *GUSB*, exon 5 of*)*, failed in the design of amplicons, probably due to the high GC content and/or presence of repeat regions.

IDUA gene had low amplification efficiency of some amplicons, as it can be seen on Figure 2a, which shows the $\log_{25}$ ratio for eight samples sequenced in different runs. Consistently, all of them fall below the median depth of the batch in most of exons, as MPS I is the third most frequent MPS in our routine analysis, we decided to re-design the panel 1, improving the exons coverage.

A run with 8 samples was performed, in order to evaluate the new panel design for *the IDUA* gene. We could identify 12 out of 14 mutations previously detected by Sanger sequencing. The other 2 mutations not detected were in exon I, and no amplicons were designed by Ampliseq^{TM} design software for this region. The median depth of the batch showed a high coverage depth when compared with the first designed panel (Figure 2b).
HYAL1 gene, which is associated with MPS IX, was included in the re-designed panel in order to have a complete coverage of the MPSs genes. This analysis was validated only using 8 normal samples, once this type of MPS is very rare and no positive case was available. The analysis showed a great coverage and depth for all exons. Although the finding of a specific enzyme deficiency in leucocytes or fibroblasts confirms the diagnosis, molecular analysis of the respective gene is recommended, whenever possible (Figure 3).

Discussion

The MPSs are genetic metabolic diseases which could be caused by mutations in several distinct genes, each one coding for a specific enzyme in the GAG degradation pathway. Symptoms may be similar in some MPS types, and nowadays definitive diagnosis is achieved by enzyme assay, usually confirmed by genotype analysis.

Once a clinical suspicion of MPS is raised, biochemical studies are usually performed with diagnostic purposes. The first step in this process aims to identify if the levels of total GAGs in urine are increased and which are the GAGs species excreted in urine. Quantitation of urinary GAGs is usually colorimetric (14), and the qualitative pattern

can be obtained by thin-layer chromatography (TLC), electrophoresis or tandem mass spectrometry (TMS) techniques (Dembure eta al, 1990; Hopwood et al, 1982; Tomatsu et al, 2014; Auray-Blais et al, 2012). The gold standard diagnosis of MPS is based on determination of the enzyme activity in plasma, leukocytes or fibroblasts by fluorometric methods (Vozny et al,2001). Dried blood spots (DBS) could also be used for most enzyme assays, but it is usually recommended to confirm positive results in leukocytes or fibroblasts. These biochemical tests are often laborious and require specialized personnel and specific substrates. Sometimes, it is hard to obtain viable samples, especially when there is the need to travel long distances and across international borders.

Although the biochemical tests (enzyme assays and GAG analysis) are sufficient to confirm the diagnosis of MPS, genetic analysis of the specific genes involved are necessary for the characterization of the molecular defect and prognosis prediction. This enables the detection of carrier status of relatives (that is especially important for female relatives of MPS II patients) and accurate genetic counseling. The knowledge of the causative mutation also allows a faster and more precise prenatal diagnosis in future pregnancies in affected families.

Over the past 30 years, Sanger sequencing technology has been regarded as the gold standard to identify sequence alterations in the target region, being an accurate approach for molecular diagnosis. This method relies on analyzing individual genes, usually exon-by-exon, being expensive and time consuming (Sanger et al, 1977).

In this sense, we developed a custom panel for molecular analysis of MPS patients using NGS, thus allowing the simultaneous sequence of multiple genes grouped according to the MPS phenotype. In our study we chose to group MPS patients according to their clinical symptoms, instead of developing a comprehensive panel for all disease types. This allows maximizing the number of patients per run, avoiding unnecessary sequencing of genes that do not fit the clinical hypothesis.

The use of NGS mutation detection method compared with Sanger sequencing lead to the identification of 250 variants and 90% coverage of the 11 genes involved in MPSs etiology. The high sensitivity of our assay could be demonstrated in the detection of 94% mutations, including missense, nonsense, splice sites, in-frame deletions and large deletions, such as a deletion of four exons in the *IDS* gene.

On the other hand, in five samples the genotype could not be defined. This was due to mutations located in regions not covered or with low amplification (<100X), caused by a high GC content or repeat regions. For these regions which Ion Torrent PGM is unable to

sequence or for regions not designed by Ampliseq design software, other methodologies such as Sanger sequencing must be used for a complete gene amplification.

