

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
DEPARTAMENTO DE BIOQUÍMICA PROFESSOR TUISKON DICK
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS: BIOQUÍMICA

**ANÁLISE MOLECULAR EM PACIENTES COM DOENÇA DE GAUCHER: UMA
ABORDAGEM ABRANGENTE PARA IDENTIFICAÇÃO DE ALELOS
MUTANTES**

MARINA SIEBERT

Orientadora

Profa. Dra. Maria Luiza Saraiva Pereira

Porto Alegre, fevereiro de 2010.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE
DEPARTAMENTO DE BIOQUÍMICA PROFESSOR TUISKON DICK
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS: BIOQUÍMICA

**ANÁLISE MOLECULAR EM PACIENTES COM DOENÇA DE GAUCHER: UMA
ABORDAGEM ABRANGENTE PARA IDENTIFICAÇÃO DE ALELOS
MUTANTES**

MARINA SIEBERT

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Biológicas - Bioquímica da Universidade Federal do Rio Grande do Sul como requisito parcial à obtenção do grau de Mestre em Bioquímica.

Porto Alegre, fevereiro de 2010.

PENSAMENTO

“Tudo o que a mente humana pode conceber, ela pode conquistar.”

W. Clement Stone (1902-2002)

AGRADECIMENTOS

À minha orientadora Profa. Dra. Maria Luiza Saraiva Pereira, por ter acreditado no meu potencial, pelo apoio e incentivo, pelos ensinamentos e, principalmente, pela amizade.

Aos colegas do Laboratório de Identificação Genética pela amizade, pela motivação, pelas ideias e risadas compartilhadas.

Ao meu amigo Hugo Bock, pelos ensinamentos, pela ajuda e pelo convívio.

Às amigas, Juliana, Mariana e Vanessa, que conquistei ao longo desse período, mas que levarei comigo para a vida inteira.

Aos meus pais, Marli e Mauro, pelo amor e pelo apoio incondicional, que sempre torceram pelo meu sucesso.

Aos meus irmãos, pelo carinho, pelo interesse e pelo apoio.

Ao meu namorado Tiago, pelo amor, pela compreensão, pela força e incentivo para que eu pudesse concluir mais esta etapa.

Aos pacientes e seus familiares por proverem amostras biológicas para este estudo.

Ao CNPq e ao FIPE-HCPA pelo apoio financeiro.

A todos, que de alguma forma, contribuíram para a realização e conclusão deste trabalho.

RESUMO

A doença de Gaucher (DG) é uma doença lisossômica de depósito, de herança autossômica recessiva, causada pela deficiência da enzima glicocerebrosidase (GC). Essa deficiência é causada por mutações no gene (GBA) que codifica esta enzima. Até o momento, mais de 250 mutações já foram identificadas nesse gene. O objetivo deste estudo foi identificar mutações na região codificante do gene GBA de pacientes brasileiros com DG através de PCR longo, seguido de *nested* PCR e sequenciamento direto. As análises foram realizadas em 54 pacientes não-aparentados com DG, confirmados por apresentar baixa atividade enzimática e com pelo menos um alelo mutante não-identificado após a triagem para 4 mutações comuns (p.N370S, p.L444P, 84insG e IVS2+1G>A). O protocolo introduzido permitiu a identificação de 7 variações de sequência novas (p.S125N, p.F213L, p.P245T, p.W378C, p.D399H, 982-983insTGC e IVS10+1G>T) no gene GBA de pacientes com DG. Todas essas novas variações de sequência são provavelmente alterações responsáveis pela doença, pois alteram resíduos conservados da proteína, inserem um aminoácido ou prejudicam o *splicing* normal do RNA mensageiro. Consequentemente, estas mutações causam mudanças na estrutura e/ou na função da GC. Além dessas, 24 mutações raras que já foram descritas previamente, também, foram identificadas nesse trabalho, contribuindo na definição do genótipo dos pacientes. A identificação dos alelos mutantes é importante para o conhecimento do espectro de mutações no nosso país e para aumentar o conhecimento das bases moleculares da doença. Além disso, essas informações podem também contribuir para a melhor compreensão de correlações genótipo-fenótipo, assim como, para o aconselhamento genético e/ou para a oferta de análises moleculares individualizadas para famílias em risco.

