

Universidade Federal do Rio Grande do Sul

Análise molecular do gene *IDUA* em pacientes brasileiros com Mucopolissacaridose tipo I e distribuição mundial das mutações mais comuns

Édina Poletto

Dissertação submetida ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS como requisito parcial para a obtenção do grau de Mestre em Genética e Biologia Molecular

Orientador: Prof. Dr. Guilherme Baldo
Co-orientadora: Prof^ª. Dra. Ursula da Silveira Matte

Porto Alegre, março de 2017

Este trabalho foi realizado no Hospital de Clínicas de Porto Alegre (HCPA), no Centro de Terapia Gênica do Centro de Pesquisa Experimental (CTG-CPE) e no Serviço de Genética Médica (SGM). O estudo foi financiado pelo Fundo de Incentivo à Pesquisa e Eventos do HCPA (FIPE-HCPA), pela Fundação Médica do Rio Grande do Sul (FMRS) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Página de agradecimentos

Agradeço à todas pessoas que me auxiliaram na conclusão deste trabalho, em especial:

Meus queridos orientadores, Prof. Guilherme Baldo e Prof^a. Ursula Matte, pelos ensinamentos, oportunidades e, principalmente, pela presença, participação e pelo apoio que recebi nestes anos (e nos que virão);

Gabriela Pasqualim, colega querida que tem toda a minha gratidão, me ajudou do início ao fim deste trabalho (e sem a qual eu certamente não teria conseguido concluir);

Colegas do Centro de Terapia Gênica, pela paciência, pelo companheirismo de sempre e pelos momentos de descontração, que sempre tornam todas as atividades (inclusive as adversidades) mais fáceis;

Minha família, pela motivação constante, interesse, participação e presença em todas as etapas da minha vida.

Sumário

INTRODUÇÃO	1
Mucopolissacaridoses	1
Mucopolissacaridose tipo I.....	3
Tratamento	6
Diagnóstico.....	8
Aspectos moleculares.....	9
OBJETIVOS	12
Objetivo geral.....	12
Objetivos específicos.....	12
RESULTADOS	13
Molecular analysis of MPS I patients in Brazil and geographic distribution of common IDUA mutations worldwide	14
DISCUSSÃO	42
CONCLUSÕES	51
REFERÊNCIAS BIBLIOGRÁFICAS	53

Lista de abreviaturas

MPS – Mucopolissacaridoses

MPS I – Mucopolissacaridose tipo I

GAGs – Glicosaminoglicanos

IDUA – α -L-iduronidase

TRE – Terapia de Reposição Enzimática

TCTH - Transplante de Células-tronco Hematopoiéticas

M6P - Manose-6-fosfato

CRISPR-Cas9 – *Clustered Regularly Interspaced Short Palindromic Repeats* –
CRISPR associated protein 9

mRNA – RNA mensageiro

gnomAD - *Genome Aggregation Database*

Lista de Figuras

Figura 1. Organização dos glicosaminoglicanos na matriz extracelular..	2
Figura 2. Via de degradação dos GAGs sulfato de heparana e sulfato de dermatana	5
Figura 3. Distribuição de p.Trp402* em pacientes com MPS I na Europa.....	45
Figura 4. Distribuição de p.Pro533Arg em pacientes com MPS I na Europa	46
Figura 5. Distribuição de p.Gln70* em pacientes com MPS I na Europa	48

Lista de Tabelas

Tabela 1. Principais características das mucopolissacaridoses.	3
Tabela 2. Frequências no Brasil das mutações mais frequentes no mundo	44

Resumo

A mucopolissacaridose tipo I (MPS I) é uma doença autossômica recessiva causada por mutações no gene α -L-iduronidase (*IDUA*), resultando na ausência da enzima lisossomal de mesmo nome e consequente acúmulo de glicosaminoglicanos (GAGs). Há mais de 200 mutações patogênicas descritas nesse gene e o perfil mutacional varia consideravelmente entre os países. Neste estudo, foi feita a análise molecular de 92 amostras de DNA, representando a maior amostra analisada de pacientes brasileiros com MPS I. Além disso, as frequências alélicas de outros países foram analisadas para melhor entender a composição da nossa população e a distribuição no mundo das mutações mais comuns. Trinta e nove mutações foram identificadas na amostra, sendo 12 delas novas: c.78delC, c.159-1_-23del, p.Gln125*, p.Trp175*, c.590-6ins4G, c.763delC, c.973-1G>A, p.Asp349Glu, p.Asn350Lys, p.Lys384Asn, c.1403-12_-4delCAGGCCCCCG e c.1522_1525delGAGG. Análises *in silico* das mutações novas sugerem que elas são provavelmente patogênicas, confirmando o diagnóstico bioquímico e clínico. As mutações p.Trp402* e p.Pro533Arg são as mais frequentes no Brasil, com 38% e 17% dos alelos, respectivamente. A primeira também é a mais frequente na Austrália, América do Norte e em alguns países da Europa, enquanto que a p.Pro533Arg é encontrada principalmente no norte da África e nos países mediterrâneos. O segundo alelo mais frequente no mundo, p.Gln70*, é muito comum na Escandinávia e relativamente raro nos pacientes brasileiros, com apenas 3% dos alelos. Estes resultados revelam a grande heterogeneidade alélica do gene *IDUA* não apenas nos pacientes brasileiros, mas também entre populações geograficamente próximas, e contribui para o desenvolvimento de estratégias de diagnóstico e tratamento mais eficientes.

Palavras-chave: Mucopolissacaridose tipo I, *IDUA*, perfil mutacional, p.Trp402*, p.Pro533Arg.

Abstract

Mucopolysaccharidosis type I (MPS I) is an autosomal recessive disease caused by mutations in the α -L-iduronidase (*IDUA*) gene, resulting in lack of the lysosomal enzyme IDUA and consequent accumulation of glycosaminoglycans (GAGs). There are more than 200 disease-causing mutations reported and the mutational profile varies considerably from country to country. In this study, we conducted molecular analysis of 92 DNA samples, representing the largest cohort of Brazilian MPS I patients examined. Additionally, we analysed allelic frequencies from other countries to better understand the composition of our population and the allele distribution worldwide. Forty-two mutations were identified in our sample and 14 of them are newly described: c.78delC, c.159-1_-23del, p.Gln125*, p.Trp175*, c.590-6ins4G, c.763delC, c.973-1G>A, p.Asp349Glu, p.Asn350Lys, p.Lys384Asn, c.1403-12_-4delCAGGCCCCG and c.1522_1525delGAGG. *In silico* analysis of novel mutations suggests they are possibly pathogenic, supporting previous biochemical and clinical diagnosis. Among recurrent mutations, p.Trp402* and p.Pro533Arg are the most frequent in Brazil, with 38% and 17% of alleles, respectively. The first is also the major allele in Australian patients and in some European and American countries, while the second is found in the Mediterranean region, mainly North Africa. The second most frequent allele worldwide and highly prevalent in Scandinavia, p.Gln70*, was relatively rare in our sample, with only 3% of alleles. These results reveal the great allelic heterogeneity of *IDUA* not only in Brazilian patients, which differs from all countries analysed, but also between geographically close populations. The knowledge of the genetic background of MPS I may improve diagnosis and management of patients and families with the disease.

Keywords: Mucopolysaccharidosis type I, IDUA, mutation profile, p.Trp402*, p.Pro533Arg

INTRODUÇÃO

Mucopolissacaridoses

As mucopolissacaridoses (MPS) são um grupo de doenças lisossômicas de depósito, hereditárias, causadas pela deficiência de enzimas lisossômicas responsáveis pela degradação de glicosaminoglicanos (GAGs) (Neufeld and Muenzer, 2001).

Com exceção ao hialuronan (ou ácido hialurônico), os GAGs são cadeias polissacarídicas, covalentemente ligadas a um núcleo proteico, que formam os proteoglicanos (Figura 1) – componentes muito importantes da matriz extracelular, que conferem suporte contra forças de compressão (Alberts, et al., 2010). O *turnover* dos GAGs é contínuo, e o seu catabolismo é iniciado pela quebra da ligação proteica com o núcleo proteico e, posteriormente no lisossomo, é feita a remoção de grupos sulfato e a quebra das ligações entre os dissacarídeos, envolvendo 10 enzimas lisossômicas diferentes em ação sequencial (Coutinho, et al., 2012; Neufeld and Muenzer, 2001). A deficiência de qualquer uma dessas enzimas resulta em interrupção na degradação dos GAGs sulfato de condroitina, sulfato de dermatana, sulfato de heparana e sulfato de queratana, individualmente ou em conjunto, resultando nas mucopolissacaridoses I a VII, dependendo da enzima ausente e do substrato acumulado. As MPS descritas, inclusive a mais recente, MPS IX – que envolve o acúmulo de hialuronan – estão detalhadas na Tabela 1.

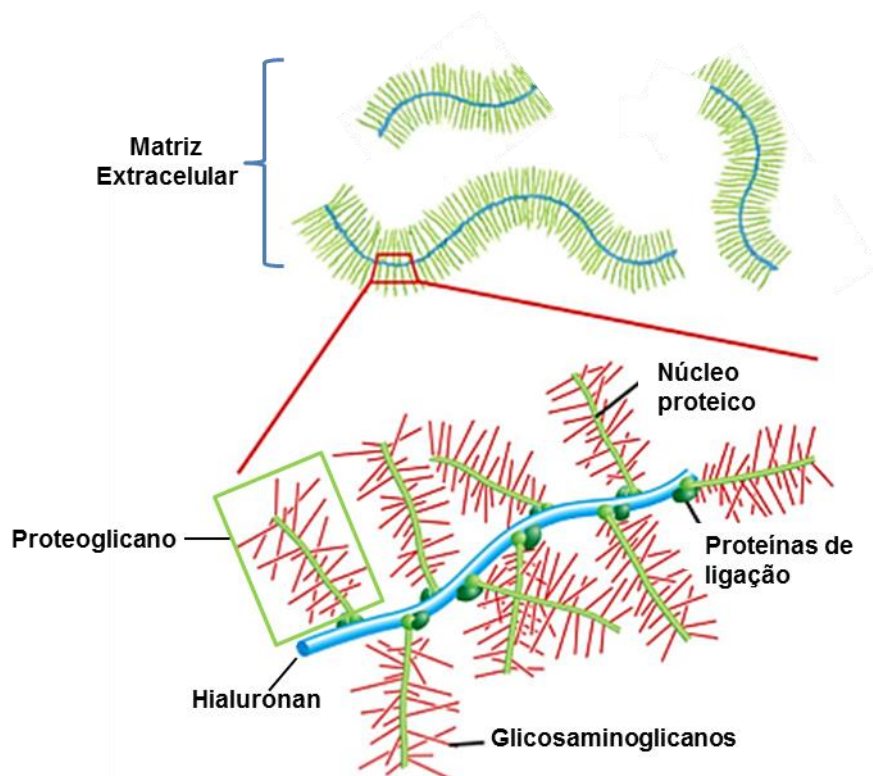


Figura 1. Organização dos glicosaminoglicanos na matriz extracelular. Os glicosaminoglicanos (GAGs) sulfato de condroitina, sulfato de dermatana, sulfato de heparana e sulfato de queratana estão ligados a núcleos proteicos formando os proteoglicanos. Estes, por sua vez, se ligam à uma molécula de hialuronan por proteínas de ligação, formando grandes proteoglicanos, como agrecana, versicana e neurocana, dependendo do tecido onde se encontram. Adaptado de (Alberts, et al., 2010).

Os GAGs não degradados ou parcialmente degradados se acumulam nos lisossomos, eventualmente causando diversas disfunções celulares e teciduais, como aumento de estruturas, aumento de metabólitos secundários, inflamação e ativação do sistema imune, alteração de sinalização celular e superexpressão de proteases, por exemplo (Archer, et al., 2014; Bigg, et al., 2013; Coutinho, et al., 2012; Neufeld and Muenzer, 2001). GAGs excedentes também são excretados na urina, sendo um auxiliar diagnóstico capaz de diferenciar as MPS nos grandes grupos (Neufeld and Muenzer, 2001).

Tabela 1. Principais características das mucopolissacaridoses.

Tipo	Síndrome	Enzima deficiente	GAGs acumulados
MPS I	Hurler e Scheie	α -L-Iduronidase	HS, DS
MPS II	Hunter	Iduronato sulfatase	HS, DS
MPS IIIA	Sanfilipo A	Heparan N-sulfatase	HS
MPS IIIB	Sanfilipo B	α -N-acetil-glicosaminidase	HS
MPS IIIC	Sanfilipo C	Acetil-Coa: α -glicosaminideo acetiltransferase	HS
MPS IIID	Sanfilipo D	N-acetilglicosamina 6-sulfatase	HS
MPS IVA	Morquio A	Galactose-6-sulfatase	KS, CS
MPS IVB	Morquio B	β -galactosidase	KS, CS
MPS VI	Maroteaux-Lamy	Arilsulfatase B	DS, CS
MPS VII	Sly	β -glicuronidase	HS, DS, CS
MPS IX	Natowicz	Hialuronidase	Hialuronan

HS: Sulfato de heparana; DS: sulfato de dermatana; KS: sulfato de queratana; CS: sulfato de condroitina. Adaptado de Neufeld and Muenzer, 2001.

Mucopolissacaridose tipo I

A mucopolissacaridose tipo I (MPS I) é causada pela deficiência da enzima α -L-iduronidase (IDUA), uma glicosidase que hidroliza resíduos terminais ácidos α -L-idurônicos dos GAGs sulfato de dermatana e sulfato de heparana (Figura 2). A síntese da enzima inicia com uma forma precursora com 100 aminoácidos a mais que a forma madura. Ao ser processada intracelularmente, a enzima perde os aminoácidos extras e recebe resíduos de manose-6-fosfato (M6P) que permitirão o seu direcionamento ao lisossomo, onde o pH ácido tornará a molécula uma enzima ativa (Saftig and Klumperman, 2009).

A deficiência na atividade enzimática de IDUA pode levar a um amplo espectro de manifestações clínicas, classicamente agrupadas nas formas grave, intermediária e atenuada, embora a transição entre os grupos seja contínua e a classificação muitas vezes é complexa e imprecisa, preferindo-se a utilização de apenas dois grupos: grave e atenuado (Giugliani, et al., 2010). A forma grave de MPS I é também historicamente conhecida como Síndrome de Hurler (MIM#607014); a principal característica que normalmente a diferencia da forma atenuada é o comprometimento neurológico evidente. Além disso, os pacientes apresentam face infiltrada, hepatoesplenomegalia, insuficiência respiratória e cardíaca, valvulopatia, opacificação da córnea e perda de audição. As manifestações musculoesqueléticas incluem rigidez nas articulações, mãos em garra, contraturas e disostose múltipla. Pacientes com a forma grave apresentam os sinais e sintomas logo no primeiro ano de vida e a progressão é rápida, com expectativa de vida inferior aos 10 anos, se não tratado (Giugliani, et al., 2010). Formas intermediárias da doença, classicamente designadas Síndrome de Hurler-Scheie (MIM#607015), apresentam manifestações viscerais similares e comprometimento neurológico presente ou ausente, manifestando-se em alguns casos apenas como dificuldades de aprendizado (Giugliani, et al., 2010).

A forma atenuada, ou Síndrome de Scheie (MIM#607016), por definição, exclui sinais neurológicos. Os pacientes apresentam a face infiltrada característica, mas a estatura pode ser normal. Rigidez articular e mãos em garra são comuns. Além da opacificação da córnea, também podem ocorrer glaucoma e degeneração da retina que, juntos, podem comprometer significativamente a visão. Manifestações somáticas ainda incluem obstrução das vias aéreas, valvulopatia e hérnias, muitas vezes necessitando de cirurgias corretivas. O surgimento dos sinais e sintomas é notado normalmente após os 5 anos e o diagnóstico comumente efetuado entre 10 e 20 anos de idade (Neufeld and Muenzer, 2001).