Moreover, due to the inability of NGS to adequately cover some GC-rich regions, as seen for IDUA gene (MPS I) we re-designed panel 1 changing primers for these regions in order to improve read depth, because there are pathogenic mutations previously reported in these regions. We did not redesign the other panels, which present genes without coverage or with low coverage depth, due to the presence of high GC content in the segments.

Other studies had used NGS to diagnose lysosomal storage disorders by whole-exome approaches (Prada et al, 2014; Fernández-Marmiesse et al, 2014). Selmer et.al, 2012 described a mild form of MPS IIIB and illustrated the diagnostic potential of targeted NGS in Mendelian disease with unknown etiology. Wei et al., 2011, reported a novel disease-causing mutation in the *IDS* gene using this strategy.

NGS is limited by: enormous amounts of genomic data generated after sequencing of large genomic regions making interpretation a big challenge. We were able to minimize this by our targeted sequencing strategy. To assign pathogenicity for previously undetected variants can also be difficult (Barba et al, 2014; Grada and Weinbrecht,2013). In the specific scenario, the existence of validated biochemical tests can help establishing which alterations are disease causing. However, if more than two variants of unknown significance are found in the same gene (or one in the case of X-linked *IDS*) there might remain the doubt of which one (if any) is a non-pathogenic variant.

Additionally, complex rearrangements and large gene deletions could not be detected by Ion Torrent platform requiring the use of other technologies, including aCGH, MLPA and cDNA analysis by Sanger sequencing, to complement the NGS molecular analysis.

In conclusion, the Ion Torrent PGM showed to be a fast tool with a high capacity of sequencing, allowing the mutation detection of multiple genes and patients at the same time improving the molecular diagnosis approach. This massively parallel sequencing technology allows the sequencing of large genomic regions in a short period of time and at a lower cost.

In addition, NGS platforms are becoming more popular and consequently more affordable to smaller centers. Thus, we suggest this NGS protocol as a first-tier for MPS diagnosis followed by a second-tier biochemical confirmation.

Acknowledgements

This study was supported by grants from FAPERGS, CNPq and FIPE-HCPA funds. ACBF receives a governmental postdoctoral scholarship from CAPES. SLS and RG are recipients of CNPq research grants.

References

1. Neufeld EF and Muenzer J (2001) The mucopolysaccharidoses. In: Scriver CR SW, Childs B, Beaudet AL, Valle D, Kinzler KW, Vogelstein B The metabolic and molecular basis of inherited disease. 8th edition.McGraw-Hill, Medical Publishing Division, vol3. p. 3421-52
2. Muenzer J, Beck M, Giugliani R, Suzuki Y, Tylki-Szymanska A, Valayannopoulos V, Vellodi A, Wraith JE. (2011) Idursulfase treatment of Hunter syndrome in children younger than 6 years: results from the Hunter Outcome Survey. Genet Med 13,102-109
3. Martin R, Beck M, Eng C, Giugliani R, Harmatz P, Muñoz V, Muenzer J (2008) Recognition and diagnosis of mucopolysaccharidosis II (Hunter syndrome). Pediatrics 121, 377-86.
4. Muenzer J, Wraith JE, Clarke LA; International Consensus Panel on Management and Treatment of Mucopolysaccharidosis I. (2009) Mucopolysaccharidosis I: management and treatment guidelines. Pediatrics 123, 19-29
5. Montaño AM, Lock-Hock N, Steiner RD, Graham BH, Szlago M, Greenstein R, Pineda M, Gonzalez-Meneses A, Çoker M, Bartholomew D, et al. (2016) Clinical course of sly syndrome (mucopolysaccharidosis type VII). J Med Genet. 53(6):403-18

6. Valayannopoulos V, Nicely H, Harmatz P, et al. (2005) Mucopolysaccharidosis VI. Orphanet J Rare Dis 5:5.

7. Valstar MJ, Neijs S, Bruggenwirth HT, Olmer R, Ruijter GJ, Wevers RA, van Diggelen OP, Poorthuis BJ, Halley DJ, Wijburg FA. (2010) Mucopolysaccharidosis type IIIA: clinical spectrum and genotype-phenotype correlations. Ann Neurol 68, 876-87.

8. Northover H, Cowie RA, Wraith JE (1996) Mucopolysaccharidosis type IVA (Morquio syndrome): a clinical review. J Inherit Metab Dis 19, 357-65.