ABSTRACT

Gaucher disease (GD) is an autosomal recessive lysosomal storage disorder caused by deficiency of the glucocerebrosidase (GC). This deficiency is caused by mutations in the gene (GBA) coding for this enzyme. To date, more than 250 mutations have been identified associated with GD. The aim of this study was to identify mutations in the coding region of the GBA gene in Brazilian patients with GD through long-range PCR, followed by nested PCR and direct sequencing. Analyses were carried out in 54 unrelated GD patients who presented low enzymatic activity and had at least one unidentified disease-associated allele after screening for 4 common mutations (p.N370S, p.L444P, 84insG, and IVS2+1G>A). Protocol described here allowed the identification of 7 novel sequence variations (p.S125N, p.F213L, p.P245T, p.W378C, p.D399H, 982-983insTGC, and IVS10+1G>T) in the GBA gene of patients with GD. As these new sequence variations change conserved protein residues, insert an amino acid or affect mRNA normal processing, they are likely to be disease causing mutations. Consequently, these mutations cause changes in the GC structure and/or function. Besides, 24 rare mutations that were previously described were also identified in this work leading to the definition of patients' genotype. The identification of mutant alleles is crucial for improving the knowledge of GBA mutation spectrum in our country and to the better understanding of molecular basis of the disease. Furthermore, such information may also contribute to the establishment of genotype-phenotype correlations as well as for more specific genetic counseling and/or to offer a customized molecular analysis for families at risk.

LISTA DE ABREVIATURAS

A	adenina
ARMS-PCR	<i>Amplification Refractory Mutation System - PCR</i>
C	citossina
cDNA	ácido desoxirribonucléico complementar
DG	doença de Gaucher
DNA	ácido desoxirribonucléico
G	guanina
GBA	gene que codifica a enzima glicocerebrosidase
GC	glicocerebrosidase
kb	quilobase
OMIM	<i>Online Mendelian Inheritance in Man</i>
pb	pares de bases
PCR	reação em cadeia da polimerase (<i>Polimerase Chain Reaction</i>)
PSGBA	pseudogene da glicocerebrosidase
RFLP	<i>Restriction Fragment Length Polymorphism</i>
RNA	ácido ribonucléico
mRNA	ácido ribonucléico mensageiro
SAP-C	saposina C
SNC	sistema nervoso central
T	timina
<i>Taq</i> DNA polimerase	DNA polimerase de <i>Thermus aquaticus</i>
TRE	terapia de reposição enzimática
TRS	terapia de redução de substrato
UV	ultravioleta

LISTA DE FIGURAS

Figura 1. Reação catalisada pela enzima glicocerebrosidase	3
Figura 2. Representação da estrutura tridimensional da glicocerebrosidase	12

NOMENCLATURA DAS MUTAÇÕES

Neste estudo, a nomenclatura adotada para as mutações no gene GBA foi baseada nas recomendações da *Human Genome Variation Society* - HGVS (<http://www.hgvs.org>) (den Dunnen e Antonarakis, 2000).

SUMÁRIO

1. INTRODUÇÃO	1
1.1. Doença de Gaucher	1
1.1.1. Manifestações clínicas	1
1.1.2. Defeito metabólico na DG	2
1.1.3. Diagnóstico da DG	3
1.1.4. Tratamento.....	4
1.2. Gene da GC e seu pseudogene	5
1.2.1. Mutações no gene GBA.....	6
1.2.2. Correlações genótipo-fenótipo	8
1.3. Análise estrutural da GC	10
1.4. Justificativa	13
2. OBJETIVOS	15
2.1. Objetivo Geral	15
2.2. Objetivos Específicos	15
3. RESULTADOS	16
3.1. Artigo científico	16
4. DISCUSSÃO	47
5. CONCLUSÕES	51
6. REFERÊNCIAS	52
7. ANEXO	58

1. INTRODUÇÃO

1.1. Doença de Gaucher

A doença de Gaucher (DG) é uma doença lisossômica de depósito caracterizada pelo acúmulo do substrato glicosilceramida (glicocerebrosídeo), um composto intermediário no catabolismo de globosídeos e gangliosídeos. Este acúmulo é causado pela deficiência total ou parcial da enzima glicocerebrosidase (ou β -glicosidase ácida, E.C. 3.2.1.45) ou, em casos raros, pela deficiência do ativador dessa enzima, a saposina C (SAP-C). A DG é herdada de forma autossômica recessiva e é uma patologia comum em judeus Ashkenazi (Beutler e Grabowski, 2001).

1.1.1. Manifestações clínicas

A DG é caracterizada por hepatoesplenomegalia, anemia e trombocitopenia, podendo também estar presentes sintomas ósseos e pulmonares (Grabowski, 2008). Esta doença é classificada em três tipos clínicos tendo como base a presença e a gravidade do envolvimento neurológico (Blech-Hermoni *et al.*, 2009).