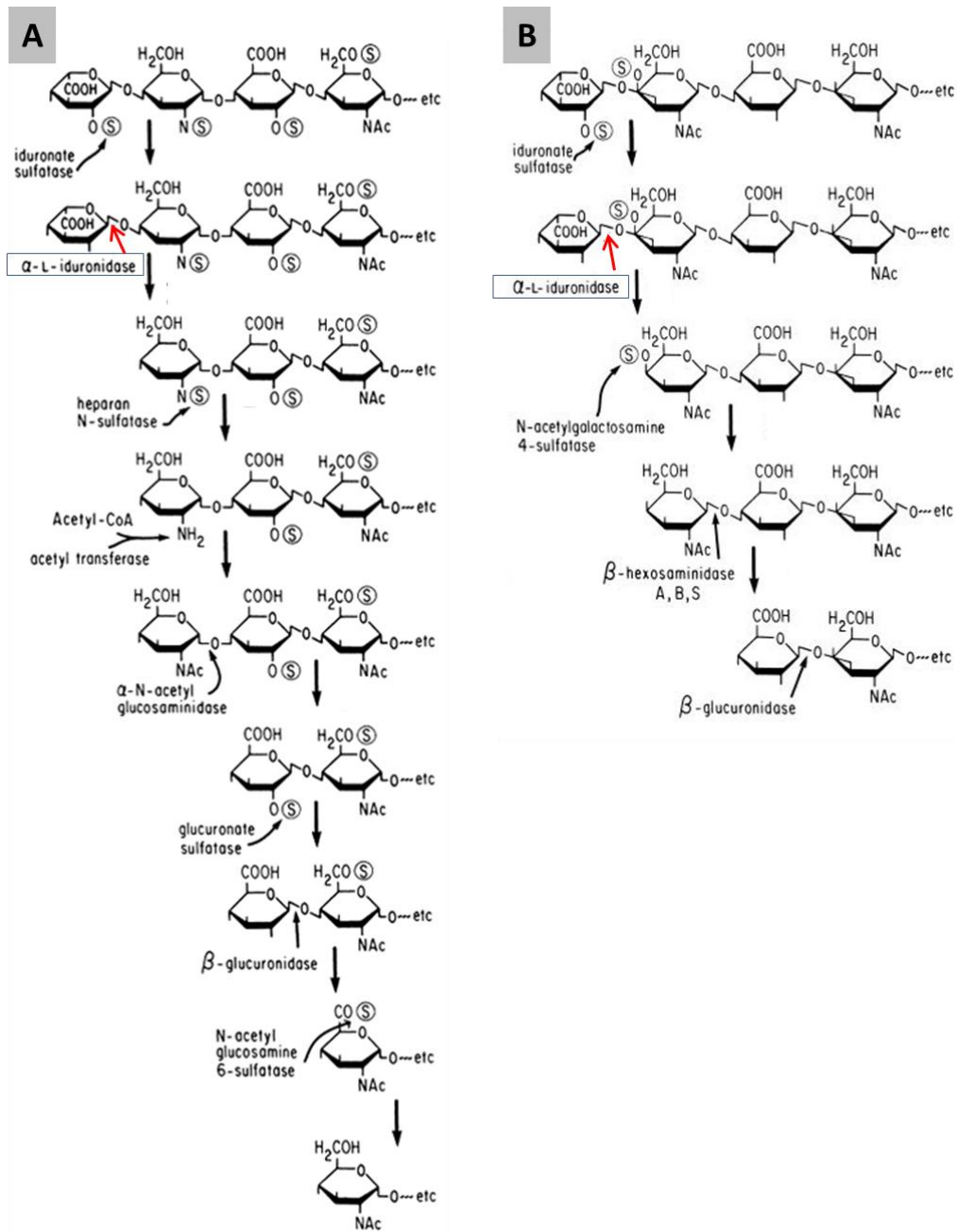


Figura 2. Via de degradação dos GAGs sulfato de heparana (A) e sulfato de dermatana (B). A enzima α -L-iduronidase (em destaque) atua na segunda etapa da degradação destes GAGs, removendo o ácido L-idurônico na posição terminal da molécula. Imagem adaptada de Neufeld and Muenzer, 2001.

A incidência de MPS I pode variar muito entre países, sendo de 1:26.000 nascimentos na Irlanda, 1:75.000 em Portugal, 1:88.000 na Austrália e 1:145.000

na Alemanha (Baehner, et al., 2005; Fuller, et al., 2006; Murphy, et al., 2009). A incidência exata de MPS I no Brasil ainda está sendo investigada, mas a Rede MPS Brasil já registra 1069 diagnósticos de MPS entre 2004 e 2013, sendo que 197 (18%) destes são de MPS I (Federhen, et al., 2016).

Tratamento

Atualmente, existem dois tratamentos aprovados para a MPS I, a terapia de reposição enzimática (TRE) e o transplante de células tronco hematopoiéticas (TCTH), que podem, inclusive, ser usados concomitantes para um melhor resultado (Ghosh, et al., 2016).

A TRE para MPS I foi desenvolvida baseada na ideia de que a enzima deficiente poderia ser fornecida de maneira exógena, captada pelos lisossomos e, então, atuar no catabolismo dos GAGs de maneira suficiente para impedir as manifestações da doença. Aprovada comercialmente em 2003, a enzima recombinante desenvolvida, laronidase (Aldurazyme[®], Genzyme Corporation), contém os resíduos de M6P necessários para ser reconhecida e internalizada no lisossomo (Jameson, et al., 2016). O tratamento é contínuo e para a vida toda, com infusões semanais da enzima. Além de reduzir GAGs urinários a níveis normais, o tratamento beneficia os pacientes com melhora significativa nas manifestações viscerais, como redução do volume hepático e melhora da função pulmonar (com diminuição de hipoventilação noturna e apnéia do sono). Na maioria dos casos, há um aumento da resistência física avaliada no teste de caminhada de 6 minutos, aumento de flexibilidade em ombros e joelhos e crianças com baixa estatura tendem a ter um aumento na taxa de crescimento. Alguns pacientes apresentam melhora na acuidade visual, embora a opacificação da córnea não seja prevenida (Clarke, et al., 2009; Jameson, et al., 2016). Por outro lado, o tratamento só é tão eficaz em casos mais atenuados da doença e quando iniciado antes do surgimento dos sintomas mais graves, não tendo ação sobre manifestações cardíacas e neurológicas já estabelecidas (Al-Sanna, et al., 2015; Gabrielli, et al., 2016). Efeitos adversos também são relatados, como febre,

dores de cabeça e erupções cutâneas, além do risco de choque anafilático (Clarke, et al., 2009).

O TCTH é usado desde 1981 no tratamento de MPS I, sendo a melhor alternativa para pacientes com o fenótipo grave, uma vez que é um tratamento realizado uma única vez e que tem potencial para atuar em órgãos nos quais a TRE é ineficiente, como o sistema nervoso central (Beck, et al., 2014). Quando o transplante é bem sucedido, as células doadoras produzem enzima suficiente para impedir o avanço da doença e a expectativa de vida do paciente pode aumentar significativamente. Entretanto, devido à mortalidade associada ao transplante, este tratamento é indicado somente para pacientes com a forma grave da doença, antes de 2,5 anos de idade e que ainda não tenham manifestado sintomas neurológicos (Giugliani, et al., 2016).

O uso combinado de TRE e TCTH ajuda a diminuir a mortalidade associada ao transplante, uma vez que melhora a condição somática do paciente e o condiciona para o procedimento, mantendo-se a TRE até que as células doadoras se estabeleçam no receptor (Ghosh, et al., 2016). Adicionalmente, a associação dos tratamentos pode preservar significativamente a cognição em comparação com o TCTH isolado, sendo considerada a melhor opção de tratamento atualmente (Eisengart, et al., 2013).

A busca por novas terapias é constante e, recentemente, novos protocolos estão sendo desenvolvidos – como a utilização de moléculas de transleitura de códons de parada prematuros (*stop-codon read through*) e a edição gênica por CRISPR-Cas9 (*Clustered Regularly Interspaced Short Palindromic Repeats – CRISPR associated protein 9*). A primeira técnica utiliza moléculas, como aminoglicosídeos, capazes de alterar um códon de terminação prematuro por outro códon qualquer no momento da tradução, de maneira randômica, que resulta na produção de uma proteína de tamanho normal com atividade enzimática normal ou residual – ao invés da degradação prematura do RNA mensageiro (mRNA) instável e da formação de proteínas truncadas sem atividade alguma. Algumas vantagens incluem a administração por via oral, que facilita a adesão ao tratamento, e a permeabilidade à barreira hemato-encefálica, que

permitiria a ação do fármaco no sistema nervoso central (Giugliani, et al., 2016). A segunda proposta de tratamento, com a utilização de edição gênica, almeja a correção das mutações presentes no DNA do paciente numa abordagem *in vivo* – com a injeção dos vetores de edição direto no paciente – ou *ex vivo* – com a correção em laboratório de células coletadas do paciente e posterior transplante no indivíduo. Esta última considera a utilização de, principalmente, células-tronco hematopoiéticas, o que mimetizaria o TCTH já consolidado na clínica – a vantagem, neste caso, é da possibilidade de um transplante autólogo, eliminando a necessidade de doador compatível e reduzindo consideravelmente a mortalidade associada ao transplante observada atualmente (Carvalho, et al., 2015).

Diagnóstico

Pacientes com a forma grave da doença costumam ser diagnosticados por volta de 1 ano de idade, aproximadamente 6 meses após o surgimento dos sintomas mais perceptíveis; enquanto que pacientes com a forma atenuada podem levar alguns anos (Beck, et al., 2014). Em frente a uma suspeita clínica de MPS, análises bioquímicas ou moleculares podem ser conduzidas. O diagnóstico bioquímico consiste em análise de GAGs urinários e ensaio enzimático da enzima deficiente em amostras de plasma, soro, leucócitos e/ou de fibroblastos obtidos e cultivados a partir de uma biópsia, sendo este último considerado definitivo (Coutinho, et al., 2012).

O diagnóstico molecular consiste na busca pelas mutações mais frequentes ou no sequenciamento completo do gene *IDUA*. Ele é realizado em pacientes já diagnosticados bioquimicamente para fins de prognóstico, auxílio na escolha do tratamento e aconselhamento genético, entre outras aplicações.

Aspectos moleculares

Localizado no locus 4p16.3, o gene *IDUA* foi clonado em 1992 (Scott, et al., 1992a), revelando aproximadamente 19 mil pares de base divididos em 14 éxons – que codificam uma proteína de 653 aminoácidos. Atualmente, mais de 200 mutações patogênicas já foram descritas, sendo que a maioria é sem sentido ou de sentido trocado (58%), seguido de mutações de *splicing* (16%) e pequenas deleções e inserções (14% e 7%, respectivamente) (Stenson, et al., 2014).

A primeira mutação patogênica descrita no gene *IDUA* foi a sem sentido p.Trp402*, encontrada em 31% dos alelos em uma amostra de 64 pacientes com MPS I e associada com o fenótipo grave da doença. Na ocasião, os autores já sugeriam que a presença dessa mutação poderia ser um indicador para realização de TCTH (Scott, et al., 1992b). No mesmo ano, em uma amostra de 73 pacientes, o mesmo grupo descreveu as mutações p.Gln70* e p.Pro533Arg, com frequências de 15% e 3%, respectivamente – representando, junto com a p.Trp402*, 53% dos alelos dos pacientes analisados (Scott, et al., 1992c).

A relação genótipo-fenótipo para pacientes com MPS I não está completamente estabelecida. Sabe-se que pacientes com duas mutações sem sentido apresentam a forma mais grave da doença (Terlato and Cox, 2003), uma vez que a maior parte do mRNA transcrito vai para a via de degradação mediada por mutação sem sentido (*Nonsense-mediated decay*), um sistema de proteção celular contra proteínas truncadas (Menon and Neufeld, 1994). Por essa razão, a atividade enzimática desses pacientes é praticamente indetectável, resultando em manifestações mais graves da doença. Mutações em sítios de *splicing* também costumam gerar mRNAs instáveis e proteínas defeituosas, gerando fenótipos graves; porém, há casos de alterações menores que resultam em atividade enzimática residual e consequente atenuação do fenótipo (Terlato and Cox, 2003).

Mutações de sentido trocado podem ter resultados variados, dependendo da posição na proteína e do aminoácido trocado. Mutações nos sítios de glicosilação ou no sítio ativo são bastante deletérias e resultam na perda de função da

enzima, provocando formas mais graves da doença quando em homozigose (Bie, et al., 2013). Por outro lado, a troca de aminoácidos em posições de função secundária, que acarretem somente pequenas modificações estruturais, permitem a síntese de enzimas com atividade residual suficiente para prevenir grande acúmulo de substrato. Nestes casos, a atividade residual é significativa mesmo que proveniente apenas de um alelo, resultando em formas intermediárias e leves de MPS I (Saito, et al., 2014). Pequenas deleções e inserções estão relacionadas com fenótipos variados, com casos graves e atenuados descritos (Terlato and Cox, 2003).

A maioria das mutações mais frequentes relacionadas à MPS I tem uma relação genótipo-fenótipo praticamente estabelecida. Uma exceção é a mutação p.Pro533Arg, encontrada em homozigose em pacientes com diferentes espectros da doença – desde formas muito graves à formas intermediárias e leves. O motivo da variação ainda não é conhecido, entretanto o prognóstico para estes pacientes costuma ser melhor do que o de homozigotos para mutações sem sentido (Terlato and Cox, 2003).

A relação genótipo-fenótipo pode auxiliar na identificação de importantes sítios e estruturas da proteína que, conseqüentemente, melhoram o entendimento sobre a doença e oportunizam novas abordagens terapêuticas. Além disso, por ser um indicativo de prognóstico possível de ser obtido precocemente, o genótipo do paciente pode ajudar na escolha do tratamento – em caso de mutações conhecidamente relacionadas à forma grave, pode-se aconselhar a realização de TCTH antes mesmo do surgimento dos sintomas mais graves, o que seria a conduta ideal. A genotipagem dos pacientes também é benéfica para a família dos indivíduos afetados pois facilita e agiliza o diagnóstico pré-natal e pré-implantacional de novas gestações e permite o aconselhamento genético adequado.

Por fim, a identificação de mutações frequentes em diferentes populações pode auxiliar no desenvolvimento de novas terapias, uma vez que o genótipo do paciente interfere em alguns possíveis tratamentos – a terapia com transleitura de códons de parada prematuros só terá aplicação em pacientes com mutações

sem-sentido; já a edição gênica, por ser direcionada especificamente aos sítios das mutações, só é possível após a genotipagem dos pacientes, o que reforça a importância da caracterização molecular da MPS I.

OBJETIVOS

Objetivo geral

O objetivo deste trabalho é genotipar os pacientes brasileiros com MPS I e relacionar as mutações observadas com dados descritos em outras populações no mundo.

Objetivos específicos

- Sequenciar a região codificante do gene *IDUA* em pacientes brasileiros com MPS I;
- Caracterizar mutações novas com análises de bioinformática e relacionar com os dados clínicos dos pacientes;
- Atualizar o perfil mutacional dos pacientes brasileiros com MPS I, descrevendo quais as mutações no gene *IDUA* são mais frequentes nesta população;
- Analisar as frequências das mutações patogênicas no gene *IDUA* de pacientes com MPS I em todas as populações com dados da literatura;
- Avaliar a presença dos alelos patogênicos mais comuns no gene *IDUA* em indivíduos saudáveis do banco de dados gnomAD e correlacionar com os dados obtidos em pacientes de mesma origem étnica;
- Sugerir a possível origem dos alelos mais comuns em diferentes populações, especialmente na população brasileira.

RESULTADOS

Os resultados desta dissertação serão apresentados a seguir, em formato de artigo científico com o título “Molecular analysis of MPS I patients in Brazil and geographic distribution of common *IDUA* mutations worldwide” a ser submetido para o periódico Human Mutation.

Molecular analysis of MPS I patients in Brazil and geographic distribution of common IDUA mutations worldwide

Édina Poletto^{1,2}, Ursula Matte^{1,2} e Guilherme Baldo^{1,2}.

¹Post-Graduate Program on Genetics and Molecular Biology, UFRGS, Porto Alegre, RS, Brazil

²Gene Therapy Center, Experimental Research Center, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil

Corresponding Author

Guilherme Baldo

Gene Therapy Center, Experimental Research Center, Hospital de Clínicas de Porto Alegre

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

Tel +55 51 3359 8838

Abstract

Mucopolysaccharidosis type I (MPS I) is an autosomal recessive disease caused by mutations in the α -L-iduronidase (*IDUA*) gene, resulting in decreased activity of the lysosomal enzyme with the same name and consequent accumulation of glycosaminoglycans (GAGs) in lysosomes. There are more than 200 disease-causing mutations reported and the mutational profile varies considerably from country to country. In this study, we performed molecular analysis of 92 DNA samples, representing the largest cohort of Brazilian MPS I patients examined. Additionally, we analysed allelic frequencies from other countries to better understand the composition of our population and the allele distribution worldwide. Thirty-nine mutations were identified in our sample and 12 of them are newly described: c.78delC, c.159-1_-23del, p.Gln125*, p.Trp175*, c.590-6ins4G, c.763delC, c.973-1G>A, p.Asp349Glu, p.Asn350Lys, p.Lys384Asn, c.1403-12_-4delCAGGCCCCCG and c.1522_1525delGAGG. *In silico* analysis of novel mutations suggest they are possibly pathogenic, supporting previous biochemical and clinical diagnosis. Among recurrent mutations, p.Trp402* and p.Pro533Arg are the most frequent in Brazil, with 38% and 17% of alleles, respectively. The first is also the major allele in European descent countries, while the second is found in the Mediterranean region, mainly North Africa. The second most frequent allele worldwide and highly prevalent in Scandinavia, p.Gln70*, was relatively rare in our sample, with only 3% of alleles. These results reveals the great allelic heterogeneity of *IDUA* not only in Brazilian patients, which differs from all countries analysed, but also between geographically close populations. The knowledge of the genetic background of MPS I may improve diagnosis and management of patients and families with the disease.

Introduction

Mucopolysaccharidosis type I is an autosomal recessive lysosomal storage disorder caused by mutations in the α -L-iduronidase (*IDUA*) gene, which encodes for the lysosomal enzyme with same name. The *IDUA* protein (EC 3.2.1.76) is involved in the breakdown of glycosaminoglycans (GAGs) dermatan sulfate and heparan sulfate. When defective, it leads to accumulation of partially degraded GAGs in many organs and tissues, resulting in a progressive multisystemic disease. A wide range of clinical manifestations is observed, including: coarse facial features, dysostosis multiplex, visceromegaly, corneal clouding, hearing loss, inguinal and umbilical hernias, valvular heart disease, dilated cardiomyopathy, restrictive lung disease, airway obstruction and cognitive and developmental delay (absent in the attenuated forms of the disease). According to the severity of the symptoms, MPS I is classically divided into three phenotypes – the severe form, or Hurler syndrome (MPS IH; MIM#607014); the attenuated form, or Scheie syndrome (MPS IS; MIM#607016) and the intermediate form, also known as Hurler-Scheie syndrome (MPS IH/S; MIM#607015) (Neufeld and Muenzer, 2001). However, the clinical spectrum is continuous and the classification is frequently complex.