9. Tomatsu S., Montaño A.M., Oikawa H., Smith M., Barrera L., Chinen Y., Thacker M.M., Mackenzie W.G., Suzuki Y., Orii T. (2011) Mucopolysaccharidosis type IVA (Morquio A disease): clinical review and current treatment. Curr Pharm Biotechnol., 12, 931-945.

10. Triggs-Raine B1, Salo TJ, Zhang H, Wicklow BA, Natowicz MR (1999) Mutations in HYAL1, a member of a tandemly distributed multigene family encoding disparate hyaluronidase activities, cause a newly described lysosomal disorder, mucopolysaccharidosis IX.Proc Natl Acad Sci USA 96, 6296-630.

11. Buermans HP, den Dunnen JT (2014) Next generation sequencing technology: Advances and applications. Biochim Biophys Acta,1842,1932-1941.

12. Merriman B, Ion Torrent R&D Team, Rothberg JM (2012) Progress in Ion Torrent semicondutor chip based sequencing. Electrophoresis, 33, 3397-417.

13. Thorvaldsdóttir H, Robinson JT, Mesirov JP (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings in Bioinformatics 14, 178-192.

14. Dembure PP, Drumheller JE, Barr SM, Elsas LJ (1990) Selective urinary screening for mucopolysaccharidoses. Clin Biochem 23(1):91-6

15. Hopwood JJ, Harrison JR. (1982) High-resolution electrophoresis of urinary glycosaminoglycans: an improved screening test for the mucopolysaccharidoses. Anal Biochem 119(1):120-7

16. Tomatsu S, Shimada T, Mason RW, Kelly J, LaMarr WA, Yasuda E, Shibata Y, Futatsumori H, Montaño AM, Yamaguchi S, Suzuki Y, (2014) Assay for Glycosaminoglycans by Tandem Mass Spectrometry and its Applications. J Anal Bioanal Tech (Suppl 2):006

17. Auray-Blais C, Lavoie P, Zhang H, Gagnon R, Clarke JT, Maranda B, Young SP, An Y, Millington DS. (2012) An improved method for glycosaminoglycan

analysis by LC-MS/ MS of urine samples collected on filter paper. Clin Chim Acta 413(7-8):771-8

18. Voznyi Y.V., Keulemans J.L., van Diggelen O.P.(2001) A fluorimetric enzyme assay for the diagnosis of MPS II (Hunter disease) J Inherit Metab Dis.,24, 675-680.

19. Sanger F, Nicklen S, Coulson AR.(1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. 74(12), 5463–5467.

20. Prada CE, Gonzaga-Jauregui C, Tannenbaum R, Penney S, Lupski JR, Hopkin RJ, Sutton VR (2014) Clinical utility of whole-exome sequencing in rare diseases: Galactosialidosis. Eur. J. Med. Genet. 57(7), 339–344.

21. Fernández-Marmiesse A1, Morey M, Pineda M, Eiris J, Couce ML, Castro-Gago M, Fraga JM, Lacerda L, Gouveia S, Pérez-Poyato MS (2014) Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. Orphanet J. Rare Dis. 9, 59.

22. Wei X1, Jin F, Ye Y, Xu C, Qu N, Ju X, Yi X (2011) A novel mutation of IDS gene in a Chinese patient with mucopolysaccharidosis II by next-generation sequencing. Clin Chim Acta 412, 2340-2

23. Selmer KK1, Gilfillan GD, Strømme P, Lyle R, Hughes T, Hjorthaug HS, Brandal K, Nakken S, Misceo D, Egeland T, Munthe LA, Braekken SK, Undlien DE (2012) A mild form of Mucopolysaccharidosis IIIB diagnosed with targeted next-generation sequencing of linked genomic regions. Eur J Hum Genet 20(1):58-63

24. Barba M, Czosnek H, Hadidi A (2014) Historical perspective, development and applications of next-generation sequencing in plant virology. Viruses 6, 106–36