O tipo 1 (não-neuropático; OMIM 230800) é a forma mais comum da DG e não causa prejuízos ao sistema nervoso central (SNC) (Hughes, 2009). A variação fenotípica abrange desde pacientes assintomáticos podendo chegar ao risco de morte. O início das manifestações clínicas pode ocorrer em qualquer idade, inclusive na vida adulta. O aparecimento precoce dos sinais clínicos geralmente está associado a um grave e rápido comprometimento dos órgãos (Grabowski, 2004). Essa forma da doença pode afetar 1 em

cada 60.000 indivíduos na população em geral. A sua incidência é pan-étnica, mas pode ser alta em judeus Ashkenazi, chegando a afetar 1 em cada 850 indivíduos (Hughes, 2009).

O tipo 2 (neuropático agudo; OMIM 230900) é caracterizado pelo aparecimento precoce das manifestações clínicas e envolvimento do SNC, resultando em óbito logo nos primeiros anos de vida (Martins *et al.*, 2009). É o tipo mais grave, o qual apresenta um rápido envolvimento neurológico devido ao acúmulo de glicosilceramida também nos neurônios (Grabowski, 2004). Essa forma da doença pode afetar 1 em cada 100.000 indivíduos na população em geral (Martins *et al.*, 2009).

O tipo 3 (neuropático subagudo; OMIM 231000) é uma forma neurodegenerativa que progride mais lentamente do que o tipo 2. Os sintomas iniciam entre a infância e a adolescência (Alfonso *et al.*, 2007). Por definição, o tipo 3 inclui todos os pacientes com algum grau de comprometimento neurológico, mas que sobreviveram aos primeiros anos de vida (Koprivica *et al.*, 2000). A frequência estimada da DG tipo 3 é de 1 em cada 50.000 indivíduos na população em geral (Chen e Wang, 2008).

1.1.2. Defeito metabólico na DG

A DG é um erro inato do metabolismo dos glicoesfingolípídeos, que resulta no acúmulo lisossômico de um material parcialmente degradado dentro das células da linhagem monócito-macrófago (Germain, 2004). Esse acúmulo é causado pela ausência ou deficiência da enzima glicocerebrosidase (GC), que ocasiona um prejuízo na hidrólise da ligação β -glicosídica do substrato glicosilceramida (figura 1) (Beutler e Grabowski, 2001).

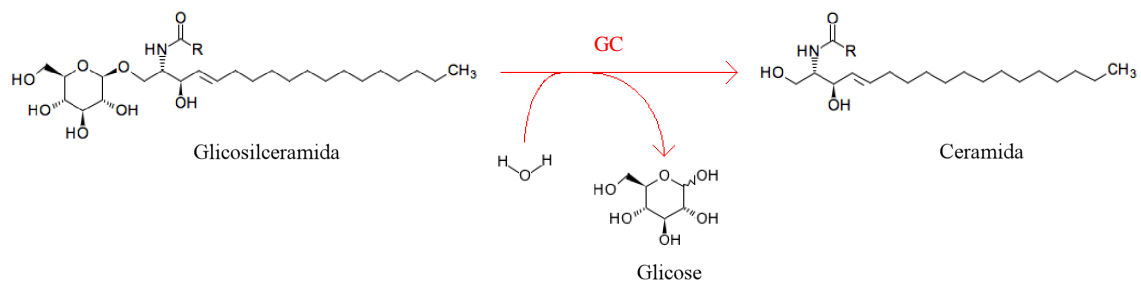


Figura 1. Reação catalisada pela enzima glicocerebrosidase. A glicocerebrosidase (GC) catalisa a hidrólise da ligação β -glicosídica da glicosilceramida produzindo glicose e ceramida (adaptado de http://www.genome.jp/dbget-bin/www_bget?rn:R01498).

O acúmulo de glicosilceramida dentro dos macrófagos resulta nas “células de Gaucher”, as quais têm aparência de “papel de seda amassado” e são repletas de lipídeos (Martins *et al.*, 2009). Estas células anômalas estão distribuídas pelo corpo todo, mas principalmente no fígado, baço e medula óssea. Conseqüentemente, essas células anômalas prejudicam e impedem o funcionamento normal destes órgãos e tecidos (Harmanci e Bayraktar, 2008).