The *IDUA* gene is located at the locus 4p16.3, it is 19 Kb long and composed of 14 exons. More than 200 pathogenic mutations have been described so far; most of them are missense/nonsense (58%), followed by mutations on splice sites (16%) and small deletions and insertions (14% and 7%, respectively) (Stenson, et al., 2014). Patients with two nonsense mutations normally manifest the most severe form of the disease, due to complete absence of the functional enzyme. On the other hand, many missense mutations result in residual enzyme production, leading to attenuated phenotypes even when the patient also present a nonsense mutation in the other allele (Terlato and Cox, 2003).

The knowledge of the mutation profile in a given population enables the provision of proper patient care. Our group had previously described the mutation frequency in a sample of Brazilian patients diagnosed until the early 2000's (Matte, et al., 2000; Matte, et al., 2003). In this study, we present a more

comprehensive view of the Brazilian mutation profile, describing findings from 92 MPS I patients diagnosed over the last 16 years, with the description of 12 novel mutations. In addition, we provide an analysis of the *IDUA* mutation profile of MPS I patients from all populations with published data and the combined allelic frequencies worldwide, highlighting the high allelic heterogeneity of this gene and the significant differences between populations.

Methodology

Patients

This study comprises a total of 92 unrelated DNA samples collected from 2000 to 2016. These patients are from all Brazilian regions, biochemically diagnosed with MPS I by both quantification of urinary GAGs and IDUA activity in plasma and/or in leukocytes. The study was approved by the Ethics Research Committee of our institution, and the patients gave written informed consent.

Molecular Analysis

Whole blood genomic DNA was extracted from EDTA tubes either by standard salting-out method or commercial kits according to manufacturer's instructions (Wizard Genomic DNA Purification from Promega, Easy-DNA or iPrep PureLink gDNA Blood Kit, from Thermo Fisher Scientific, USA). All 14 *IDUA* exons, their intronic flanking regions and part of the 5' and 3' untranslated regions were amplified by PCR, using the primers and conditions listed in suppl. table 1. PCR was performed in Veriti thermal cycler (Applied Biosystems Thermo Fisher Scientific, USA), as follows: initial denaturation at 95°C for 5 min, then 40 times of 95°C for 30", T_m for 30" and 72°C for 30", with final extension at 72°C for 10min. Exceptionally for primers pairs amplifying exons 9/10 and 7/8, cycling was 40x 95°C 1', T_m 1', 72°C 1' and 40x 95°C 30", T_m 30", 72°C 1', respectively. Amplicons were checked in agarose gel, then purified with PEG 8000 20% NaCl 2.5M (v/v) and quantified with Low Mass Reader (Thermo

Fisher Scientific, USA). Sequencing was done in ABI3500 genetic analyzer using BigDye Terminator v3.1 (both from Applied Biosystems Thermo Fisher Scientific, USA) and the same set of primers. Sequences were analysed by comparison to genomic reference sequence (GenBank accession number NG_008103) and all alterations were confirmed by reverse strand sequencing of PCR duplicate.

In one case, RNA sample from a patient with a novel splicing mutation was available, thus RNA was extracted using PAXgene Blood RNA System (PreAnalytiX QIAGEN, Germany) and cDNA was obtained with High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems Thermo Fisher Scientific, USA), both according to manufacturer's protocol. PCR amplification of cDNA was performed using primers that covered exons 3 to 7 and reaction described in suppl. table 1, in a Veriti thermal cycler with the following configuration: initial denaturation at 95°C for 5 min, then 40 times of 95°C for 30", Tm for 30" and 72°C for 30", with final extension at 72°C for 10min. Amplicons were checked in agarose gel for contamination and product size. In the case of multiple bands, they were purified from the gel using QIAquick Gel Extraction Kit (QIAGEN) and then sequenced as aforementioned and compared with the reference sequence NM_000203.4, from GeneBank.

Bioinformatics

Novel mutations were analysed with Polyphen2, PROVEAN, SIFT, PhD-SNP, PANTHER, SNPs&GO, MutPred, NetGene 2, Human Splice Finder and Variant Effect Predictor. Alignment of IDUA amino acid sequences across different species was performed with Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo) to check for evolutionary conservation. Species and sequences utilized are described in Figure 1.

Allelic frequencies analysis

Literature search was performed using the key terms “*IDUA* mutation”, “mutation frequency *IDUA*” and “MPS I mutation” on Pubmed (www.ncbi.nlm.nih.gov/pubmed) and through references cited in related papers. Inclusion criteria used were: information about the country of origin of patients and the absolute number of alleles in the study and for each mutation. Patients from multicenter studies were included if their nationality was declared. Thirty-five papers were selected and used in the analysis.

Patients from all studies were grouped according to their country of origin and, for the countries with more than 10 alleles described in total, the combined frequencies of each mutation were calculated. The three most frequent mutations from each country were tabled separately and used to analyse the differences among populations. The most common mutation from each country was searched in Genome Aggregation Database (gnomAD) (Lek, et al., 2016) to evaluate its allelic frequency in the general population, grouping them in eight different ethnic groups: European, Finnish European, Ashkenazi Jewish, African, Latino, East Asian, South Asian and Other (corresponding to individuals not clustered in the major ethnic groups).

Results

Molecular analysis

The entire *IDUA* gene from 92 DNA samples was sequenced for screening of sequence variations possibly related to MPS I. We were unable to identify the second pathogenic mutation in two patients (1%) with confirmed clinical and biochemical MPS I diagnosis. In the whole sample, a total of 39 mutations were identified, including 12 novel (Figure 1). Overall, nonsense mutations represented 47% of alleles, closely followed by 40% of missense mutations; while small deletions or duplications and splice site mutations corresponded to 7% and 5%, respectively. Doubtlessly, the most frequent mutation observed was p.Trp402*, corresponding to 38% of alleles (71 in total, including 27 homozygous patients). Other nonsense mutations found were p.Gln70* with 3% of total alleles, p.Arg619* and p.Arg621*, with 1% each. The missense

p.Pro533Arg was the second most common mutation in the sample, counting for 17% of alleles (32 in total, 10 homozygous patients), followed by p.Gly208Asp (6%) and p.Tyr625Cys and p.Arg383His, with frequencies of 3% each. Splice site mutations had low frequencies, about only 1% each mutation identified, as well as duplications or small deletions, except for the most frequent, c.144_146delGAG, which was observed in only 6 alleles (3%). No gross deletions, gross insertions or complex rearrangements were found. All mutations identified and their frequencies are summarized in Table 1 and their location along the *IDUA* gene is illustrated in Figure 2.

Clinical characterization of novel mutations

Twelve novel mutations were identified (Table 2). The splicing mutation c.973-1G>A was diagnosed along with p.Tyr625Cys in a patient at 2 years of age (pt.7), but the phenotype classification and clinical manifestations was not available. Two patients were homozygous for novel mutations. Patient 3 (p.Gln125*) presented the severe form of the disease, usually related to nonsense mutations. Similarly, patient 2, homozygous for the deletion c.159-1_-23del, had almost undetectable *IDUA* levels and presented relevant manifestations, finally classified as having the severe phenotype.

Patients with the missense mutations p.Asp349Glu and p.Lys384Asn, the nonsense p.Trp175* and deletions c.763delC, c.1403-12_-4delCAGGCCCCCG and c.1522_1525delGAGG were classified as having the severe phenotype. However, with the exception of pt. 12, neurological regression was not reported at the time of diagnosis (probably because some patients were very young). Apart from this, they presented many important clinical manifestations, including motor delay, cardiovascular diseases, respiratory difficulties, hearing loss, dysmorphias and dysostosis multiplex.

Two patients with the attenuated form, pt.1 and pt.5, had the mutations c.78delC and c.590-6ins4G, respectively. They both were diagnosed when older than 20 years of age and present milder clinical symptoms, with no motor delay nor neurologic regression. Patient 6 is currently 32 years old and presents

hepatomegaly, coarse facial features, corneal clouding, valvulopathy, hernia, joint stiffness, respiratory difficulties and claw hands. Unfortunately, there are no available data about patient 1 besides his age (42 years old) and the absence of neurologic and motor delay.

In silico characterization of novel mutations

All novel variants identified in this study were submitted to pathogenicity prediction softwares in order to confirm the causative alleles for the disease (Table 3). All missense mutations (p.Asp349Glu, p.Asn350Lys, p.Lys384Asn) were considered pathogenic by at least 5 out of the 7 tools used, suggesting they are potentially deleterious. In addition, the multiple alignment analyses shows that the amino acids altered at these sites are in conserved positions across several different species (Figure 1A).

Four variations were classified as splice site mutations by at least one of the softwares used. The mutation c.159-1_-23del is predicted to result in inclusion of part of intron 1 and partial loss of exon 2 due to activation of cryptic acceptor and donor splice sites upstream the wild-type ones. Similarly, mutations c.973-1G>A and c.1403-12_-4delCAGGCCCCCG lose the splice acceptor sites, possibly provoking complete loss of exons 8 and 10, respectively. Moreover, in addition to the software prediction, the effect of the splicing mutation c.590-6ins4G was analysed also in the cDNA of the patient – it results in loss of most of exon 6 and inclusion of intron 6 in (Figure 1D).

The other mutations were classified as frameshift variants with creation of premature stop codons (c.78delC and c.763delC) or as nonsense mutations (p.Gln125* and p.Trp175*). Although not recognized by any softwares used, the deletion c.1522_1525delGAGG alters the reading frame from amino acid 508 leading to a premature stop codon after 16 amino acids, possibly resulting in a non-functional enzyme.

Mutational profile worldwide

Results from our analysis of combined frequencies of the most frequent *IDUA* mutations in MPS I patients worldwide are presented in Table 4, for all countries with published studies using samples larger than 10 alleles. The p.Trp402* is by far the most common mutation, being the major allele in North America, Colombia, Brazil, United Kingdom, Netherlands, Germany, Czech Republic and Slovakia, Spain and Australia (Figure 3). It is followed by p.Gln70*, present mainly in North Europe, as Norway and Russia, but also in Poland and Austria. Italy has the most variable profile observed: although results showed p.Gln70* as the most frequent mutation, it corresponds to only 14% of alleles, closely followed by p.Pro496Arg (13%), p.Gly51Asp (12%) p.Trp402*(11%) and p.Pro533Arg (10%).

North African countries Morocco, Algeria and Tunisia have p.Pro533Arg as the most frequent mutation, corresponding to 92%, 81% and 54% of alleles, respectively. This missense mutation is also present in other Mediterranean countries, such as Turkey, Spain and Italy – mainly in the Sicily region (42%) (Gatti, et al., 1997) (Figure 3) – though it is very rare in United Kingdom, Australia and North America.

Mutations frequently observed in East Asians were not found in Western populations, as c.1190-1G>A, p.Ala79Val, p.Leu346Arg and c.613_617dupTGCTC. Contrarily, the most frequent mutation among Japanese patients, p.Arg89Gln, was found in patients from Brazil and United Kingdom.

Israel has the most distinct mutational profile analysed, as its two more frequent mutations were not found in any other population (p.Thr366Pro and p.Gln310*). Mutations in Turkish patients were also present in other populations, except for the most frequent, c.494-1G>A, found only in one homozygous patient in Brazil. India shares alleles mainly with United Kingdom (p.Arg619* and p.Ala75Thr), although p.Leu121Pro seems exclusive.

In gnomAD, allelic frequency among individuals without the disease mirrors the geographic localization of most mutations analysed in patients. For example, the p.Trp402* allele was observed in Europeans (0.0014), Others (0.0012), Latinos (0.0004) and Finnish Europeans (0.0003), but it was not present in Asians or Africans (Table 5). Accordingly, frequent mutations

observed in patients from East Asia, p.Leu346Arg and p.Arg89Gln, seem exclusive to East Asians or at least 5-fold more frequent in this population than in Europeans, respectively.

The second most frequent allele in patients, p.Gln70*, is also the second in healthy subjects, with the very high proportion of 0.002 in Finnish Europeans, followed by 0.0007 in Non-Finnish Europeans and none allele found in Latinos nor Asians. This is in agreement with the very low frequency of p.Gln70* in our Brazilian MPS I sample.

Besides p.Trp402*, Latinos presented only the p.Pro533Arg allele (0.0002), also found in Others and Europeans. Nonsense p.Arg619* is shared by Others and Europeans, being more frequent in the first group. Turkey's most frequent mutation c.494-1G>A was observed in Europeans only. Ashkenazi Jews had none of the alleles analysed and the most frequent mutation in Israel, p.Thr366Pro, and China, c.1190-1G>A, were not listed on gnomAD.

Discussion

Characterizing a given population is important to provide improved medical care, with faster diagnosis, prenatal testing or assessing the choice of treatment. This study updates the allelic frequencies of the *IDUA* gene among Brazilian patients with MPS I, previously reported (Matte, et al., 2000; Matte, et al., 2003; Pereira, et al., 2008), though with a bigger sample – in the last 16 years, 92 DNA samples were sequenced.

Of the 184 alleles analysed, only two remained without the pathogenic mutation identified, probably due to limitations in our methodology, which analyses only coding sequences and their flanking regions. Thirty-nine mutations were identified in our sample, of which 12 were novel (around 7% of alleles). This highlights the allelic heterogeneity of *IDUA*, also observed in other populations and discussed further. All novel mutations found in this study were analysed *in silico* to evaluate their pathogenicity and detailed biochemical and clinical data from the patients carrying those variants were collected when possible, to allow some correlation between genotype and phenotype.

According to the *in silico* analysis, all novel missense mutations found in our sample are probably pathogenic – they were considered deleterious by at least 5 out of 7 prediction softwares and the multiple alignment with IDUA homologue sequences showed they were placed on conserved regions of the protein. In addition, patients carrying these mutations had urinary GAGs up to 3 times the normal range and IDUA activity at undetectable levels at the time of diagnosis.

Few of these novel mutations occur in positions already affected by other known disease-causing mutations, as the p.Asn350Ile, a substitution of an amino acid with polar neutral side chain (asparagine) for one with hydrophobic side chain (isoleucine), which causes the severe phenotype due to failure in protein processing and thus absent activity (Matte, et al., 2003). Since the novel p.(Asn350Leu) substitutes the asparagine for a leucine, a non-polar amino acid as isoleucine, it is possible that the effect in the protein will be very similar to the mutation already described. Accordingly, p.(Asp349Glu) was found in a severe form patient and the described mutations in the same site (p.Asp349Asn and p.Asp349Tyr) were associated with this phenotype as well. Moreover, the residue Asp349 is involved in substrate binding, therefore essential to the enzyme catalytic activity (Bie, et al., 2013).

There are two known mutations affecting the flanking amino acids (p.Arg383His and p.Pro385Arg) of the novel missense p.(Lys384Asn), identified in a patient with severe phenotype. By all *in silico* analysis, this location seems conserved and an amino acid change may be pathogenic.

Four splice site mutations were identified. The sequencing of cDNA from patient 5 confirms the *in silico* analysis, showing that the insertion of 4 guanines at intron 5 (c.590-6ins4G) alters the splice site, since the transcript skips almost all exon 6 and has the sequence of intron 6 instead, followed by exon 7 (Figure 1D). Considering that the other allele contains a known severe nonsense mutation (p.Tyr343*) and this patient has the attenuated form of the disease, the variant transcript probably results in a slightly active enzyme, enough to prevent the more severe manifestations. The mutation c.973-1G>A was found in a patient (pt. 7) without clinical data available, therefore it is not possible to

rule out further consequences besides causing the disease. On the other hand, the other splice site mutations, c.159-1_-23del and c.1403-12_-4delCAGGCCCCG, are related to the severe form of MPS I, since both were identified in severely affected patients in either homozygosis or in heterozygosis with stop codon mutation (pt. 2 and pt 11, respectively).

Regarding worldwide allelic frequencies, p.Trp402* is clearly the most frequent mutation, followed by p.Gln70* and p.Pro533Arg. The first is present mainly in Europe (United Kingdom, Spain, Germany) and overseas countries colonized by Europeans (Australia, United States, Mexico, Colombia, Brazil). This mirrors the allelic frequency in the general population, according to gnomAD: p.Trp402* is only found in Europeans, Latinos and Others, without a single allele in the ethnic groups Africans or Asians. The second most frequent, p.Gln70*, is localized mainly in Scandinavian countries, being present in 0.2% of Finnish Europeans and corresponding to 50% of alleles in MPS I patients from Norway. This is the main allele among patients in Poland and Italy also, which could explain its presence in Brazilian patients through 19th century migrations, although very rare (8 alleles out of 312).

According to gnomAD, Africans have p.Gln70* in a proportion of approximately 0.0002 in the general population, though there are no studies with Sub-Saharan African MPS I patients to cross-reference this, since the genetic profile from North Africans (with available data) is significantly diverged (Botigue, et al., 2013). Consistently, the mutation p.Pro533Arg is very frequent in North Africa, corresponding to 92% of mutated alleles in Morocco and 81% in Algeria, although it was not found in the genome of the 3702 Africans from gnomAD. p.Pro533Arg probably originated in North Africa and spread to southwestern European countries with Moorish conquests in the Middle Age, already considered an important gene flow event (Botigue, et al., 2013). Southwestern Europeans, specially Portuguese, Spanish and Italians, eventually migrated to Brazil and other American countries (as Mexico), which justifies the high allelic frequency observed in Latinos (0.02%) and MPS I patients (58 out of 312 alleles in Brazil). Migrations also influenced other countries, as the most frequent mutation in India, p.Arg619*, is also relatively frequent in the UK, for example.

The most frequent mutations in western countries are practically not observed in Asians, though. Accordingly, mutations present in East Asians remained almost exclusively to this population, with very few cases in other countries (as p.Arg89Gln in Brazil). Korea, geographically located between China and Japan, shares p.Ala79Val and p.Leu346Arg with China, and c.613_617dupTGCTC with Japan.