25. Grada A, Weinbrecht K (2013) Next-Generation Sequencing: Methodology and Application. J. Invest. Dermatol. 133(8)

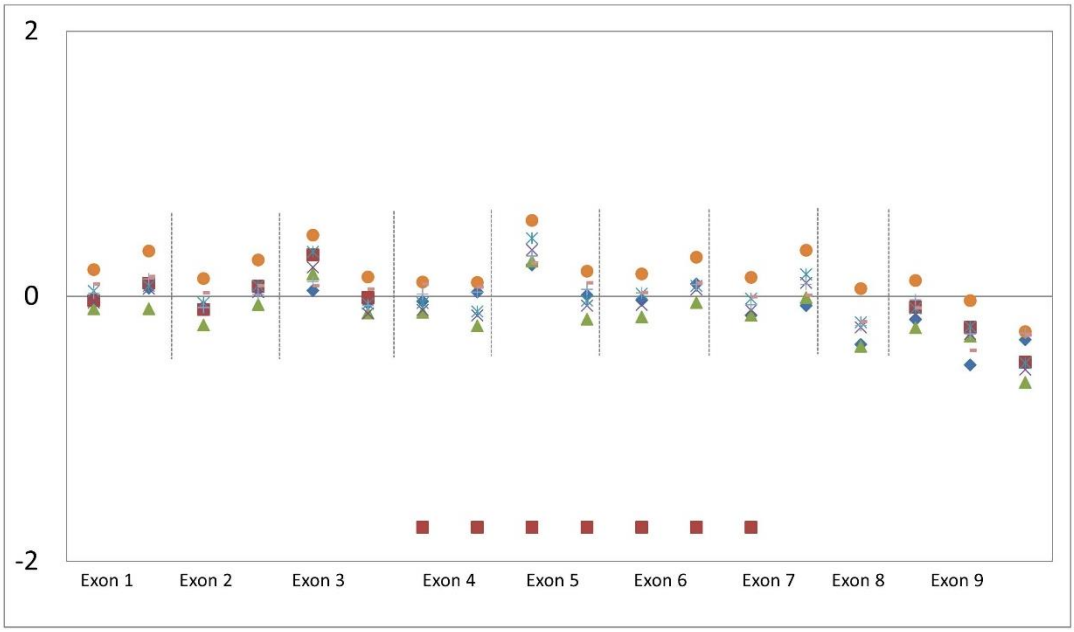

Figure 1: Deletion visualized in the decrease coverage of log25 values. Relative copy number at relative copy number derived from coverage data per target. The red squares mark the probable deletion of the exons IV to VIII of *IDS* gene.

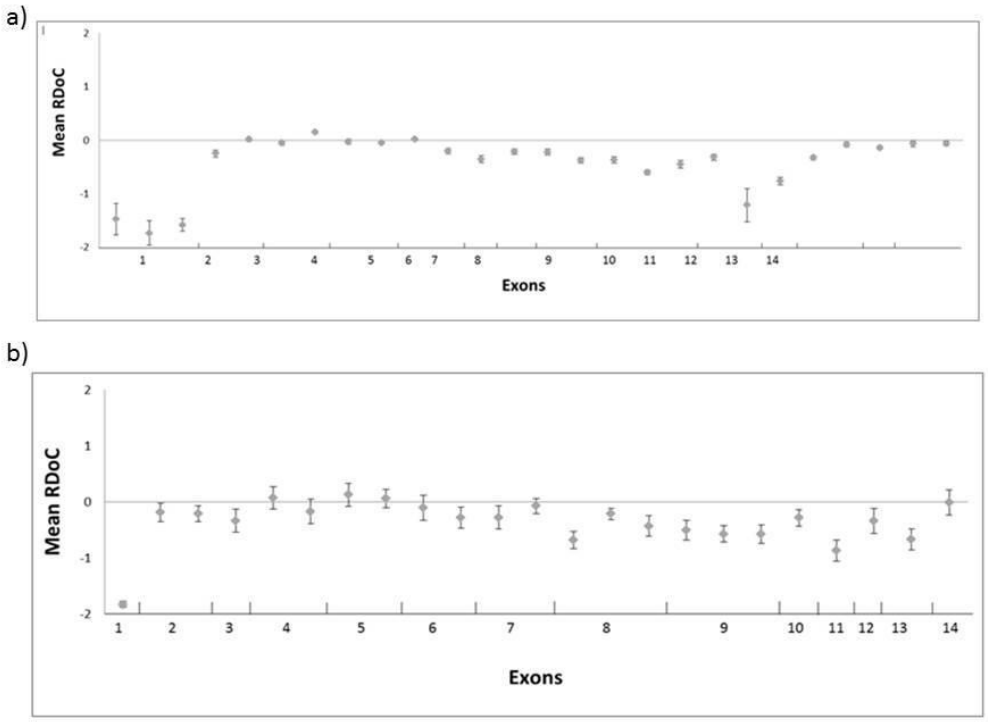

Figure 2: a) *IDUA* gene amplicon coverage. Intra-run relative depth of coverage (RDoC) at 14 exons corresponding to 8 samples sequenced in 6 runs; **b)** *IDUA* gene amplicon coverage of second design. Intra-run relative depth of coverage (RDoC) at 14 exons corresponding to 8 samples sequenced in 1 run.