1.1.3. Diagnóstico da DG

Antes da identificação da deficiência enzimática, o diagnóstico da DG era realizado através de exame histológico de biópsia da medula óssea ou, em alguns casos, do fígado ou do baço (Beutler, 2006). Entretanto, a presença de células semelhantes às de Gaucher (chamadas de pseudocélulas de Gaucher) em outras doenças, tais como leucemia mielóide crônica, leucemia mielóide aguda, mieloma múltiplo, talassemia e doença de Hodgkin, poderia muitas vezes levar a diagnósticos incorretos (Alterini *et al.*, 1996).

Atualmente, o diagnóstico laboratorial da DG é feito através da medida da atividade enzimática em leucócitos e/ou de fibroblastos, obtidos a partir de sangue periférico e a partir de biópsia de pele, respectivamente (Coelho e Giugliani, 2000; Michelin *et al.*, 2005).

A análise molecular é complementar ao diagnóstico enzimático. A identificação de mutações no gene GBA de pacientes com DG é comumente usada em busca de mutações mais frequentes, tais como p.N370S e p.L444P (Beutler e Grabowski, 2001).

1.1.4. Tratamento

Antes dos anos 90, o tratamento existente para DG era meramente paliativo, incluindo alívio sintomático e intervenções cirúrgicas, tais como correções ortopédicas e esplenectomia (Zimran e Elstein, 2003). Atualmente, duas formas de tratamento estão disponíveis para DG: a terapia de reposição enzimática (TRE) e a terapia de redução de substrato (TRS) (Martins *et al.*, 2009).

Desde 1992, a TRE tornou-se uma realidade para pacientes com DG (Grabowski, 2008). A TRE teve início com o uso da enzima purificada de placenta humana (alglucerase, Ceredase[®]). Posteriormente, a alglucerase foi substituída pela forma recombinante da GC (imiglucerase, Cerezyme[®]) (Beck, 2010). A terapia com enzima tem provado ser segura e efetiva em mais de 3000 pacientes no mundo todo (Sawkar *et al.*, 2006). A TRE com imiglucerase previne as manifestações da DG, melhora a anemia, a trombocitopenia, a organomegalia, as dores e crises ósseas associadas à doença, bem como, a qualidade de vida dos pacientes (Martins *et al.*, 2009). As duas enzimas sofreram

uma modificação nos resíduos de carboidrato a fim de atingir preferencialmente os macrófagos via receptor de manose (Beck, 2010).

A TRE é mais efetiva para o tratamento de pacientes com DG tipo 1, pois a enzima não atravessa a barreira hematoencefálica (Sawkar *et al.*, 2006). Entretanto, as manifestações sistêmicas da DG tipo 3 também respondem bem à TRE. A dose empregada é específica para cada paciente, dependendo da idade do indivíduo e da progressão da doença (Martins *et al.*, 2009). O alto custo do tratamento pode limitar o número de pacientes beneficiados pela TRE, principalmente em países de terceiro mundo (Jmoudiak e Futerman, 2005).

A TRS é outra forma de tratamento da DG e consiste na ação indireta no substrato armazenado dentro dos macrófagos (Zimran e Elstein, 2003). Nesta terapia, usa-se um inibidor da enzima glicosilceramida sintase, que corresponde à primeira etapa na biossíntese dos glicosfingolípídeos (Aerts *et al.*, 2006). Diferentemente da TRE, o acúmulo de substrato é evitado pela redução da quantidade do substrato glicosilceramida. Assim, a TRS atenua a taxa de síntese para alcançar um equilíbrio com a atividade reduzida da GC endógena nos pacientes com DG, evitando o armazenamento de substrato patológico (Zimran e Elstein, 2003). Um dos inibidores usados é o *N*-butil-deoxinojirimicina (Miglustat, Zavesca[®]), o qual é indicado para tratamento via oral (Beutler, 2006).

1.2. Gene da GC e seu pseudogene

O gene funcional da glicocerebrosidase (GBA; GenBank-EMBL J03059) está localizado na região q21 do cromossomo 1 (Horowitz *et al.*, 1989). Esta região é muito rica em genes, contendo pelo menos 7 genes funcionais e 2 pseudogenes (Hruska *et al.*, 2008). O gene GBA engloba 7,6 kb de DNA genômico, sendo dividido em 11 éxons (Horowitz *et al.*, 1989). O cDNA do gene GBA abrange 2 kb e seu mRNA apresenta 2 códons de iniciação. A transcrição iniciada em ambos os códons resultará em uma proteína funcional. A proteína produzida contém um polipeptídeo sinalizador de 19 ou de 39 resíduos, dando origem a uma proteína madura contendo 497 aminoácidos, com um peso molecular de 56 kDa (Hruska *et al.*, 2008).