Conclusion

This study comprises 92 Brazilian MPS I patients sequenced from 2000 to 2016, which is a considerable sample size since this is a rare disease, with an estimate incidence of 1:100.000 live births (Beck, et al., 2014). Moreover, twelve mutations were newly identified in patients with confirmed biochemical and clinical diagnosis of MPS I and most of them were predicted to be pathogenic according to the *in silico* analysis performed.

Although MPS I has a continuous clinical spectrum and to define a genotype-phenotype correlation is frequently difficult, it is important to perform molecular and clinical characterization of variants, so it can facilitate the diagnosis (including prenatal), improve prognostic and choice of treatment and provide proper genetic counseling.

The mutations p.Trp402* and p.Pro533Arg are the most frequent in Brazilian MPS I patients, representing together 55% of mutated alleles in our sample. Target analysis of exons 9, 11, 1, 12 and 8 covers more than 70% of all genetic variations found in this population, which could reduce time and cost of diagnosis by standard Sanger sequencing. Moreover, other methodologies as the use of Taqman assays for p.Trp402* and p.Pro533Arg could be applied as a first tier screening of patients.

The mutational profile in Brazilian MPS I patients is very characteristic of the populations that colonized and migrated to the country, mainly Europeans, and it is relatively similar to other European former colonies, as Australia and United States, which also have p.Trp402* as the major allele. However, only Brazil has high occurrence of p.Pro533Arg, a mutation present in Mediterranean

countries, while the second most frequent allele among MPS I patients in the world, p.Gln70*, is rare in Brazilian patients.

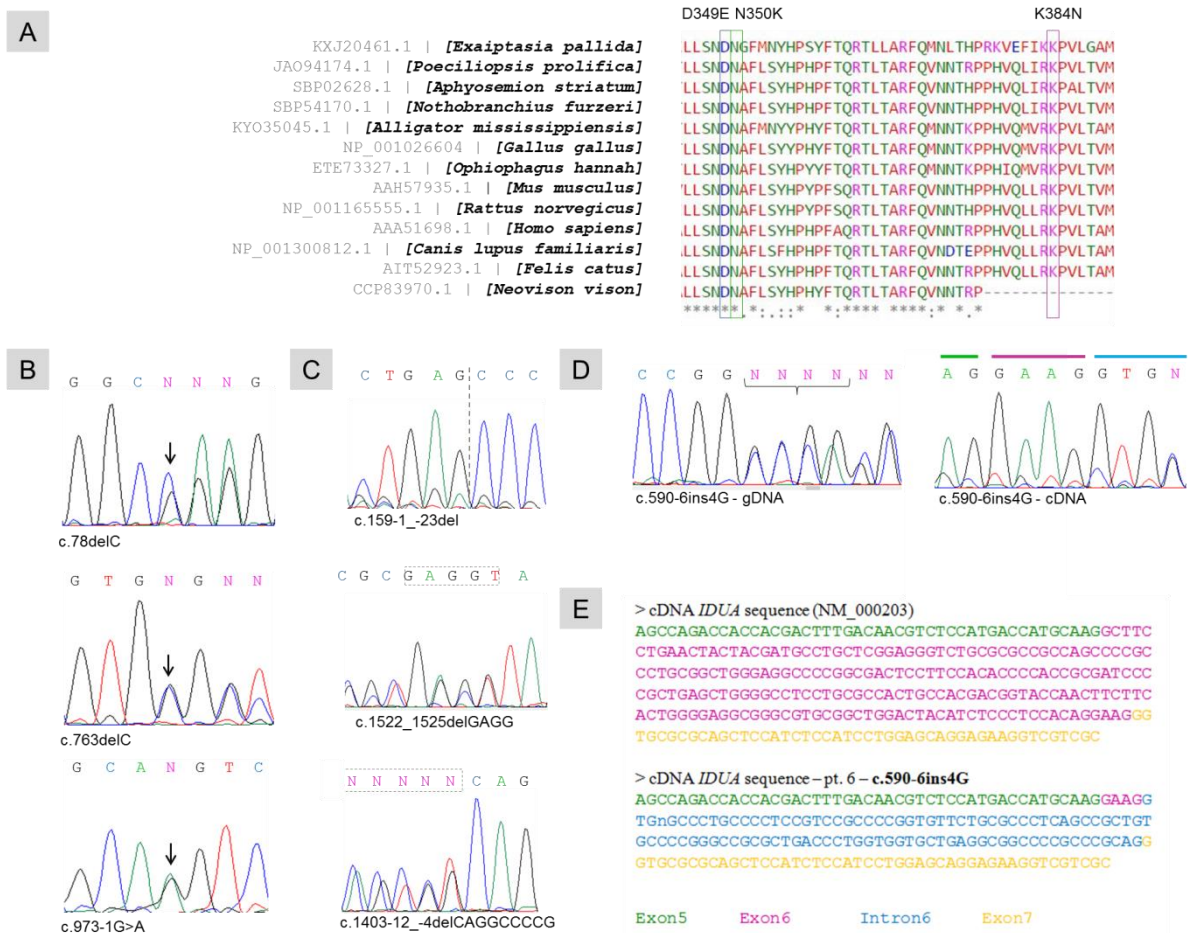


Figure 1. Analysis of novel mutations identified in Brazilian patients. **A** – Multiple alignment of IDUA amino acid (AA) sequences from different species made with ClustalO. Positions of the novel missense mutations are highlighted with coloured boxes; *stands for conserved AA across all 13 species analysed. **B** – Part of electropherograms of gDNA sequencing of *IDUA* gene showing single base deletions or substitutions (c.78delC, c.763delC, c.973-1G>A, all heterozygous), indicated by black arrows; **C** – Part of electropherograms of gDNA sequencing of *IDUA* gene showing small deletions (homozygous c.159-1_-23del, heterozygous c.1522_1525delGAGG and heterozygous c.1403-12_-4delCAGGCCCG); dashed line indicates the 23bp deletion site and dashed box highlights the deleted nucleotides. **D** – Part of electropherograms of gDNA and cDNA sequencing of the novel mutation c.590-6ins4G. Curly bracket indicates the inserted nucleotides; continuous lines above nucleotides represent exons and introns, coloured according to the sequences represented on E; **E** – partial representation of normal *IDUA* cDNA (NM_000203) and mutated sequences in patient with c.590-6ins4G.

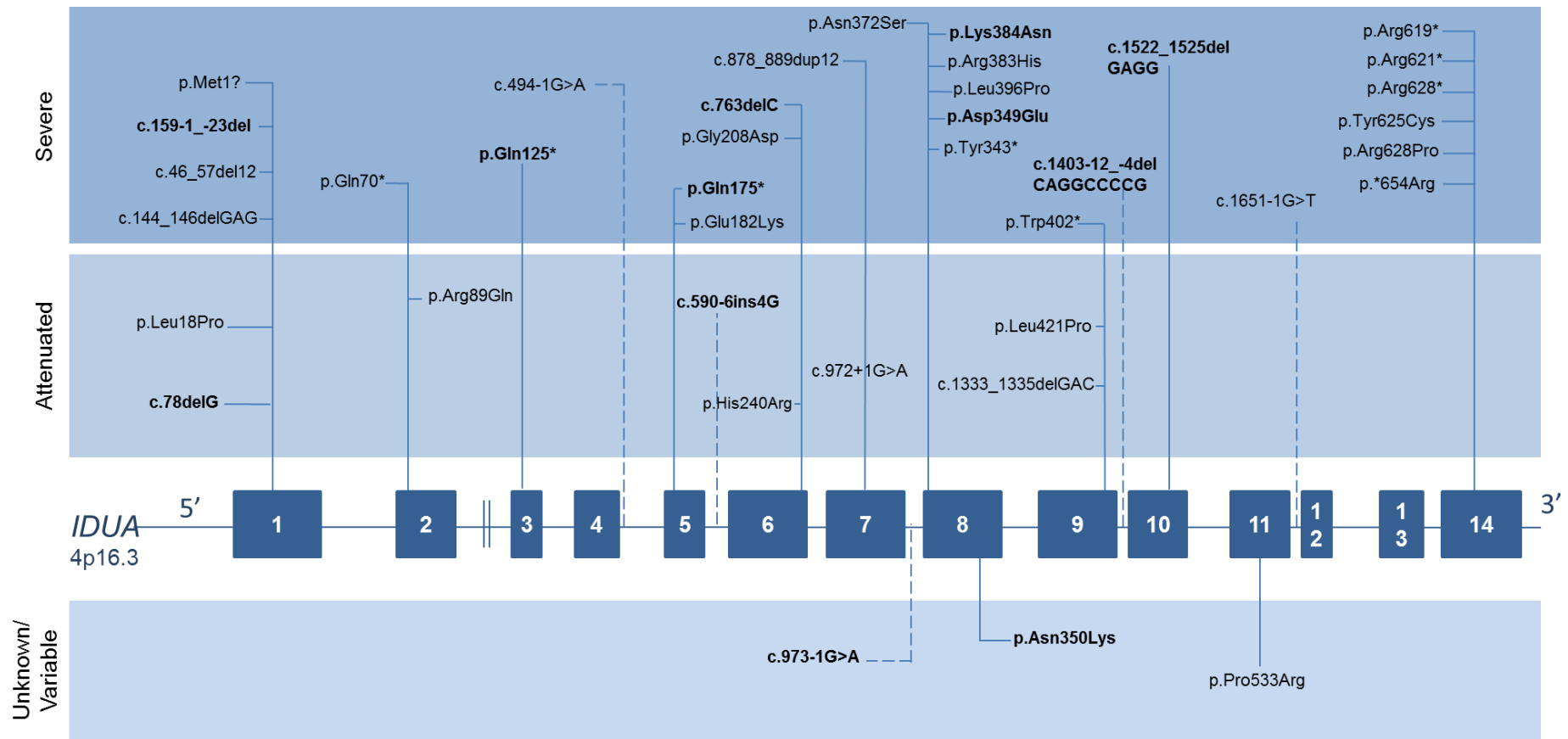


Figure 2. Mutations identified in Brazilian patients and correlation with phenotype. All mutations observed were classified in Severe or Attenuated when data was available in the literature regarding the correlation genotype-phenotype. Known mutations with variable phenotype or novel mutations with insufficient data were classified as Variable/Unknown. Novel mutations are in bold. Continuous lines indicate mutations in exons, while dashed lines represent intronic mutations.

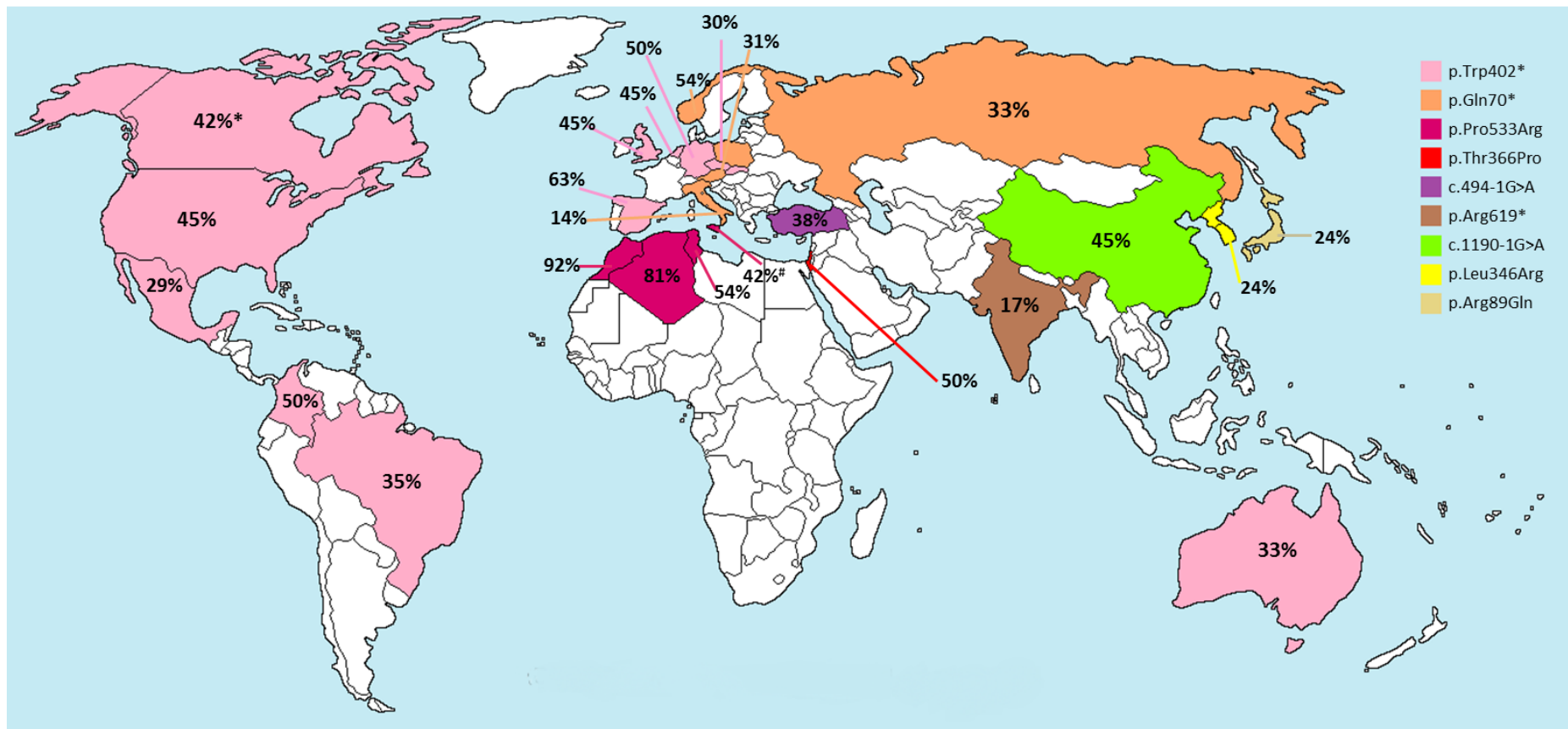


Figure 3. Most frequent *IDUA* mutations among MPS I patients worldwide. P.Trp402* is the most frequent mutation in Australia, Europe, South and North America; p.Gln70* is very frequent in Scandinavia, but also in Russia, Poland and Austria. p.Gln70* is also the most frequent mutation in Italy, although the region of Sicily has distinct frequencies, with p.Pro533Arg as the major allele. p.Pro533Arg is very frequent in Morocco, Algeria and Tunisia. The most frequent alleles in Turkey, Israel, India, China, Korea and Japan are, respectively, c.494-1G>T, p.Thr366Pro, p.Arg619*, c.1190-1G>A, p.Leu346Arg and p.Arg89Gln. *42% is the frequency of p.Trp402* in a population not specified of North Americans, from a Canadian study, but does not necessarily represent the frequency in Canada. #42% of p.Pro533Arg in Sicily was calculated from Gatti et al, 1997 (Gatti, et al., 1997). White-coloured countries: less than 10 alleles analysed or no available data.