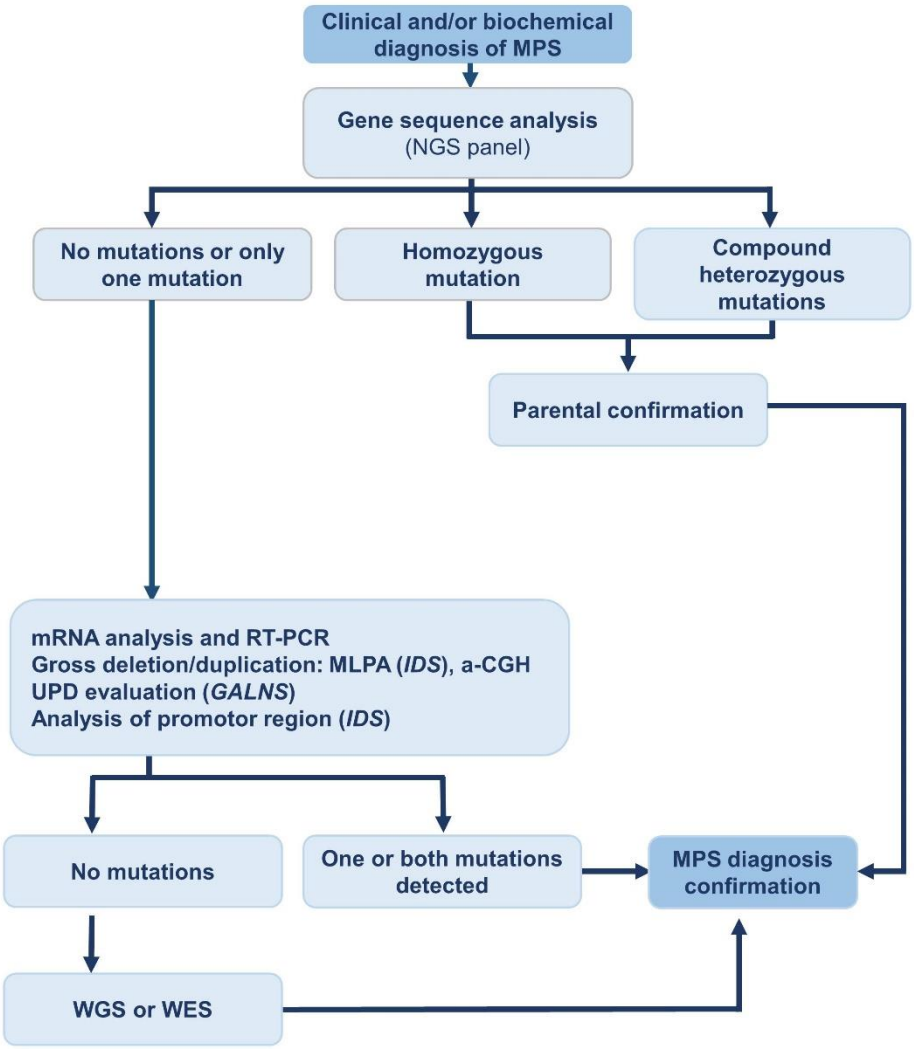

Figure 3. Flowchart of molecular investigation for MPS. MPS, mucopolysaccharidosis; MLPA, multiplex ligation-dependent probe amplification; NGS, next-generation sequencing; mRNA, RNA messenger; RT-PCR, real-time PCR; a-CGH, microarray-based comparative genomic hybridization; UPD, uniparental disomy; WES, whole exome sequencing; WGS, whole genome sequencing.