O gene GBA tem um pseudogene não-funcional (PSGBA, GenBank-EMBL J03060) localizado a 16 kb da porção 3' terminal do gene funcional. Este pseudogene abrange 5,7 kb de DNA, devido a algumas perdas de sequências nucleotídicas em relação ao gene funcional (Horowitz *et al.*, 1989). O PSGBA é altamente homólogo ao gene GBA, correspondendo a 96% de identidade entre seus nucleotídeos (Beutler e Grabowski, 2001). Algumas das mutações encontradas em pacientes com DG são idênticas à sequência do PSGBA e outras resultam da recombinação entre o gene GBA e seu pseudogene (Blech-Hermoni *et al.*, 2009). Por isso, ao investigar mutações em pacientes com DG, as semelhanças e as diferenças existentes entre o gene e o pseudogene devem ser levadas em consideração (Hruska *et al.*, 2008).

1.2.1. Mutações no gene GBA

Até o momento, mais de 250 mutações já foram identificadas no gene GBA, incluindo mutações de ponto, deleções, inserções, mutações de *splicing* e diversos

rearranjos (Hruska *et al.*, 2008). A distribuição das mutações no gene GBA varia em diferentes populações. Mas, quatro destas mutações são encontradas com frequência em várias populações (Emre *et al.*, 2008).

A mutação mais comum em pacientes com DG é a transição de uma adenina para uma guanina no nucleotídeo 1226 do cDNA, localizada no éxon 9 do gene. Essa transição determina a substituição do aminoácido asparagina por serina na posição 370 da proteína GC (p.N370S; c.1226A>G) (Beutler e Grabowski, 2001).

Outra mutação é a transição de uma timina para uma citosina no nucleotídeo 1448 do cDNA, correspondente ao éxon 10 do gene GBA. Essa transição provoca a substituição de leucina por prolina no resíduo 444 da proteína madura (p.L444P; c.1448T>C). A posição homóloga na sequência do PSGBA é ocupada por citosina, exatamente como na mutação p.L444P que produz a DG (Beutler e Grabowski, 2001).

A inserção de uma guanina no nucleotídeo 84 do cDNA, denominada comumente como 84GG, será designada neste trabalho como 84insG (c.84dupG), de acordo com as recomendações internacionais de denominação das mutações. Essa mutação provoca uma mudança no quadro de leitura (*frameshift*) e, conseqüentemente, a proteína não é produzida (Beutler, 2006).

A mutação de *splicing*, decorrente da substituição de guanina por adenina no primeiro nucleotídeo do segundo íntron (IVS2+1G>A; c.115+1G>A), impede a formação da enzima GC (Beutler, 2006).

Uma característica importante observada em pacientes com DG é a ocorrência de alelos mutantes complexos (também conhecidos como “Rec”, de alelos recombinantes), os

quais incluem mais de uma mutação de ponto em um mesmo alelo. Esses alelos são possivelmente gerados por *crossing over* desigual entre o gene GBA e seu pseudogene ou por eventos de conversão do gene funcional (Beutler e Grabowski, 2001). Os sítios destes eventos de recombinação são variáveis, estendendo-se desde o íntron 2 até o éxon 11 (Hruska *et al.*, 2008). Os alelos complexos mais prevalentes são o RecNciI, que abrange as mutações p.L444P, p.A456P (c.1483G>C) e a mutação silenciosa p.V460V (c.1497G>C), e o RecTL, que compreende as mesmas mutações do alelo RecNciI mais a mutação p.D409H (c.1342G>C) (Cormand *et al.*, 1998). A prevalência do alelo RecNciI é geralmente inferior a 8% em todas as populações estudadas até o momento. Esse alelo é raro em pacientes judeus Ashkenazi e mais frequente em populações não-Ashkenazi, onde varia de 2,8% em pacientes espanhóis a 7,2% em pacientes australianos (Cormand *et al.*, 1998). É importante mencionar que, em alguns casos, o alelo RecNciI pode ter sido considerado equivocadamente como p.L444P, caso as outras mutações presentes no alelo complexo não tenham sido analisadas (Cormand *et al.*, 1998).

As mutações p.N370S, p.L444P, 84insG e IVS2+1G>A representam 90% dos alelos mutantes na população de judeus Ashkenazi, apesar de constituírem menos de 75% dos alelos nas populações não-Ashkenazi (Emre *et al.*, 2008).