Table 1. Mutations identified in Brazilian population from 2000 to 2016 with corresponding allelic frequencies

Location	Nucleotide change*	Effect on protein [§]	Alleles	%	Reference
MISSENSE			75	40%	
exon 1	c.1A>G	p.Met1?	2	1%	(Uttarilli, et al., 2016)
exon 1	c.53T>C	p.Leu18Pro	4	2%	(Pasqualim, et al., 2015)
exon 2	c.266G>A	p.Arg89Gln	4	2%	(Scott, et al., 1995)
exon 5	c.544G>A	p.Glu182Lys	1	0,5%	(Brooks, et al., 2001)
exon 6	c.623G>A	p.Gly208Asp	11	6%	(Li, et al., 2002)
exon 6	c.719A>G	p.His240Arg	1	0,5%	(Beesley, et al., 2001)
exon 8	c.1047C>A	p.Asp349Glu	1	0,5%	This report
exon 8	c.1050 T>G	p.Asn350Lys	1	1%	This report
exon 8	c.1115A>G	p.Asn372Ser	1	0,5%	(Uttarilli, et al., 2016)
exon 8	c.1148G>A	p.Arg383His	6	3%	(Bunge, et al., 1995)
exon 8	c.1152G>C	p.Lys384Asn	1	0,5%	This report
exon 8	c.1187T>C	p.Leu396Pro	1	1%	(Bertola, et al., 2011)
exon 9	c.1262T>C	p.Leu421Pro	1	0,5%	(Wang, et al., 2012)
exon 11	c.1598C>G	p.Pro533Arg	32	17%	(Scott, et al., 1992c)
exon 14	c.1874A>G	p.Tyr625Cys	6	3%	(Fernandez-Marmiesse, et al., 2014)
exon 14	c.1883G>C	p.Arg628Pro	1	0,5%	(Matte, et al., 2003)
exon 14	c.1960T>C	p.*654Argext*62	1	0,5%	(Bertola, et al., 2011)
NONSENSE			87	47%	
exon 2	c.208C>T	p.Gln70*	6	3%	(Clarke and Scott, 1993)
exon 3	c.375G>A	p.Gln125*	2	1%	This report
exon 5	c.524G>A	p.Trp175*	1	0,5%	This report
exon 8	c.1029C>G	p.Tyr343*	1	0,5%	(Tieu, et al., 1994)
exon 9	c.1205G>A	p.Trp402*	71	38%	(Scott, et al., 1992b)
exon 14	c.1855C>T	p.Arg619*	2	1%	(Beesley, et al., 2001)
exon 14	c.1861C>T	p.Arg621*	3	2%	(Bunge, et al., 1994)
exon 14	c.1882C>T	p.Arg628*	1	0,5%	(Beesley, et al., 2001)
SPLICING			9	5%	
intron 1	c.159-1_-23del		2	1%	This report
intron 4	c.494-1G>A		2	1%	(Bertola, et al., 2011)
intron 5	c.590-6ins4G		1	0,5%	This report
intron 7	c.972+1G>A		1	0,5%	(Kwak, et al., 2016)
intron 7	c.973-1G>A		1	1%	This report
intron 9	c.1190-12_-4delCAGGCCCG		1	0,5%	This report
intron 11	c.1651-1G>T		1	0,5%	(Matte, et al., 2003)
SMALL DELETIONS AND DUPLICATIONS			13	7%	
exon 1	c.46_57del12	p.Ser16_Ala19del	2	1%	(Bunge, et al., 1994)
exon 1	c.78delG		1	0,5%	This report
exon 1	c.144_146delGAG	p.Arg48del	6	3%	(Matte, et al., 2003)
exon 6	c.763delC	p.Arg255Glyfs*62	1	0,5%	This report
exon 7	c.878_889dup12	p.Thr293_Tyr296dup	1	0,5%	(Bunge, et al., 1994)
exon 9	c.1333_1335delGAC	p.Asp445del	1	0,5%	(Bunge, et al., 1994)
exon 10	c.1522_1525delGAGG	p.Glu508Thrfs*16	1	0,5%	This report
	Not identified		2	1,1%	
Total			184	100%	

* Nomenclature according to HGVS guidelines, using the reference sequence NM_000203; [§] nomenclature using reference sequence NP_000194;

Table 2. Clinical data of MPS I patients with novel mutations

P t.	Age at diagnosis	Genotype	Phenotype	GAGs [§] µg/mg creatinine (ref) [†]	IDUA activity nmoles/h/ mg in leucocytes (ref) ^{* †}	ERT	Age at beginning of ERT	Hepatomegaly	Esplenomegaly	Growth delay	Coarse facial features	Corneal clouding	Cardiopathy/Valvulopathy	Hernia	Joint stiffness	Thoracic gibbus	Motor delay	Neurologic regression	Hearing loss	Respiratory difficulties	Claw hands	Macrocephaly/Hydrocephaly	Dysostosis multiplex	Dysmorphias	Tunnel carpal syndrome	Seizures	Hypotonia
1		c.78delC p.Arg383His	attenuated			yes											-	-									
2	3 yo	c.159-1_-23del c.159-1_-23del	severe	450 (68-186)	0,03 (32-52)			+	+		+						+			+	+	+					
3		p.(Gln125*) p.(Gln125*)	severe	631 (68-186)	2,7 (27-171)						+			+			+								+	+	
4	10 mo	p.(Trp175*) p.Trp402*	severe	996 (133-274)	0,8 (27-171)			+			+		+			+						+	+				
5	20 yo	c.590-6ins4G p.Tyr343*	attenuated	10 (1,5-5)	0,16 (32-56)	yes	21 yo	+	-	-	+	+	+	+	+	-	-	-	-	+	+	-	-	-	+	-	-
6	2 yo	c.763delC p.Trp402*	severe	604 (68-186)	0,7 (32-52)			+	+	+	+	+	+	+			+		+	+	+						
7	2 yo	c.973-1G>A p.Tyr625Cys	N.A.	-	0,11 (32-56)																						
8	8 mo	p.(Asp349Glu) p.Trp402*	severe	552 (79-256)	0,25 (32-52)	yes	10 mo	+	+		+	+	+	+	+	+					+	+					
9	20 mo	p.(Asn350Lys) p.Gly208Asp	N.A.																								
10	3 yo	p.(Lys384Asn) p.Pro533Arg	severe	212 (13-59)	0,1 (32-52)	yes	7 yo	+	-	-	+	+	+	+	+	+	+	-	+	+	+	-	+	+	-	+	-
11	4 yo	c.1403-12_-4del CAGGCCCG p.Arg621*	severe	417 (67-124)	0,05 (32-52)	yes	4 yo	+	+	+	+	+	+	+	+	+	+	-	-	+	+	-	+	+	-	-	-
12	9 yo	c.1522_1525delGAGG p.Pro533Arg	severe	975 (26-97)	0,13 (32-56)	yes		+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+				

[§]GAG eletrophoresis for dermatan sulfate and heparan sulfate was positive for all assessed patients; [†]Data collected prior to treatment; ^{*}All assays were controlled by analysis of reference enzymes

activity, with values within normal range; Pt.: patient; ref: Reference values; ERT: Enzyme Replacement Therapy; Bold formatting for novel mutations;

Table 3. Molecular characterization and bioinformatic analysis of 14 novel mutations

DNA (NM_000203)	Protein (NP_000194)	Codon	Type	Mutation pathogenicity prediction *							Top 5 features from MutPred
				Polyphen2	PROVEAN	SIFT	PhD-SNP	PANTHER	SNPs&GO	MutPred	
c.1047C>A	p.(Asp349Glu)	gaC>gaA	missense	Probably damaging	Deleterious	Damaging	Disease	Neutral	Neutral	Disease	Loss of loop (P = 0.1242), Loss of sheet (P = 0.1501), Loss of helix (P = 0.1706), Gain of disorder (P = 0.2351), Loss of catalytic residue at A351 (P = 0.3378)
c.1050 T>G	p.(Asn350Lys)	aaT>aaG	missense	Probably damaging	Deleterious	Damaging	Disease	Disease	Disease	Disease	Gain of ubiquitination at N350 (P = 0.0177), Loss of sheet (P = 0.0181), Gain of methylation at N350 (P = 0.0231), Loss of helix (P = 0.079), Loss of loop (P = 0.0804), Loss of methylation at K384 (P = 0.0091), Loss of sheet (P = 0.0817), Gain of catalytic residue at R383 (P = 0.203), Loss of MoRF binding (P = 0.2269), Gain of loop (P = 0.2754)
c.1152G>C	p.(Lys384Asn)	aaG>aaC	missense	Probably damaging	Deleterious	Damaging	Neutral	Disease	Neutral	Disease	
				Nonsense, deletions and insertions							
				VEP	Human Splice Finder			NetGene2			
c.78delC	p.(E27Rfs81*)		deletion	frameshift variant							
c.159-1_-23del			splice site								Activation of cryptic acceptor and donor splice sites upstream the normal, resulting in inclusion of part of intron 1 and loss of part of exon 2.
c.375G>A	p.(Gln125*)	caG>aaA	nonsense	stop gained							
c.524G>A	p.(Trp175*)	tGg>tAg	nonsense	stop gained							
c.590-6ins4G			splice site								Activation of cryptic acceptor and donor splice sites downstream the normal, with loss of most of exon 6 and inclusion of intron 6.
c.763delC	p.(Arg255Glyfs*62)		deletion	frameshift variant							
c.973-1G>A			splice site	splice acceptor variant			Alteration of the WT acceptor site and activation of an intronic cryptic acceptor site. Possible alteration of splice site.				Loss of acceptor site, resulting in loss of exon 8
c.1403-12_-4del CAGGCCCG			splice site								Loss of acceptor site, resulting in loss of exon 10
c.1522_1525del GAGG	p.(Glu508Thrfs*16)		deletion								

* All predictions were made using transcript sequence ENST00000247933.

Table 4. Most frequent IDUA mutations among MPS I patients worldwide

	North America #	United States	Mexico	Colombia	Brazil	Netherlands	United Kingdom	Norway	Austria	Czech Republic and Slovakia	Germany	Poland	Spain	Italy	Morocco	Algeria	Tunisia	Israel	Turkey	Russia	India	China	Korea	Japan	Australia
c.1190-1G>A																						9/20 ²⁹			
c.144_146del GAG				2/24 ⁶	7/24 ² 8,10																				
c.386-2A>G	2/38 ¹						2/28 ¹²				1/10 ¹⁵														
c.46_57del12			1/14 ⁵		3/24 ² 8,10		2/28 ¹²	2/28 ¹¹				2/62 ¹⁶		3/60 ¹⁹						4/20 ¹⁶		2/114 ³⁰			1/20 ³⁴
c.494-1G>A					2/184 ¹⁰															15/40 ^{16,26}					
c.613_617dup TGCTC																							6/34 ^{31,32}	7/38 ³⁶	
p.Ala327Pro				2/24 ⁶	6/106 ^{7,8}		6/170 ¹²		5/46 ¹³	2/40 ¹⁴	1/10 ¹⁵	1/62 ¹⁶	1/24 ¹⁶	7/188 ^{16,18,19}						1/20 ¹⁶					1/20 ³⁴
p.Ala75Thr	2/38 ¹						3/170 ¹²	1/28 ¹¹													7/60 ²⁸				
p.Ala79Val																						19/114 ³⁰	1/4 ³²		1/20 ³⁴
p.Arg619*					2/184 ¹⁰		3/28 ¹²														10/60 ²⁸	1/20 ²⁹			
p.Arg628*		1/44 ²			3/242 ^{8,10}					1/40 ¹⁴				2/60 ¹⁹					4/42 ²¹⁻²³	2/20 ¹⁶				1/4 ³²	
p.Arg89Gln					7/290 ^{7,8,10}		2/170 ¹²							1/54 ¹⁸								2/114 ³⁰		9/38 ³⁶	
p.Gln310*																		2/16 ²⁵							
p.Gln380Arg					1/58 ⁸		1/28 ¹²			4/40 ¹⁴				1/60 ¹⁹							1/12 ²⁷				1/20 ³⁴
p.Gln70*	3/38 ¹	20/6 ²⁻⁴			8/310 ⁷⁻ 10	3/22 ¹¹	27/170 ¹²	15/28 ¹¹	6/11,13 ^{11,13}	7/40 ¹⁴	8/11,15 ^{11,15}	4/11,16 ^{11,16}	7/60 ^{16,17}	8/16,18,19 ^{16,18,19}						3/20 ¹⁶	4/12 ²⁷	1/114 ³⁰			23/166 ^{33,34}
p.Glu178Lys														1/60		2/26						1/114			

Table 5. Allelic frequencies of most common mutations across different populations according to gnomAD

Mutation	European (Non-Finnish)	European (Finnish)	Other	African	Latino	South Asian	East Asian	Ashkenazi Jewish
p.Trp402*	0.001409	0.0003447	0.001224	0	0.0003929	0	0	0
p.Gln70*	0.0007012	0.002022	0.000413	0.0001969	0	0	0	0
p.Pro533Arg	0.00006547	0	0.000248	0	0.0001988	0	0	0
p.Arg619*	0.00002694	0	0.0001573	0	0	0	0	0
c.494-1G>A	0.000008928	0	0	0	0	0	0	0
p.Arg89Gln	0.00001072	0	0	0	0	0.00003522	0.00006416	0
p.Leu346Arg	0	0	0	0	0	0	0.00005775	0

Color gradient: red for most frequent inside the population, green for less frequent.

Supplementary Table 1. Protocols used for PCR amplification of *IDUA* fragments

Exons amplified	Primer	Sequence 5' > 3'	Product size (bp)	Annealing temperature	Reference	MgCl ₂ concentration	Protocol
1	ID1+	ACC CAA CCC CTC CCA C	397	59°C	1	3mM	A
	ID1-	GCT CCG GTC TCT GAA GCT			1		
2	IDUA 2F	GAA CGT GTG TGT CAG CCG	303	60°C	1	2mM	D
	IDUA 2R	ACA AGG GGT CTT CCG AGC			1		
3 and 4	IDUA 3-4F	TTC CAG CCT GGA GCA TGG AG	513*	64°C	1	1,5mM	C
	ID4-	GCG TGA TAG GGG TGC AAC			1		
5 and 6	ID5+	CAT CAC CTT GCA CCC TCC	578*	65°C	1	1,5mM	C
	ID6-	CTC AGC ACC ACC AGG GTC			1		
7 and 8	7-8F	TGC GGC TGG ACT ACA TCT C	730*	56°C	2	1,5mM	B
	7-8R	CTG GAG GAA GTG CGC TCC			2		
9	IDUA 9F	CTG GGG ACT CCT TCA CCA AG	454	54°C	Present study	3mM	A
	IDUA 9R	GAG CCC GTT GTC CAG GTA G			Present study		
9 and 10	IDUA 9F	CTG GGG ACT CCT TCA CCA AG	708*	61°C	Present study	2mM	D
	IDUA 9-10R	TCC TCA GGG TTC TCC AGG G			Present study		
10	IDUA 10F	ATCTACGCGAGCGACGAC	469	60°C	3	2mM	D
	IDUA 9-10R	TCC TCA GGG TTC TCC AGG G			Present study		
11 and 12	IDUA 11-12F	GTG TGG GTG GGA GGT GGA	465*	60°C	1	1,5mM	C
	IDUA 11-12R	GTG ACC GCA TGG GTG AAG			1		
13 and 14	IDUA 13-14F	CTG CCT GCT CCC ACC TTT GAG	529*	63°C	4	1,5mM	C
	IDUA 13-14R	TGA TGG GAG GGC AGC ATG GG			4		
E3 to E7 (cDNA)	CIDUA 2F	GCT TGT CAC CAC CAG GGG G	554	54°C	Present study	1,5mM	E
	CIDUA 2R	GCG ACG ACC TTC TCC TGC TCC			Present study		

* Amplicons include whole intron; A: 10% (v/v) ammonium buffer, 0.2mM dNTP mix, 0.4µM each primer, 1U Taq DNA polymerase and 2ng/µL of DNA; B and C: same as A with 5% and 10% (v/v) of DMSO, respectively; D: 10% (v/v) AmpliTaq Gold Buffer 10x, 0.2mM dNTP mix, 10% (v/v) GC Enhancer, 0.4µM each primer, 0.5U AmpliTaqGold and 1ng/µL of DNA; E: 10% (v/v) Taq DNA Polymerase PCR Buffer 10x, 0.4mM dNTP mix, 10% (v/v) DMSO, 0.4µM each primer, 2U Taq Platinum DNA polymerase, 100ng cDNA in a 50µL volume reaction. References: 1 (Beesley, et al., 2001), 2 (Bertola, et al., 2011), 3 (Li, et al., 2002), 4 (Prommajan, et al., 2011).

References

- Alif N, Hess K, Straczek J, Sebbar S, N'Bou A, Nabet P, Dousset B. 1999. Mucopolysaccharidosis type I: characterization of a common mutation that causes Hurler syndrome in Moroccan subjects. *Ann Hum Genet* 63(Pt 1):9-16.
- Atceken N, Ozgul RK, Yucel Yilmaz D, Tokatli A, Coskun T, Sivri HS, Dursun A, Karaca M. 2016. Evaluation and identification of IDUA gene mutations in Turkish patients with mucopolysaccharidosis type I. *Turk J Med Sci* 46(2):404-8.
- Bach G, Moskowitz SM, Tieu PT, Matynia A, Neufeld EF. 1993. Molecular analysis of Hurler syndrome in Druze and Muslim Arab patients in Israel: multiple allelic mutations of the IDUA gene in a small geographic area. *Am J Hum Genet* 53(2):330-8.
- Beck M, Arn P, Giugliani R, Muenzer J, Okuyama T, Taylor J, Fallet S. 2014. The natural history of MPS I: global perspectives from the MPS I Registry. *Genet Med* 16(10):759-65.
- Beesley CE, Meaney CA, Greenland G, Adams V, Vellodi A, Young EP, Winchester BG. 2001. Mutational analysis of 85 mucopolysaccharidosis type I families: frequency of known mutations, identification of 17 novel mutations and in vitro expression of missense mutations. *Hum Genet* 109(5):503-11.
- Bertola F, Filocamo M, Casati G, Mort M, Rosano C, Tylki-Szymanska A, Tuysuz B, Gabrielli O, Grossi S, Scarpa M and others. 2011. IDUA mutational profiling of a cohort of 102 European patients with mucopolysaccharidosis type I: identification and characterization of 35 novel alpha-L-iduronidase (IDUA) alleles. *Hum Mutat* 32(6):E2189-210.
- Bie H, Yin J, He X, Kermodé AR, Goddard-Borger ED, Withers SG, James MN. 2013. Insights into mucopolysaccharidosis I from the structure and action of alpha-L-iduronidase. *Nat Chem Biol* 9(11):739-45.
- Botigue LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, Atzmon G, Burns E, Ostrer H, Flores C and others. 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc Natl Acad Sci U S A* 110(29):11791-6.
- Brooks DA, Fabrega S, Hein LK, Parkinson EJ, Durand P, Yogalingam G, Matte U, Giugliani R, Dasvarma A, Eslahpazire J and others. 2001. Glycosidase active site mutations in human alpha-L-iduronidase. *Glycobiology* 11(9):741-50.
- Bunge S, Kleijer WJ, Steglich C, Beck M, Schwinger E, Gal A. 1995. Mucopolysaccharidosis type I: identification of 13 novel mutations of the alpha-L-iduronidase gene. *Hum Mutat* 6(1):91-4.
- Bunge S, Kleijer WJ, Steglich C, Beck M, Zuther C, Morris CP, Schwinger E, Hopwood JJ, Scott HS, Gal A. 1994. Mucopolysaccharidosis type I: identification of 8 novel mutations and determination of the frequency of the two common alpha-L-iduronidase mutations (W402X and Q70X) among European patients. *Hum Mol Genet* 3(6):861-6.
- Chistiakov DA, Savost'yanov KV, Kuzenkova LM, Gevorkyan AK, Pushkov AA, Nikitin AG, Pakhomov AV, Vashakmadze ND, Zhurkova NV, Podkletnova TV and others. 2014. Molecular characteristics of patients with glycosaminoglycan storage disorders in Russia. *Clin Chim Acta* 436:112-20.
- Chkioua L, Khedhiri S, Jaidane Z, Ferchichi S, Habib S, Froissart R, Bonnet V, Chaabouni M, Dandana A, Jrad T and others. 2007. [Mucopolysaccharidosis type I: identification of alpha-L-iduronidase mutations in Tunisian families]. *Arch Pediatr* 14(10):1183-9.
- Chkioua L, Khedhiri S, Kassab A, Bibi A, Ferchichi S, Froissart R, Vianey-Saban C, Laradi S, Miled A. 2011a. Molecular analysis of mucopolysaccharidosis type I in Tunisia: identification of novel mutation and eight novel polymorphisms. *Diagn Pathol* 6:39.
- Chkioua L, Khedhiri S, Turkia HB, Tchong R, Froissart R, Chahed H, Ferchichi S, Ben Dridi MF, Vianey-Saban C, Laradi S and others. 2011b. Mucopolysaccharidosis type I: molecular characteristics of two novel alpha-L-iduronidase mutations in Tunisian patients. *Diagn Pathol* 6:47.
- Clarke LA, Nelson PV, Warrington CL, Morris CP, Hopwood JJ, Scott HS. 1994. Mutation analysis of 19 North American mucopolysaccharidosis type I patients: identification of two additional frequent mutations. *Hum Mutat* 3(3):275-82.
- Clarke LA, Scott HS. 1993. Two novel mutations causing mucopolysaccharidosis type I detected by single strand conformational analysis of the alpha-L-iduronidase gene. *Hum Mol Genet* 2(8):1311-2.
- Fernandez-Marmiesse A, Morey M, Pineda M, Eiris J, Couce ML, Castro-Gago M, Fraga JM, Lacerda L, Gouveia S, Perez-Poyato MS and others. 2014. Assessment of a targeted

- resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. *Orphanet J Rare Dis* 9:59.
- Gatti R, DiNatale P, Villani GR, Filocamo M, Muller V, Guo XH, Nelson PV, Scott HS, Hopwood JJ. 1997. Mutations among Italian mucopolysaccharidosis type I patients. *J Inher Metab Dis* 20(6):803-6.
- Gort L, Chabas A, Coll MJ. 1998. Analysis of five mutations in 20 mucopolysaccharidosis type 1 patients: high prevalence of the W402X mutation. *Mutations in brief* no. 121. Online. *Hum Mutat* 11(4):332-3.
- Kroepfl T, Milos I, Paul K, Plecko B, Paschke E. 2001. The frequency of common mutations among patients with mucopolysaccharidosis types I, II and IIIA diagnosed in Austria over the last 17 years. *Clin Genet* 60:393–394.
- Kwak M, Huh R, Kim J, Park H, Cho S, Jin D. 2016. Report of 5 novel mutations of the α -Liduronidase gene and comparison of Korean mutations in relation with those of Japan or China in patients with mucopolysaccharidosis I. *BMC Medical Genetics* 17:58.
- Laradi S, Tukul T, Erazo M, Shabbeer J, Chkioua L, Khedhiri S, Ferchichi S, Chaabouni M, Miled A, Desnick RJ. 2005. Mucopolysaccharidosis I: Alpha-L-Iduronidase mutations in three Tunisian families. *J Inher Metab Dis* 28(6):1019-26.
- Lee IJ, Hwang SH, Jeon BH, Song SM, Kim JS, Paik KH, Kwon EK, Jin DK. 2004. Mutational analysis of the alpha-L-iduronidase gene in 10 unrelated Korean type I mucopolysaccharidosis patients: Identification of four novel mutations. *Clin Genet* 66(6):575-6.
- Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB and others. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536(7616):285-91.
- Li P, Wood T, Thompson JN. 2002. Diversity of mutations and distribution of single nucleotide polymorphic alleles in the human alpha-L-iduronidase (IDUA) gene. *Genet Med* 4(6):420-6.
- Matte U, Leistner S, Lima L, Schwartz I, Giugliani R. 2000. Unique frequency of known mutations in Brazilian MPS I patients. *Am J Med Genet* 90(2):108-9.
- Matte U, Yogalingam G, Brooks D, Leistner S, Schwartz I, Lima L, Norato DY, Brum JM, Beesley C, Winchester B and others. 2003. Identification and characterization of 13 new mutations in mucopolysaccharidosis type I patients. *Mol Genet Metab* 78(1):37-43.
- Neufeld EF, Muenzer J. 2001. The mucopolysaccharidosis. In: CR BA, Sly WS and Valle D editor. *The Metabolic and Molecular Bases of Inherited Disease*. New York: McGraw-Hill Co. p 3421-3452.
- Pasqualim G, Ribeiro MG, da Fonseca GG, Szlago M, Schenone A, Lemes A, Rojas MV, Matte U, Giugliani R. 2015. p.L18P: a novel IDUA mutation that causes a distinct attenuated phenotype in mucopolysaccharidosis type I patients. *Clin Genet* 88(4):376-80.
- Pereira VG, Martins AM, Micheletti C, D'Almeida V. 2008. Mutational and oxidative stress analysis in patients with mucopolysaccharidosis type I undergoing enzyme replacement therapy. *Clin Chim Acta* 387(1-2):75-9.
- Pineda T, Marie S, Gonzalez J, Garcia AL, Acosta A, Morales M, Correa LN, Vivas R, Escobar X, Protzel A and others. 2014. Genotypic and bioinformatic evaluation of the alpha-L-iduronidase gene and protein in patients with mucopolysaccharidosis type I from Colombia, Ecuador and Peru. *Mol Genet Metab Rep* 1:468-473.
- Pollard LM, Jones JR, Wood TC. 2013. Molecular characterization of 355 mucopolysaccharidosis patients reveals 104 novel mutations. *J Inher Metab Dis* 36(2):179-87.
- Scott HS, Bunge S, Gal A, Clarke LA, Morris CP, Hopwood JJ. 1995. Molecular genetics of mucopolysaccharidosis type I: diagnostic, clinical, and biological implications. *Hum Mutat* 6(4):288-302.
- Scott HS, Litjens T, Hopwood JJ, Morris CP. 1992a. A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype. *Hum Mutat* 1(2):103-8.
- Scott HS, Litjens T, Nelson PV, Brooks DA, Hopwood JJ, Morris CP. 1992b. alpha-L-iduronidase mutations (Q70X and P533R) associate with a severe Hurler phenotype. *Hum Mutat* 1(4):333-9.
- Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. 2014. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 133(1):1-9.

- Suárez A, Sánchez L. 2014. Mutaciones genéticas y su relación con el fenotipo clínico en pacientes con mucopolisacaridosis de tipo I en el noreste de México. *Gaceta Médica de México* 150:289-96.
- Sun L, Li C, Song X, Zheng N, Zhang H, Dong G. 2011. Three novel alpha-L-iduronidase mutations in 10 unrelated Chinese mucopolysaccharidosis type I families. *Genet Mol Biol* 34(2):195-200.
- Tebani A, Zanutene-Cheriet L, Adjtoutah Z, Abily-Donval L, Brasse-Lagnel C, Laquerriere A, Marret S, Chalabi Benabdellah A, Bekri S. 2016. Clinical and Molecular Characterization of Patients with Mucopolysaccharidosis Type I in an Algerian Series. *Int J Mol Sci* 17(5).
- Terlato NJ, Cox GF. 2003. Can mucopolysaccharidosis type I disease severity be predicted based on a patient's genotype? A comprehensive review of the literature. *Genet Med* 5(4):286-94.
- Tieu PT, Menon K, Neufeld EF. 1994. A mutant stop codon (TAG) in the IDUA gene is used as an acceptor splice site in a patient with Hurler syndrome (MPS IH). *Hum Mutat* 3(3):333-6.
- Uttarilli A, Ranganath P, Matta D, Md Nurul Jain J, Prasad K, Babu AS, Girisha KM, Verma IC, Phadke SR, Mandal K and others. 2016. Identification and characterization of 20 novel pathogenic variants in 60 unrelated Indian patients with mucopolysaccharidoses type I and type II. *Clin Genet* 90(6):496-508.
- Vazna A, Beesley C, Berna L, Stolnaja L, Myskova H, Bouckova M, Vlaskova H, Poupetova H, Zeman J, Magner M and others. 2009. Mucopolysaccharidosis type I in 21 Czech and Slovak patients: mutation analysis suggests a functional importance of C-terminus of the IDUA protein. *Am J Med Genet A* 149A(5):965-74.
- Venturi N, Rovelli A, Parini R, Menni F, Brambillasca F, Bertagnolio F, Uziel G, Gatti R, Filocamo M, Donati MA and others. 2002. Molecular analysis of 30 mucopolysaccharidosis type I patients: evaluation of the mutational spectrum in Italian population and identification of 13 novel mutations. *Hum Mutat* 20(3):231.
- Wang X, Zhang W, Shi H, Qiu Z, Meng Y, Yao F, Wei M. 2012. Mucopolysaccharidosis I mutations in Chinese patients: identification of 27 novel mutations and 6 cases involving prenatal diagnosis. *Clin Genet* 81(5):443-52.
- Yogalingam G, Guo XH, Muller VJ, Brooks DA, Clements PR, Kakkis ED, Hopwood JJ. 2004. Identification and molecular characterization of alpha-L-iduronidase mutations present in mucopolysaccharidosis type I patients undergoing enzyme replacement therapy. *Hum Mutat* 24(3):199-207.

DISCUSSÃO

As mucopolissacaridoses são doenças raras, com diagnóstico relativamente tardio, normalmente depois de os principais e mais graves sintomas surgirem. Considerando que os tratamentos disponíveis devem ser iniciados preferencialmente antes das manifestações clínicas da doença, o diagnóstico deve ser rápido e eficiente. O diagnóstico molecular, além de ser útil em análises pré-natais – o que torna o tratamento possível nos primeiros dias de vida – ainda pode dar uma estimativa de prognóstico, vantagem em relação ao diagnóstico bioquímico. A análise molecular de diferentes populações também é importante, uma vez que identifica as principais mutações e acelera o diagnóstico de novos casos.

Este estudo compreende a maior amostra descrita para análise molecular do gene *IDUA* em pacientes com MPS I – são 92 amostras coletadas nos últimos 16 anos. Todas foram analisadas por sequenciamento de Sanger e a mutação patogênica não foi encontrada em apenas 2 alelos (1%), possivelmente por serem grandes rearranjos não identificáveis com a metodologia empregada. Trinta e nove mutações foram identificadas nos 184 alelos restantes, sendo que 12 destas são mutações novas, correspondendo a aproximadamente 7% dos alelos. Este dado mostra a grande heterogeneidade alélica deste gene, que já possui mais de 200 mutações descritas, com novas mutações constantemente sendo identificadas.

Todas as mutações novas de sentido trocado foram analisadas por *softwares* de predição de patogenicidade e consideradas como possivelmente patogênicas. A análise da conservação dos aminoácidos em diferentes espécies nos mostra que estas mutações ocorrem em posições altamente conservadas, inclusive com outras mutações patogênicas já descritas, acometendo aminoácidos possivelmente relacionados com funções importantes da proteína – como o sítio de ligação ao substrato (Asp349) (Bie, et al., 2013). Além disso, a substituição dos aminoácidos observada nestas mutações novas é por aminoácidos de cargas diferentes, o que afeta o processamento e a conformação

da proteína (Bie, et al., 2013). Somadas a estes dados, a baixa atividade enzimática de IDUA, a ausência destas alterações em indivíduos não afetados (do gnomAD) e as manifestações clínicas encontradas nos pacientes corroboram a sugestão de causalidade das mutações novas com a ocorrência da MPS I.

Duas mutações novas sem sentido, p.Gln125* e p.Trp175*, foram identificadas em pacientes com a forma grave da doença, reforçando a relação genótipo-fenótipo de mutações deste tipo (Terlato and Cox, 2003), embora análises adicionais às *in silico* não tenham sido feitas.

Quatro mutações novas de *splicing* foram identificadas pelos softwares de predição: c.973-1G>A e c.1403-12_-4delCAGGCCCCG, que perdem o sítio acceptor normal resultando na perda dos éxons 8 e 10, respectivamente, e c.159-1_-23del, que inclui parte do íntron 1 e perde parte do éxon 2 por alterações nos sítios acceptor e doador. Todas essas mutações foram encontradas em pacientes com a forma grave da doença ou sem dados clínicos suficientes.

A quarta mutação de *splicing*, c.590-6ins4G, foi analisada *in silico* e a alteração predita pelo software pôde ser confirmada no sequenciamento do cDNA do paciente – ocorre a substituição do éxon 6 (com exceção de 4 bases), pelo íntron 6 no mRNA final. Essa alteração gera uma enzima deficiente, porém com ação residual, uma vez que foi encontrada em um paciente com a forma leve da doença e que possui a mutação sem sentido p.Tyr343* no outro alelo. Já as demais mutações de *splicing* foram encontradas em pacientes com a forma grave da doença ou sem dados clínicos suficientes.

Neste estudo, a mutação p.Trp402* foi a mais frequentemente encontrada, representando 38% dos alelos, seguida da p.Pro533Arg, com 17%. Estudos anteriores no Brasil também concluíram que estas são as duas mutações mais frequentes na população brasileira, embora tenham observado frequências variadas, inclusive uma amostra com maior frequência da p.Pro533Arg, como apresentado na Tabela 2. Por outro lado, a mutação p.Gln70*, muito frequente em alguns países europeus, não possui uma frequência alta no Brasil (Matte, et al., 2000; Matte, et al., 2003; Pereira, et al., 2008).

Tabela 2. Frequências no Brasil das mutações mais frequentes no mundo

	Matte 2000	Matte 2003	Pereira 2008	Presente trabalho	Frequências combinadas
p.Trp402*	10/48 (21%)	22/58 (38%)	5/20 (25%)	71/184 (38%)	108/312 (35%)
p.Pro533Arg	9/48 (19%)	9/58 (16%)	8/20 (40%)	32/184 (17%)	58/312 (18,5%)
p.Gln70*	1/48 (2%)	1/58 (2%)	0/20 (0%)	6/184 (3%)	8/312 (2,5%)

Número de alelos encontrados / número de alelos totais (frequência)

A mutação p.Trp402* é a mais frequente nos pacientes com MPS I no mundo, concentrada principalmente em países europeus, como Espanha (63%), Alemanha (50%), Reino Unido (45%), Holanda (45%) e República Tcheca e Eslováquia (30%). Além disso, países de colonização predominantemente europeia herdaram este alelo, sendo o mais frequente nos Estados Unidos (45%), México (29%), Colômbia (50%), Brasil (35%) e Austrália (33%). Esta mutação também já foi descrita em pacientes da Turquia, sendo a segunda mais frequente, com 20% dos alelos. Não há registros da ocorrência desta mutação em pacientes do extremo oriente, como China, Coréia e Japão. Da mesma forma, corroborando com o observado em pacientes, a análise do gene *IDUA* em indivíduos sem doenças genéticas e agrupados por etnias no gnomAD (europeus, europeus finlandeses, latinos, africanos, sul-asiáticos, leste-asiáticos, judeus Ashkenazi e Outros (que incluem todos os outros grupos étnicos não citados) mostra que o alelo p.Trp402* é muito frequente em europeus, europeus finlandeses, latinos e outros grupos étnicos, mas está ausente de populações africanas e asiáticas, que juntas somam 42.656 alelos analisados. Ao analisarmos a distribuição do alelo pelo território europeu (Figura 3), pode-se sugerir que a mutação teve origem na Europa central ou ocidental, possivelmente com os povos celtas que povoaram a região, e espalhou-se pelas Américas e Austrália com as colonizações europeias a partir do século XV (Atceken, et al., 2016; Beesley, et al., 2001; Bertola, et al., 2011; Bunge, et al., 1995; Bunge, et al., 1994; Clarke, et al., 1994; Gatti, et al., 1997; Gort, et al., 1998; Lek, et al., 2016; Li, et al., 2002; Pineda, et al., 2014;

Pollard, et al., 2013; Suárez and Sánchez, 2014; Vazna, et al., 2009; Venturi, et al., 2002; Yogalingam, et al., 2004).

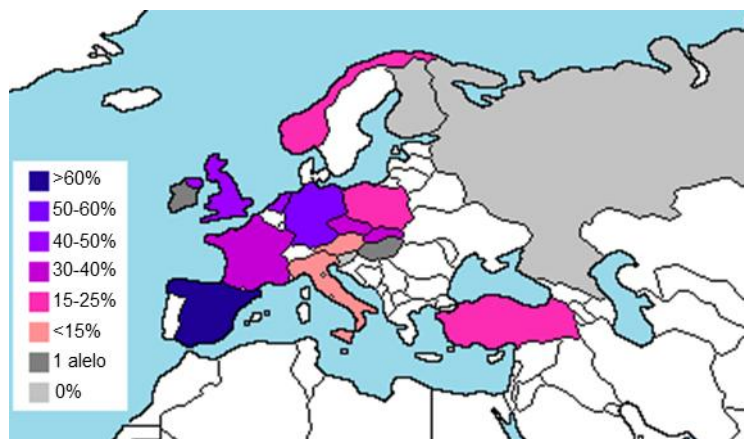


Figura 3. Distribuição de p.Trp402* em pacientes com MPS I na Europa.