Table 1: Panels coverage metrics

a)

	Size	Number of amplicons	Missed bp	Coverage (%)
Panel 1	13.42Kb	83	342	95.9
Panel 2	7.81Kb	66	435	95.1
Panel 3	8.61Kb	49	562	90.1

b)

Gene	Target size (bp)	Missed (bp)	Covered (%)
IDUA	2.256	209	90.74
IDS	2.079	0	100
ARSB	1.820	96	94.73
GUSB	2.208	37	98.32
SGSH	1.597	0	100
NAGLU	2.298	294	87.21
HGSNAT	2.106	141	93.3
GNS	1.813	0	100
GALNS	2.657	541	79.64
GLB1	3.049	21	99.31

c)

		Number of samples	Mapped reads	Reads on target	Depth of coverage	Uniformity of coverage
Panel 1						
	IDUA	8	77.000	79%	1160	88%
	IDS	12	76.000	81%	641	87%
	ARSB	6	75.000	82%	841	86%
	GUSB	4	76.000	82%	580	86%
Panel 2						

	SGSH	11	56.000	94%	693	90%
	NAGLU	8	61.000	94%	764	89%
	HGSNAT	2	58.000	76%	477	91%
	GNS	2	-	-	-	-
Panel 3						
	GALNS	13	42.000	67%	565	87%
	GLB1	4	53.000	83%	800	87%

a)Panels metrics; b) Genes target coverage and missed regions; c) Panel coverage metrics.

Table 2: Number of variants identified by Sanger and NGS

	Gene	Pathogenic point mutations		Pathogenic indels		Other* variants
Panel 1		Sanger	NGS	Sanger	NGS	
	IDUA	6	3	0	0	44
	IDS	6	6	6	6	10
	ARSB	4	4	2	2	25
	GUSB	4	4	1	1	18
Panel 2						
	SGSH	6	6	2	2	21
	NAGLU	9	9	3	3	19
	GNS	0	0	0	0	14
	HGSNAT	2	2	0	0	11
Painel 3	*GALNS*	12	12	1	1	15
	GLB1	2	2	2	2	8

* Non- pathogenic mutations: synonymous or polymorphisms

ANEXO V – Artigo 9: Mutation Analysis of Peruvian Patients Diagnosed with Duchenne Muscular Dystrophy

---

Artigo pronto para submissão à revista , 2018

MUTATION ANALYSIS OF PERUVIAN PATIENTS DIAGNOSED WITH DUCHENNE/BECKER MUSCULAR DISTROPHY

María Luisa Guevara-Fujita[1], Francia del Pilar Huaman-Dianderas[1], Alejandro Estrada-Cuzcano[2], Diana Málaga[3], Ricardo Fujita[1]

1 Centro de Investigación de Genética y Biología Molecular, Instituto de Investigación, Facultad de Medicina Humana, Universidad de San Martín de Porres, Lima, Lima – Perú.

2 Medical Genetics Laboratory, INSERM U1112, Institute of Medical Genetics of Alsace, University of Strasbourg, Strasbourg Medical School, 67000 Strasbourg, France.

3 Programa de Doutorado em Genética e Biologia Molecular, Laboratório de Genética y Biologia Molecular, Servico de Genética Médica, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil.

Short running title: Mutations in DMD/DMB patients

Key words: Dystrophin, MLPA, NGS-target, molecular diagnosis

Corresponding Author:

María Luisa Guevara-Fujita

Av. Alameda del corregidor #1531, Urb. Las Viñas, La Molina, Lima, Lima-Perú

mguevarag@usmp.pe

ABSTRACT

The study was performed in a group of 129 patients with clinical findings of Duchenne Muscular Dystrophy, ages ranging from 2-20 years old. Patients came from several hospitals of Lima-Perú. It is appropriate to note that Peruvians have about 80% of native American ancestry, therefore, studies to depict mutations for this and other diseases may give results that are not in agreement with data from other populations worldwide. Genetic tests were performed using the MLPA technique and NGS for a group of MLPA negative patients. Results obtained with the MLPA technique showed 51 patients with deletions/duplications on the DMD gene. Sixty-five patients had MLPA negative results, and in a group of 26 patients, NGS assays were performed. Our data, although small, is showing duplications/deletions in 40% of cases. Also, a high percentage of point mutations distributed as stop codons nucleotide deletions and splice site mutations were found after NGS analyses. Some of the point mutations are not listed in databases consulted. Our aim is to show results of our population, that could also benefit from treatments currently used. To our knowledge, this is the first study of molecular rearrangements performed for DMD in our country.