A frequência das mutações comuns entre pacientes brasileiros é semelhante àquela descrita para outras populações, sendo que as mutações p.N370S e p.L444P representam aproximadamente 50% e 25% dos alelos, respectivamente. A terceira mutação mais frequente no Brasil é a p.G377S (c.1246G>A) (Rozenberg *et al.*, 2006a).

1.2.2. Correlações genótipo-fenótipo

As mutações no gene GBA são comumente caracterizadas como mutações brandas, graves ou nulas. A presença de uma mutação branda (p.N370S é o protótipo deste grupo), mesmo em combinação com um alelo nulo ou grave, está associada à DG tipo 1, não sendo encontrada em pacientes com comprometimento neurológico. A combinação de duas mutações graves ou de uma mutação grave com uma mutação nula está associada às variantes neuropáticas da DG (tipo 2 e tipo 3). A ocorrência de homozigotos para mutações nulas não foi encontrada em pacientes com DG, provavelmente por este genótipo ser incompatível com a vida (Rozenberg *et al.*, 2006b).

A compreensão da correlação genótipo-fenótipo permanece limitada e incompleta. A variação fenotípica da DG é significativa, não somente entre pacientes com a mesma forma clínica, mas também entre aqueles com o mesmo genótipo. Esta observação é bastante notável entre pacientes homozigotos para a mutação p.N370S, visto que os mesmos podem ser tanto assintomáticos durante toda a vida como podem desenvolver organomegalia acentuada e ter um envolvimento esquelético incapacitante (Hruska *et al.*, 2008).

A mutação p.N370S é encontrada somente em pacientes com DG tipo 1 (Grabowski, 2008). Os pacientes com os genótipos [p.N370S] + [84insG], [p.N370S] + [IVS2+1G>A], [p.N370S] + [p.L444P] ou [p.N370S] + [1263-1317del] apresentam manifestações clínicas mais graves do que os homozigotos para p.N370S (Grabowski e Horowitz, 1997).

Os pacientes homozigotos para a mutação p.L444P frequentemente desenvolvem a forma neuropática tipo 3 da DG. Os pacientes afetados mais gravemente (tipo 2)

apresentam o genótipo [p.L444P] + [RecNciI] ou [p.L444P] + alelo nulo (Grabowski, 2008).

A DG, assim como muitas doenças de herança monogênica, não é necessariamente simples (Sidransky, 2004). Genes modificadores, genes contíguos, proteínas transportadoras, proteínas ativadoras, substratos alternativos e fatores ambientais podem contribuir significativamente no fenótipo observado nos pacientes com DG (Hruska *et al.*, 2008). Apesar das correlações genótipo-fenótipo existentes, deve-se ter cautela ao se confiar unicamente na genotipagem do paciente a fim de predizer o prognóstico ou afirmar a necessidade de terapia (Koprivica *et al.*, 2000).

1.3. Análise estrutural da GC

A enzima GC é uma hidrolase lisossômica solúvel que pode estar associada à membrana (Alattia *et al.*, 2007). A glicoproteína madura é composta por 497 aminoácidos, derivados de 516 ou 536 resíduos incluindo as sequências líder, as quais são removidas durante a passagem pelo retículo endoplasmático (Liou *et al.*, 2006). A sequência de aminoácidos da GC parece ser não variável, não existindo polimorfismos funcionais, com exceção das mutações responsáveis pela DG (Grabowski e Horowitz, 1997).

A enzima apresenta 5 sítios de *N*-glicosilação, porém somente os 4 primeiros são utilizados (Asn19, Asn59, Asn146 e Asn270) (Berg-Fussman *et al.*, 1993). A glicosilação é essencial para o desenvolvimento da conformação cataliticamente ativa da enzima, ocorrendo principalmente no primeiro sítio (no aminoácido da posição 19) (Liou *et al.*, 2006).

A enzima recém-sintetizada é levada para os lisossomos via receptores de manose-6-fosfato e rotas independentes de oligossacarídeos (Beck, 2010). Uma vez no lisossomo, a GC tem meia-vida de 60 horas em fibroblastos e a glicosilação é importante para sua manutenção (Liou *et al.*, 2006). O seu principal substrato é a glicosilceramida, mas a glicosilesfingosina (análogo deacilado) também serve como substrato, apesar de ser menos importante. O defeito na GC leva ao acúmulo desses substratos em vários tecidos do corpo (Beutler e Grabowski, 2001).