Por outro lado, a mutação p.Pro533Arg está mais concentrada nas mediações do mar Mediterrâneo. Países do norte da África possuem os maiores registros de p.Pro533Arg, correspondendo a 92% dos alelos no Marrocos, 81% na Argélia e 54% na Tunísia. Um paciente homocigoto para a mutação também foi identificado no Catar, porém possuía descendência mauritana. Esta frequência muito alta na região pode ser explicada pela alta ocorrência de casamentos consanguíneos (85% na Argélia), que acabam selecionando apenas um alelo. Países do sul europeu também apresentam relativamente alta frequência de p.Pro533Arg comparando com países mais ao norte, correspondendo a aproximadamente 10% dos alelos em ambas Espanha e Itália, chegando a 42% na ilha italiana de Sicília (Figura 4). Em outros continentes, as frequências são de 20% no México, 18,5% no Brasil e 3% na Austrália. Além destes, a mutação também foi descrita em poucos casos no Reino Unido e nos Estados Unidos (Alif, et al., 1999; Beesley, et al., 2001; Bertola, et al., 2011; Chkioua, et al., 2011a; Chkioua, et al., 2011b; Fahiminiya, et al., 2014; Gatti, et al., 1997; Gort, et al., 1998; Laradi, et al., 2005; Li, et al., 2002; Suárez and Sánchez, 2014; Tebani, et al., 2016; Venturi, et al., 2002; Yogalingam, et al., 2004)



Figura 4. Distribuição de p.Pro533Arg em pacientes com MPS I na Europa

Entre indivíduos sem doenças genéticas, a mutação p.Pro533Arg foi observada somente em europeus não-finlandeses, latinos e outros grupos étnicos; curiosamente, nenhum alelo foi identificado em indivíduos africanos, possivelmente porque a amostra que compõe o gnomAD é majoritariamente de africanos subsaarianos. Analisados em conjunto, estes dados podem sugerir que a origem da mutação se deu na região norte-africana – relativamente isolada do restante da África devido à barreira geográfica imposta pelo deserto do Saara – e migrou para países do sul da Europa, acompanhando as invasões mouras na Península Ibérica e na Sicília, que ocorreram entre os séculos VIII e XV. Embora não haja estudos moleculares publicados de MPS I em Portugal, é muito provável que pacientes portugueses apresentem um perfil mutacional no gene *IDUA* similar ao espanhol e que, assim, sejam os principais responsáveis pela alta frequência de p.Pro533Arg observada no Brasil. Além dos portugueses, imigrantes italianos e espanhóis dos séculos XIX e XX também podem ter contribuído significativamente (Botigue, et al., 2013; Lek, et al., 2016; Pena, et al., 2011; Regueiro, et al., 2015).

A segunda mutação mais frequente em pacientes com MPS I no mundo é a p.Gln70*, correspondendo à principal mutação na Finlândia (88%), Noruega (54%), Rússia (33%), Polónia (31%), Áustria (30%) e Itália (14%). No Reino Unido, Estados Unidos e Austrália, a frequência desta mutação também é relativamente alta – 16%, 32% e 14%, respectivamente – sendo a segunda mutação mais frequente nestes países. Além destes, a p.Gln70* também foi

descrita em um paciente na China, um na ilha de Guernsey e em alguns casos no Brasil, porém representando somente 2,5% dos alelos nos pacientes brasileiros (Beesley, et al., 2001; Bertola, et al., 2011; Bunge, et al., 1995; Bunge, et al., 1994; Chistiakov, et al., 2014; Clarke, et al., 1994; Gatti, et al., 1997; Kroepfl, et al., 2001; Scott, et al., 1992c; Venturi, et al., 2002; Wang, et al., 2012; Yogalingam, et al., 2004).

A frequência de p.Gln70* nos pacientes reflete o observado em indivíduos normais – em europeus finlandeses, a frequência deste alelo chega a 0,2%. Ele também é observado em menores proporções em europeus não-finlandeses, africanos e outros grupos étnicos, enquanto está ausente em latinos e asiáticos. A distribuição deste alelo na Europa segue um gradiente, sendo muito frequente na região Fino-Escandinava e diminuindo a frequência mais a sul (Figura 5), situação já observada com alelos de outros genes. Com base nesta observação, pode-se sugerir que a mutação p.Gln70* teve origem nos povos vikings que habitavam a Escandinávia e que migraram para diversas regiões da Europa durante a Era Viking, entre os séculos VIII e X. Historicamente conhecidos como bons navegadores, os vikings conquistaram regiões das ilhas da Grã-Bretanha e da Irlanda, que colonizaram ativamente neste período. Também migraram para países bálticos e navegaram pela costa do continente até o mar Mediterrâneo, onde alguns poucos se estabeleceram nos países europeus banhados pelo mar. Além disso, houve movimentos migratórios à leste – com estabelecimento de vikings varangianos em regiões das atuais Finlândia e Rússia – e criação de rotas comerciais que iam até Constantinopla, atual Istambul. Em outros continentes, a mutação pode ter acompanhado a colonização inglesa na América do Norte e na Austrália, resultando nas altas frequências de p.Gln70* observadas nos pacientes destes países. Sob a mesma ótica, a baixa frequência desta mutação no Brasil é explicada pela pouca contribuição de imigrantes norte-europeus, sendo que os casos de p.Gln70* encontrados possivelmente são de pacientes descendentes de populações da Europa Central e Sul, como alemães e italianos (Bowden, et al., 2008; Lucotte, 2001; Milman and Pedersen, 2003).

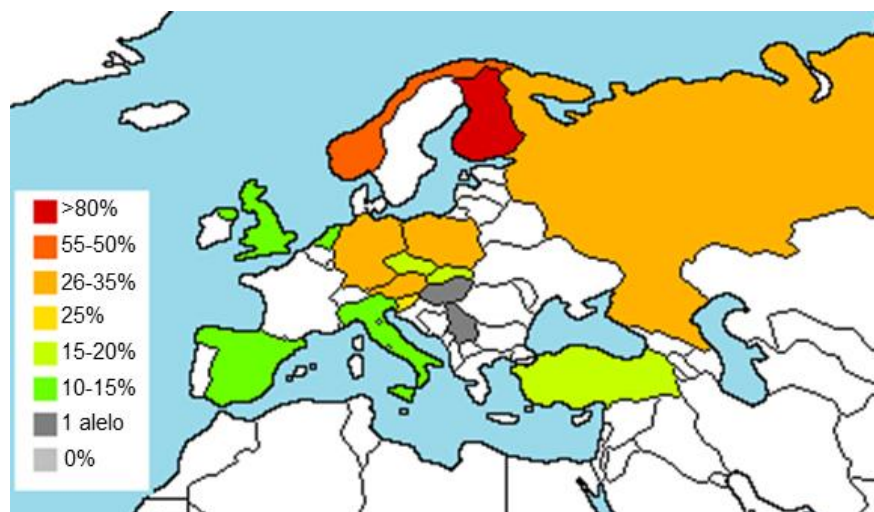


Figura 5. Distribuição de p.Gln70* em pacientes com MPS I na Europa

Algumas mutações mais frequentes no extremo Oriente parecem ser exclusivas da região. A mutação c.613_617dupTGCTC, por exemplo, só pode ser encontrada no Japão (18%) e na Coreia (17%), sendo a segunda mutação mais frequente nestes países, assim como a mutação mais frequente da China, c.1190-1G>A, que só há registros em chineses – com 45% dos alelos de pacientes com MPS I. Por outro lado, a mutação mais frequente do Japão, p.Arg89Gln, está descrita em pacientes da Itália, do Reino Unido e do Brasil, e também pode ser encontrada em indivíduos normais do grupo étnico europeu não-finlandês, além dos indivíduos asiáticos (Kwak, et al., 2016; Wang, et al., 2012; Yamagishi, et al., 1996). Na Coreia, a mutação mais frequente é p.Leu346Arg, correspondendo a 24% dos alelos. Esta mutação também é encontrada na China com uma frequência considerável, sendo responsável por 12% das mutações encontradas nos pacientes com MPS I. A p.Leu346Arg parece ser exclusiva desta região, uma vez que não foi observada em nenhuma outra população, seja de pacientes ou de indivíduos normais (Kwak, et al., 2016; Wang, et al., 2012).

Algumas mutações presentes no leste asiático possuem uma distribuição interessante: p.Ala79Val, muito frequente na China, também é encontrada na Coreia, mas não no Japão. Da mesma maneira, c.613_617dupTGCTC é muito frequente no Japão e na Coreia, mas não na China. Essa particularidade pode

indicar que os três países podem ter surgido de uma mesma população, e que as frequências de cada mutação se diferenciaram dessa maneira por movimentos migratórios e consequente efeito fundador que possa ter ocorrido em cada país (Kwak, et al., 2016).

O perfil mutacional da Índia reflete a colonização britânica na região. As duas mutações mais frequentes no país são p.Arg619* (17%) e p.Ala75Thr (12%), ambas observadas no Reino Unido (11% e 3%, respectivamente); a última está presente, inclusive, na América do Norte (5%), outro local colonizado principalmente por britânicos. Já a mutação p.Leu121Pro (10%) parece ser exclusiva da região, uma vez que não foi descrita em outras populações. Índia e China, apesar da proximidade, não compartilham frequências similares (Uttarilli, et al., 2016).

Turquia e Israel possuem perfis bem variados dos demais apresentados: a Turquia compartilha várias mutações com países europeus, como p.Trp402* e p.Gln70*, além da p.Pro533Arg; embora a mais frequentemente encontrada entre os turcos, c.494-1G>A, só tenha sido observada em brasileiros – com frequências de 38% e 1%, respectivamente. Em indivíduos normais, o alelo foi visto somente em europeus não-finlandeses (Atceken, et al., 2016). As mutações observadas em Israel, p.Thr366Pro e p.Gln310*, com frequências de 50% e 12,5% respectivamente, não foram encontradas em outras populações. A mutação p.Tyr64*, porém, foi encontrada na Turquia (Bach, et al., 1993).

Por fim, analisando-se em conjunto todos os trabalhos publicados com dados moleculares de pacientes brasileiros com MPS I (Matte, et al., 2000; Matte, et al., 2003; Pereira, et al., 2008), as mutações mais frequentes (além de p.Trp402* e p.Pro533Arg, já discutidas) são p.Ala327Pro (6%) e p.Gly208Asp (5%). A primeira é encontrada principalmente na Europa Central, com frequência máxima de 10% na Áustria e Alemanha, diminuindo em países mais a sul e a oeste (4% na Espanha e 3,5% na Itália). Já a mutação p.Gly208Asp foi observada em um paciente chinês e um norte-americano, além dos brasileiros. (Beesley, et al., 2001; Bertola, et al., 2011; Bunge, et al., 1995; Kroepfl, et al., 2001; Pineda, et al., 2014)

Estes dados reforçam que o conhecimento das diferenças moleculares no gene *IDUA* entre as populações facilita, em muito, o diagnóstico molecular de pacientes com MPS I, diminuindo tempo e custo para sua realização. Para a população brasileira, por exemplo, mais de 50% dos genótipos podem ser identificados pela análise de duas mutações apenas (p.Trp402* e p.Pro533Arg, que possuem uma frequência combinada de 54%). De acordo com a amostra analisada neste estudo, se a análise for ampliada para os 7 éxons com maior frequência de mutações (em ordem decrescente: 9, 11, 1, 14, 6, 8 e 2), correspondendo a aproximadamente metade do gene), a identificação dos genótipos passa de 90%.

Além disso, terapias que se baseiam no genótipo do paciente, como moléculas de transleitura de códons de parada prematuros e edição gênica, sofrem grande influência da variação genotípica entre países. A edição gênica da mutação p.Gln70*, por exemplo, pode ser muito eficiente em populações Fino-Escandinavas, podendo tratar mais de 80% dos pacientes na Finlândia; entretanto, no Brasil, apenas alguns poucos pacientes se beneficiariam do tratamento. Da mesma forma, uma terapia baseada em transleitura de códons de parada prematuros no norte da África não terá efeito algum, uma vez que a imensa maioria dos pacientes possui a mutação de sentido trocado p.Pro533Arg – neste caso, a edição gênica específica dessa mutação contemplaria mais de 90% dos pacientes, sendo potencialmente muito eficaz.

Este entendimento sobre as variações genotípicas entre populações e o conhecimento das frequências em cada uma é fundamental não só para o desenvolvimento de novos tratamentos, mas também para a implantação de políticas públicas de saúde, como programas nacionais de triagem neonatal – que para MPS I é de extremo valor, uma vez que o sucesso dos tratamentos disponíveis atualmente depende diretamente da intervenção ágil, preferencialmente nos primeiros meses de vida.

CONCLUSÕES

A partir dos dados apresentados por esta dissertação, pode-se concluir que:

- O sequenciamento da região codificante e junções éxon-íntron do gene *IDUA* é uma maneira eficaz para genotipagem dos pacientes com MPS I, sendo possível identificar aproximadamente 99% dos alelos;

- Foram identificadas 39 mutações nos 92 pacientes analisados, sendo que 12 destas estão sendo descritas pela primeira vez, ilustrando a alta heterogeneidade alélica do gene *IDUA*;

- Todas as mutações novas foram consideradas patogênicas por um ou mais dos critérios a seguir: 1) ausência de outras mutações no gene; 2) ausência da mutação em indivíduos normais; 3) alto escore de patogenicidade em programas de predição; 4) alta conservação do aminoácido alterado em diferentes espécies; 5) relação descrita do aminoácido trocado com a função da proteína e 6) mutação patogênica descrita na mesma posição. Entretanto, ensaios de mutagênese sítio-dirigida seriam interessantes para confirmar as predições;

- A correlação completa genótipo-fenótipo não foi possível pela ausência de dados clínicos de alguns pacientes, mas pôde-se concluir que ao menos duas das mutações novas (c.78delC e c.590-6ins4G) são responsáveis pela forma leve da doença;

- A população brasileira tem um perfil mutacional único, sendo que as mutações mais frequentes são p.Trp402* e p.Pro533Arg, com 35% e 18,5% dos alelos, e a mesma relação não é vista em nenhum outro país;

- A mutação p.Trp402* é frequente em populações europeias, principalmente na Espanha, Alemanha e Reino Unido, e nos países colonizados por eles, como Estados Unidos e Austrália;

- A p.Pro533Arg ocorre principalmente em países mediterrâneos, sendo a frequência desta mutação no Brasil proveniente possivelmente de imigrantes portugueses, espanhóis e italianos;

- Apesar de muito frequente na Europa e América do Norte, a mutação p.Gln70* não é frequente no Brasil;

- Embora as frequências das mutações possam se assemelhar entre países geograficamente próximos, cada população tem características únicas.

REFERÊNCIAS BIBLIOGRÁFICAS

- Al-Sannaa NA, Bay L, Barbouth DS, Benhayoun Y, Goizet C, Guelbert N, Jones SA, Kyosen SO, Martins AM, Phornphutkul C and others. 2015. Early treatment with laronidase improves clinical outcomes in patients with attenuated MPS I: a retrospective case series analysis of nine sibships. *Orphanet J Rare Dis* 10:131.
- Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. 2010. *Biologia Molecular da Célula*. Porto Alegre: Artmed.
- Alif N, Hess K, Straczek J, Sebbar S, N'Bou A, Nabet P, Dousset B. 1999. Mucopolysaccharidosis type I: characterization of a common mutation that causes Hurler syndrome in Moroccan subjects. *Ann Hum Genet* 63(Pt 1):9-16.
- Archer LD, Langford-Smith KJ, Bigger BW, Fildes JE. 2014. Mucopolysaccharide diseases: a complex interplay between neuroinflammation, microglial activation and adaptive immunity. *J Inherit Metab Dis* 37(1):1-12.
- Atceken N, Ozgul RK, Yucel Yilmaz D, Tokatli A, Coskun T, Sivri HS, Dursun A, Karaca M. 2016. Evaluation and identification of IDUA gene mutations in Turkish patients with mucopolysaccharidosis type I. *Turk J Med Sci* 46(2):404-8.
- Bach G, Moskowitz SM, Tieu PT, Matynia A, Neufeld EF. 1993. Molecular analysis of Hurler syndrome in Druze and Muslim Arab patients in Israel: multiple allelic mutations of the IDUA gene in a small geographic area. *Am J Hum Genet* 53(2):330-8.
- Baehner F, Schmiedeskamp C, Krummenauer F, Miebach E, Bajbouj M, Whybra C, Kohlschutter A, Kampmann C, Beck M. 2005. Cumulative incidence rates of the mucopolysaccharidoses in Germany. *J Inherit Metab Dis* 28(6):1011-7.
- Beck M, Arn P, Giugliani R, Muenzer J, Okuyama T, Taylor J, Fallet S. 2014. The natural history of MPS I: global perspectives from the MPS I Registry. *Genet Med* 16(10):759-65.
- Beesley CE, Meaney CA, Greenland G, Adams V, Vellodi A, Young EP, Winchester BG. 2001. Mutational analysis of 85 mucopolysaccharidosis type I families: frequency of known mutations, identification of 17 novel mutations and in vitro expression of missense mutations. *Hum Genet* 109(5):503-11.
- Bertola F, Filocamo M, Casati G, Mort M, Rosano C, Tylki-Szymanska A, Tuysuz B, Gabrielli O, Grossi S, Scarpa M and others. 2011. IDUA mutational profiling of a cohort of 102 European patients with mucopolysaccharidosis type I: identification and characterization of 35 novel alpha-L-iduronidase (IDUA) alleles. *Hum Mutat* 32(6):E2189-210.
- Bie H, Yin J, He X, Kermode AR, Goddard-Borger ED, Withers SG, James MN. 2013. Insights into mucopolysaccharidosis I from the structure and action of alpha-L-iduronidase. *Nat Chem Biol* 9(11):739-45.
- Bigg PW, Baldo G, Sleeper MM, O'Donnell PA, Bai H, Rokkam VR, Liu Y, Wu S, Giugliani R, Casal ML and others. 2013. Pathogenesis of mitral valve disease in mucopolysaccharidosis VII dogs. *Mol Genet Metab* 110(3):319-28.
- Botigue LR, Henn BM, Gravel S, Maples BK, Gignoux CR, Corona E, Atzmon G, Burns E, Ostrer H, Flores C and others. 2013. Gene flow from North Africa contributes to differential human genetic diversity in southern Europe. *Proc Natl Acad Sci U S A* 110(29):11791-6.
- Bowden GR, Balaesque P, King TE, Hansen Z, Lee AC, Pergl-Wilson G, Hurley E, Roberts SJ, Waite P, Jesch J and others. 2008. Excavating past population structures by surname-based sampling: the genetic legacy of the Vikings in northwest England. *Mol Biol Evol* 25(2):301-9.
- Brooks DA, Fabrega S, Hein LK, Parkinson EJ, Durand P, Yogalingam G, Matte U, Giugliani R, Dasvarma A, Eslahpazire J and others. 2001. Glycosidase active site mutations in human alpha-L-iduronidase. *Glycobiology* 11(9):741-50.