APÉNDICE I

HCPA - HOSPITAL DE CLÍNICAS DE PORTO ALEGRE
GRUPO DE PESQUISA E PÓS-GRADUAÇÃO

COMISSÃO CIENTÍFICA

A Comissão Científica do Hospital de Clínicas de Porto Alegre analisou o projeto:

Projeto: 150165
Data da Versão do Projeto: 30/06/2015

Pesquisadores:
ROBERTO GIUGLIANI
DIANA ELIZABETH ROJAS MÁLAGA
URSULA DA SILVEIRA MATTE
ANA CAROLINA BRUSIUS-FACCHIN

Título: Validação de uma Nova Abordagem para o Diagnóstico de Doenças Lisossômicas Selecionadas Utilizando o Sequenciamento de Nova Geração

Este projeto foi APROVADO em seus aspectos éticos, metodológicos, logísticos e financeiros para ser realizado no Hospital de Clínicas de Porto Alegre.
Esta aprovação está baseada nos pareceres dos respectivos Comitês de Ética e do Serviço de Gestão em Pesquisa.

- Os pesquisadores vinculados ao projeto não participaram de qualquer etapa do processo de avalição de seus projetos.

- O pesquisador deverá apresentar relatórios semestrais de acompanhamento e relatório final ao Grupo de Pesquisa e Pós-Graduação (GPPG)

Porto Alegre, 01 de julho de 2015.

Prof. José Roberto Goldim
Coordenador CEP/HCPA

APÉNDICE II

HCPA - HOSPITAL DE CLÍNICAS DE PORTO ALEGRE
GRUPO DE PESQUISA E PÓS-GRADUAÇÃO

COMISSÃO CIENTÍFICA

A Comissão Científica do Hospital de Clínicas de Porto Alegre analisou o projeto:

Projeto: 160435
Data da Versão do Projeto: 23/08/2016

Pesquisadores:
ROBERTO GIUGLIANI
DIANA ELIZABETH ROJAS MÁLAGA
URSULA DA SILVEIRA MATTE
ANA CAROLINA BRUSIUS-FACCHIN

Título: Desenho e Validação de um painel de Sequenciamento de Nova Geração para o Diagnóstico Molecular de Lipofuscinose Ceróide Neuronal

Este projeto foi APROVADO em seus aspectos éticos, metodológicos, logísticos e financeiros para ser realizado no Hospital de Clínicas de Porto Alegre.
Esta aprovação está baseada nos pareceres dos respectivos Comitês de Ética e do Serviço de Gestão em Pesquisa.

- Os pesquisadores vinculados ao projeto não participaram de qualquer etapa do processo de avalição de seus projetos.

- O pesquisador deverá apresentar relatórios semestrais de acompanhamento e relatório final ao Grupo de Pesquisa e Pós-Graduação (GPPG)

Porto Alegre, 27 de outubro de 2016.

Prof. José Roberto Goldim
Coordenador CEP/HCPA

APÉNDICE III

HOSPITAL DE CLÍNICAS DE PORTO ALEGRE

Grupo de Pesquisa e Pós Graduação

Carta de Aprovação

Projeto

2018/0249

Pesquisadores:

ANA CAROLINA BRUSIUS-FACCHIN

DIANA ELIZABETH ROJAS MALAGA;

Número de Participantes: 0

Título: Avaliação de três métodos de extração de DNA a partir de sangue impregnado em papel filtro para uso no diagnóstico molecular de doenças lisossômicas

Este projeto foi APROVADO em seus aspectos éticos, metodológicos, logísticos e financeiros para ser realizado no Hospital de Clínicas de Porto Alegre.
Esta aprovação está baseada nos pareceres dos respectivos Comitês de Ética e do Serviço de Gestão em Pesquisa.

- Os pesquisadores vinculados ao projeto não participaram de qualquer etapa do processo de avalição de seus projetos.

- O pesquisador deverá apresentar relatórios semestrais de acompanhamento e relatório final ao Grupo de Pesquisa e Pós-Graduação (GPPG).

Profª Patrícia Ashton Prolla
Coordenadora GPPG/HCPA

Impresso do sistema AGHU-Pesquisa por RAFAEL LEAL ZIMMER em 16/07/2018 16:35:36