O ciclo catalítico da GC tem um mecanismo de reação com 2 etapas, sendo elas: glicosilação do sítio ativo pelo substrato e subsequente deglicosilação com a liberação de uma β -glicose. O nucleófilo neste sítio ativo é o Glu340 e o ácido/base é o Glu235. A ligação *O*-glicosídica do substrato é protonada pelo Glu235 e, logo depois, ela é atacada pelo Glu340. A glicose torna-se covalentemente ligada à enzima e o grupo de saída (ceramida) é removido. Em seguida, o complexo enzima-glicose resultante é hidrolisado pela água e uma segunda substituição nucleofílica no carbono anomérico gera um produto com a mesma estereoquímica do substrato. Isso acontece devido à desprotonação da água pelo Glu235 e ao ataque do complexo enzima-glicose pelo OH⁻, liberando o Glu340 e a β -glicose (Liou *et al.*, 2006).

A SAP-C é uma glicoproteína composta por 80 aminoácidos que ativa a GC e a protege da proteólise (Sun *et al.*, 2010). Em pH ácido, a afinidade da SAP-C pela membrana lisossômica está fortemente aumentada (Kolter e Sandhoff, 2009). A SAP-C induz alterações na membrana que facilitam a interação entre a GC e a glicosilceramida, favorecendo a degradação desse substrato (Sun *et al.*, 2010). Na ausência da proteína ativadora e de fosfolípídeos carregados negativamente, a enzima GC não consegue acessar seu substrato, pois ele está dentro da bicamada lipídica (Alattia *et al.*, 2007).

A estrutura da GC está organizada em 3 domínios. O domínio I (resíduo 1 a 27 e 383 a 414) é uma fita tripla, que contém 2 pontes dissulfeto (entre os resíduos 4 e 16 e entre os resíduos 18 e 23). O domínio II (resíduo 30 a 75 e 431 a 497) é formado por 2 folhas β intimamente associadas, que geram um domínio independente semelhante ao dobramento da imunoglobulina. O domínio III (resíduo 76 a 381 e 416 a 430) é uma estrutura barril (β/α)₈, que contém o sítio catalítico. Os domínios II e III estão ligados por uma alça flexível, enquanto o domínio I está interagindo firmemente com o domínio III (figura 2) (Dvir *et al.*, 2003).

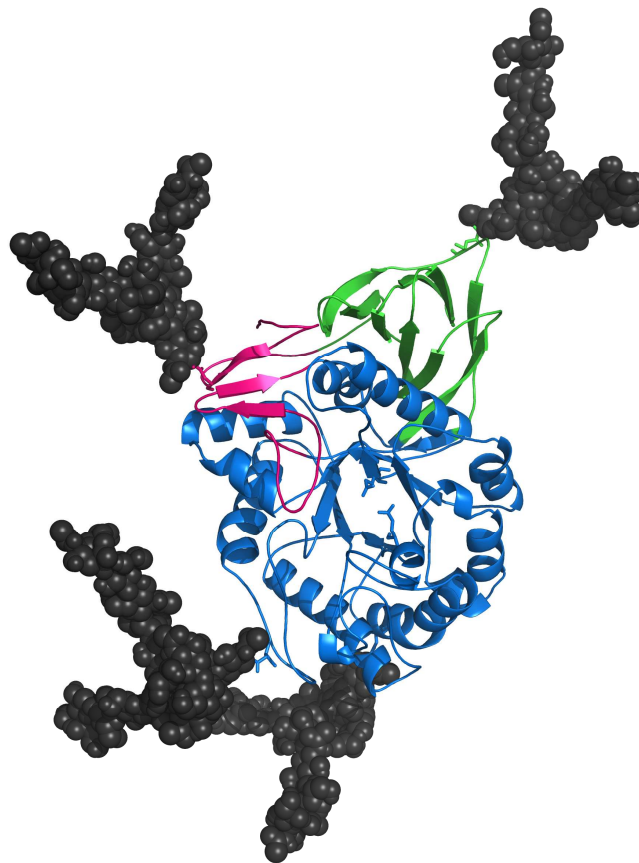


Figura 2. Representação da estrutura tridimensional da glicocerebrosidase. O domínio I está em rosa, o domínio II em verde e o domínio III em azul. Os resíduos Glu235 e Glu340 presentes no sítio catalítico da GC estão representados na forma de bastão no centro da figura. A parte sacarídica está mostrada na forma de esferas cinza.

Muitas mutações no gene GBA anulam parcial ou completamente a atividade catalítica da GC, podendo também reduzir a estabilidade da enzima. A mutação p.N370S causa predisposição à DG tipo 1, o que indica que essa mutação causa poucas mudanças na estrutura e na ação catalítica da enzima. Esta mutação localiza-se na interface dos domínios II e III, estando bem distante do sítio ativo para participar diretamente da catálise. As mutações p.R463C (c.1504C>T) e p.R496H (c.1604G>A) estão localizadas no domínio II, sendo também associadas à DG tipo 1 (Dvir *et al.*, 2003).