- Bunge S, Kleijer WJ, Steglich C, Beck M, Schwinger E, Gal A. 1995. Mucopolysaccharidosis type I: identification of 13 novel mutations of the alpha-L-iduronidase gene. *Hum Mutat* 6(1):91-4.
- Bunge S, Kleijer WJ, Steglich C, Beck M, Zuther C, Morris CP, Schwinger E, Hopwood JJ, Scott HS, Gal A. 1994. Mucopolysaccharidosis type I: identification of 8 novel mutations and determination of the frequency of the two common alpha-L-iduronidase mutations (W402X and Q70X) among European patients. *Hum Mol Genet* 3(6):861-6.
- Carvalho T, Matte U, Giugliani R, Baldo G. 2015. Genome Editing: Potential Treatment for Lysosomal Storage Diseases. *Curr Stem Cell Rep* 1:9-15.
- Chistiakov DA, Savost'anov KV, Kuzenkova LM, Gevorkyan AK, Pushkov AA, Nikitin AG, Pakhomov AV, Vashakmadze ND, Zhurkova NV, Podkletnova TV and others. 2014. Molecular characteristics of patients with glycosaminoglycan storage disorders in Russia. *Clin Chim Acta* 436:112-20.
- Chkioua L, Khedhiri S, Jaidane Z, Ferchichi S, Habib S, Froissart R, Bonnet V, Chaabouni M, Dandana A, Jrad T and others. 2007. [Mucopolysaccharidosis type I: identification of alpha-L-iduronidase mutations in Tunisian families]. *Arch Pediatr* 14(10):1183-9.
- Chkioua L, Khedhiri S, Kassab A, Bibi A, Ferchichi S, Froissart R, Vianey-Saban C, Laradi S, Miled A. 2011a. Molecular analysis of mucopolysaccharidosis type I in Tunisia: identification of novel mutation and eight Novel polymorphisms. *Diagn Pathol* 6:39.
- Chkioua L, Khedhiri S, Turkia HB, Tcheng R, Froissart R, Chahed H, Ferchichi S, Ben Dridi MF, Vianey-Saban C, Laradi S and others. 2011b. Mucopolysaccharidosis type I: molecular characteristics of two novel alpha-L-iduronidase mutations in Tunisian patients. *Diagn Pathol* 6:47.
- Clarke LA, Nelson PV, Warrington CL, Morris CP, Hopwood JJ, Scott HS. 1994. Mutation analysis of 19 North American mucopolysaccharidosis type I patients: identification of two additional frequent mutations. *Hum Mutat* 3(3):275-82.
- Clarke LA, Scott HS. 1993. Two novel mutations causing mucopolysaccharidosis type I detected by single strand conformational analysis of the alpha-L-iduronidase gene. *Hum Mol Genet* 2(8):1311-2.
- Clarke LA, Wraith JE, Beck M, Kolodny EH, Pastores GM, Muenzer J, Rapoport DM, Berger KI, Sidman M, Kakkis ED and others. 2009. Long-term efficacy and safety of laronidase in the treatment of mucopolysaccharidosis I. *Pediatrics* 123(1):229-40.
- Coutinho MF, Lacerda L, Alves S. 2012. Glycosaminoglycan storage disorders: a review. *Biochem Res Int* 2012:471325.
- Eisengart JB, Rudser KD, Tolar J, Orchard PJ, Kivisto T, Ziegler RS, Whitley CB, Shapiro EG. 2013. Enzyme replacement is associated with better cognitive outcomes after transplant in Hurler syndrome. *J Pediatr* 162(2):375-80 e1.
- Fahiminiya S, Almuriekh M, Nawaz Z, Staffa A, Lepage P, Ali R, Hashim L, Schwartzentruber J, Abu Khadija K, Zaineddin S and others. 2014. Whole exome sequencing unravels disease-causing genes in consanguineous families in Qatar. *Clin Genet* 86(2):134-41.
- Federhen A, Burin M, Leistner-Segal S, Matte U, Tirelli K, Facchin A, Pasqualim G, Bender F, Rafaelli C, Giugliani R. 2016. Minimal Estimated Incidence of MPS I, II, IV-A and VI in Brazil and Comparison to the rest of the world. 14th International Symposium on MPS and Related Diseases, Bonn, Germany:1-82.
- Fernandez-Marmiesse A, Morey M, Pineda M, Eiris J, Couce ML, Castro-Gago M, Fraga JM, Lacerda L, Gouveia S, Perez-Poyato MS and others. 2014. Assessment of a targeted resequencing assay as a support tool in the diagnosis of lysosomal storage disorders. *Orphanet J Rare Dis* 9:59.
- Fuller M, Meikle PJ, Hopwood JJ. 2006. Epidemiology of lysosomal storage diseases: an overview. In: Mehta A, Beck M, Sunder-Plassmann G, editors. *Fabry Disease: Perspectives from 5 Years of FOS*. Oxford.
- Gabrielli O, Clarke LA, Ficcadenti A, Santoro L, Zampini L, Volpi N, Coppa GV. 2016. 12 year follow up of enzyme-replacement therapy in two siblings with attenuated mucopolysaccharidosis I: the important role of early treatment. *BMC Med Genet* 17:19.

- Gatti R, DiNatale P, Villani GR, Filocamo M, Muller V, Guo XH, Nelson PV, Scott HS, Hopwood JJ. 1997. Mutations among Italian mucopolysaccharidosis type I patients. *J Inher Metab Dis* 20(6):803-6.
- Ghosh A, Miller W, Orchard PJ, Jones SA, Mercer J, Church HJ, Tylee K, Lund T, Bigger BW, Tolar J and others. 2016. Enzyme replacement therapy prior to haematopoietic stem cell transplantation in Mucopolysaccharidosis Type I: 10 year combined experience of 2 centres. *Mol Genet Metab* 117(3):373-7.
- Giugliani R, Federhen A, Rojas MV, Vieira T, Artigalás O, Pinto LL, Azevedo AC, Acosta A, Bonfim C, Lourenço CM and others. 2010. Mucopolysaccharidosis I, II, and VI: Brief review and guidelines for treatment. *Genet Mol Biol* 33(4):589-604.
- Giugliani R, Federhen A, Vairo F, Vanzella C, Pasqualim G, da Silva LM, Giugliani L, de Boer AP, de Souza CF, Matte U and others. 2016. Emerging drugs for the treatment of mucopolysaccharidoses. *Expert Opin Emerg Drugs* 21(1):9-26.
- Gort L, Chabas A, Coll MJ. 1998. Analysis of five mutations in 20 mucopolysaccharidosis type 1 patients: high prevalence of the W402X mutation. *Mutations in brief no. 121*. Online. *Hum Mutat* 11(4):332-3.
- Jameson E, Jones S, Remington T. 2016. Enzyme replacement therapy with laronidase (Aldurazyme((R))) for treating mucopolysaccharidosis type I. *Cochrane Database Syst Rev* 4:CD009354.
- Kroepfl T, Milos I, Paul K, Plecko B, Paschke E. 2001. The frequency of common mutations among patients with mucopolysaccharidosis types I, II and IIIA diagnosed in Austria over the last 17 years. *Clin Genet* 60:393–394.
- Kwak M, Huh R, Kim J, Park H, Cho S, Jin D. 2016. Report of 5 novel mutations of the α -Liduronidase gene and comparison of Korean mutations in relation with those of Japan or China in patients with mucopolysaccharidosis I. *BMC Medical Genetics* 17:58.
- Laradi S, Tukul T, Erazo M, Shabbeer J, Chkioua L, Khedhiri S, Ferchichi S, Chaabouni M, Miled A, Desnick RJ. 2005. Mucopolysaccharidosis I: Alpha-L-Iduronidase mutations in three Tunisian families. *J Inher Metab Dis* 28(6):1019-26.
- Lee IJ, Hwang SH, Jeon BH, Song SM, Kim JS, Paik KH, Kwon EK, Jin DK. 2004. Mutational analysis of the alpha-L-iduronidase gene in 10 unrelated Korean type I mucopolysaccharidosis patients: Identification of four novel mutations. *Clin Genet* 66(6):575-6.
- Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB and others. 2016. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* 536(7616):285-91.
- Li P, Wood T, Thompson JN. 2002. Diversity of mutations and distribution of single nucleotide polymorphic alleles in the human alpha-L-iduronidase (IDUA) gene. *Genet Med* 4(6):420-6.
- Lucotte G. 2001. Distribution of the CCR5 gene 32-basepair deletion in West Europe. A hypothesis about the possible dispersion of the mutation by the Vikings in historical times. *Hum Immunol* 62(9):933-6.
- Matte U, Leistner S, Lima L, Schwartz I, Giugliani R. 2000. Unique frequency of known mutations in Brazilian MPS I patients. *Am J Med Genet* 90(2):108-9.
- Matte U, Yogalingam G, Brooks D, Leistner S, Schwartz I, Lima L, Norato DY, Brum JM, Beesley C, Winchester B and others. 2003. Identification and characterization of 13 new mutations in mucopolysaccharidosis type I patients. *Mol Genet Metab* 78(1):37-43.
- Menon KP, Neufeld EF. 1994. Evidence for degradation of mRNA encoding alpha-L-iduronidase in Hurler fibroblasts with premature termination alleles. *Cell Mol Biol (Noisy-le-grand)* 40(7):999-1005.
- Milman N, Pedersen P. 2003. Evidence that the Cys282Tyr mutation of the HFE gene originated from a population in Southern Scandinavia and spread with the Vikings. *Clin Genet* 64(1):36-47.
- Murphy AM, Lambert D, Treacy EP, O'Meara A, Lynch SA. 2009. Incidence and prevalence of mucopolysaccharidosis type 1 in the Irish republic. *Arch Dis Child* 94(1):52-4.

- Neufeld EF, Muenzer J. 2001. The mucopolysaccharidosis. In: CR BA, Sly WS and Valle D editor. *The Metabolic and Molecular Bases of Inherited Disease*. New York: McGraw-Hill Co. p 3421-3452.
- Pasqualim G, Ribeiro MG, da Fonseca GG, Szlago M, Schenone A, Lemes A, Rojas MV, Matte U, Giugliani R. 2015. p.L18P: a novel IDUA mutation that causes a distinct attenuated phenotype in mucopolysaccharidosis type I patients. *Clin Genet* 88(4):376-80.
- Pena SD, Di Pietro G, Fuchshuber-Moraes M, Genro JP, Hutz MH, Kehdy Fde S, Kohlrausch F, Magno LA, Montenegro RC, Moraes MO and others. 2011. The genomic ancestry of individuals from different geographical regions of Brazil is more uniform than expected. *PLoS One* 6(2):e17063.
- Pereira VG, Martins AM, Micheletti C, D'Almeida V. 2008. Mutational and oxidative stress analysis in patients with mucopolysaccharidosis type I undergoing enzyme replacement therapy. *Clin Chim Acta* 387(1-2):75-9.
- Pineda T, Marie S, Gonzalez J, Garcia AL, Acosta A, Morales M, Correa LN, Vivas R, Escobar X, Protzel A and others. 2014. Genotypic and bioinformatic evaluation of the alpha-L-iduronidase gene and protein in patients with mucopolysaccharidosis type I from Colombia, Ecuador and Peru. *Mol Genet Metab Rep* 1:468-473.
- Pollard LM, Jones JR, Wood TC. 2013. Molecular characterization of 355 mucopolysaccharidosis patients reveals 104 novel mutations. *J Inherit Metab Dis* 36(2):179-87.
- Prommajan K, Ausavarat S, Srichomthong C, Puangsricharern V, Suphapeetiporn K, Shotelersuk V. 2011. A novel p.E276K IDUA mutation decreasing alpha-L-iduronidase activity causes mucopolysaccharidosis type I. *Mol Vis* 17:456-60.
- Regueiro M, Garcia-Bertrand R, Fadhlouzi-Zid K, Alvarez J, Herrera RJ. 2015. From Arabia to Iberia: A Y chromosome perspective. *Gene* 564(2):141-52.
- Saftig P, Klumperman J. 2009. Lysosome biogenesis and lysosomal membrane proteins: trafficking meets function. *Nat Rev Mol Cell Biol* 10(9):623-35.
- Saito S, Ohno K, Maita N, Sakuraba H. 2014. Structural and clinical implications of amino acid substitutions in alpha-L-iduronidase: insight into the basis of mucopolysaccharidosis type I. *Mol Genet Metab* 111(2):107-12.
- Scott HS, Bunge S, Gal A, Clarke LA, Morris CP, Hopwood JJ. 1995. Molecular genetics of mucopolysaccharidosis type I: diagnostic, clinical, and biological implications. *Hum Mutat* 6(4):288-302.
- Scott HS, Guo XH, Hopwood JJ, Morris CP. 1992a. Structure and sequence of the human alpha-L-iduronidase gene. *Genomics* 13(4):1311-3.
- Scott HS, Litjens T, Hopwood JJ, Morris CP. 1992b. A common mutation for mucopolysaccharidosis type I associated with a severe Hurler syndrome phenotype. *Hum Mutat* 1(2):103-8.
- Scott HS, Litjens T, Nelson PV, Brooks DA, Hopwood JJ, Morris CP. 1992c. alpha-L-iduronidase mutations (Q70X and P533R) associate with a severe Hurler phenotype. *Hum Mutat* 1(4):333-9.
- Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. 2014. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 133(1):1-9.
- Suárez A, Sánchez L. 2014. Mutaciones genéticas y su relación con el fenotipo clínico en pacientes con mucopolisacaridosis de tipo I en el noreste de México. *Gaceta Médica de México* 150:289-96.
- Sun L, Li C, Song X, Zheng N, Zhang H, Dong G. 2011. Three novel alpha-L-iduronidase mutations in 10 unrelated Chinese mucopolysaccharidosis type I families. *Genet Mol Biol* 34(2):195-200.
- Tebani A, Zanutene-Cheriet L, Adjtoutah Z, Abily-Donval L, Brasse-Lagnel C, Laquerriere A, Marret S, Chalabi Benabdellah A, Bekri S. 2016. Clinical and Molecular Characterization of Patients with Mucopolysaccharidosis Type I in an Algerian Series. *Int J Mol Sci* 17(5).

- Terlato NJ, Cox GF. 2003. Can mucopolysaccharidosis type I disease severity be predicted based on a patient's genotype? A comprehensive review of the literature. *Genet Med* 5(4):286-94.
- Tieu PT, Menon K, Neufeld EF. 1994. A mutant stop codon (TAG) in the IDUA gene is used as an acceptor splice site in a patient with Hurler syndrome (MPS IH). *Hum Mutat* 3(3):333-6.
- Uttarilli A, Ranganath P, Matta D, Md Nurul Jain J, Prasad K, Babu AS, Girisha KM, Verma IC, Phadke SR, Mandal K and others. 2016. Identification and characterization of 20 novel pathogenic variants in 60 unrelated Indian patients with mucopolysaccharidoses type I and type II. *Clin Genet* 90(6):496-508.
- Vazna A, Beesley C, Berna L, Stolnaja L, Myskova H, Bouckova M, Vlaskova H, Poupetova H, Zeman J, Magner M and others. 2009. Mucopolysaccharidosis type I in 21 Czech and Slovak patients: mutation analysis suggests a functional importance of C-terminus of the IDUA protein. *Am J Med Genet A* 149A(5):965-74.
- Venturi N, Rovelli A, Parini R, Menni F, Brambillasca F, Bertagnolio F, Uziel G, Gatti R, Filocamo M, Donati MA and others. 2002. Molecular analysis of 30 mucopolysaccharidosis type I patients: evaluation of the mutational spectrum in Italian population and identification of 13 novel mutations. *Hum Mutat* 20(3):231.
- Wang X, Zhang W, Shi H, Qiu Z, Meng Y, Yao F, Wei M. 2012. Mucopolysaccharidosis I mutations in Chinese patients: identification of 27 novel mutations and 6 cases involving prenatal diagnosis. *Clin Genet* 81(5):443-52.
- Yamagishi A, Tomatsu S, Fukuda S, Uchiyama A, Shimosawa N, Suzuki Y, Kondo N, Sukegawa K, Orii T. 1996. Mucopolysaccharidosis type I: identification of common mutations that cause Hurler and Scheie syndromes in Japanese populations. *Hum Mutat* 7(1):23-9.
- Yogalingam G, Guo XH, Muller VJ, Brooks DA, Clements PR, Kakkis ED, Hopwood JJ. 2004. Identification and molecular characterization of alpha-L-iduronidase mutations present in mucopolysaccharidosis type I patients undergoing enzyme replacement therapy. *Hum Mutat* 24(3):199-207.