A mutação p.L444P está localizada no domínio II da GC. A substituição de leucina por prolina no resíduo 444 da proteína provoca uma mudança conformacional, alterando o enovelamento desse domínio. A mutação p.L444P gera uma proteína instável, sugerindo uma importante função regulatória ou estrutural para o domínio II, talvez na interação com a SAP-C e/ou com fosfolipídeos (Dvir *et al.*, 2003).

1.4. Justificativa

A DG é a doença lisossômica de depósito mais frequente e seu diagnóstico se tornou muito relevante após a introdução da TRE no início dos anos 90. O nosso grupo realiza o diagnóstico da DG a vários anos e também desenvolveu protocolos para a identificação de mutações comuns no gene GBA (p.N370S, p.L444P, 84insG, IVS2+1G>A). Entretanto, a análise dessas mutações permite a identificação de aproximadamente 50% dos alelos mutantes em pacientes brasileiros com DG. Além da experiência do nosso laboratório, outro grupo de pesquisa brasileiro publicou dados semelhantes aos observados por nós através da avaliação de outros pacientes brasileiros com DG (Rozenberg *et al.*, 2006a; Rozenberg *et al.*, 2006b). Portanto, a introdução de uma

avaliação abrangente de toda a região codificante do gene GBA é importante para a identificação dos alelos mutantes, propiciando o conhecimento do espectro de mutações encontrado no nosso país. Além disso, essas informações podem também contribuir para novas correlações genótipo-fenótipo assim como para o melhor aconselhamento genético e/ou para a oferta de análises moleculares individualizadas para famílias em risco.

2. OBJETIVOS

2.1. Objetivo geral

Identificar mutações na região codificante do gene GBA de pacientes brasileiros com DG.

2.2. Objetivos específicos

- Introduzir protocolos para amplificação específica da região codificante e regiões adjacentes do gene GBA;
- Realizar o sequenciamento das regiões de DNA amplificadas das amostras de pacientes com DG;
- Identificar as variações de sequência (mutações ou polimorfismos) nas regiões analisadas;
- Avaliar o possível efeito patogênico e/ou deletério das variações encontradas por análises *in silico*.

3. RESULTADOS

Os resultados deste trabalho serão apresentados sob forma de um artigo que será submetido à publicação na revista *Blood Cells, Molecules and Diseases*, cujas instruções encontram-se em anexo.

3.1. Artigo científico

Identification of seven novel mutations in the glucocerebrosidase gene of Brazilian patients with Gaucher disease.

Identification of seven novel mutations in the glucocerebrosidase gene of Brazilian patients with Gaucher disease

Marina Siebert^{1,2,3}; Hugo Bock^{1,2,3}; Kristiane Michelin-Tirelli²; Janice Carneiro Coelho^{2,3,4}; Roberto Giugliani^{1,2,5}; Maria Luiza Saraiva-Pereira^{1,2,3,4}.

¹Laboratório de Identificação Genética, Centro de Pesquisa Experimental, Hospital de Clínicas de Porto Alegre, Brazil

²Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Brazil

³Programa de Pós-Graduação em Ciências Biológicas: Bioquímica, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Brazil

⁴Departamento de Bioquímica, Instituto de Ciências Básicas da Saúde, Universidade Federal do Rio Grande do Sul, Brazil

⁵Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Brazil

Corresponding author:

Maria Luiza Saraiva-Pereira, PhD

Serviço de Genética Médica – Hospital de Clínicas de Porto Alegre

Ramiro Barcelos, 2350 – CEP 90035-903

Porto Alegre, Rio Grande do Sul, Brazil

E-mail address: mlpereira@hcpa.ufrgs.br

Phone: +55 51 33598011

Fax: +55 51 33598010

Abstract

Gaucher disease (GD), an autosomal recessive disorder, is caused by deficiency of glucocerebrosidase (GC) due to mutations in the gene coding for this enzyme (GBA gene). In this study, we described a strategy to analyze the entire GBA coding region and applied this approach to 54 unrelated Brazilian patients with GD. Different techniques including long-range PCR, real time PCR using Taqman[®] assay, nested PCR and direct sequencing were employed to define sequence variations in the gene. We report here 7 novel mutations p.S125N (c.491G>A), p.F213L (c.756T>G), p.P245T (c.850C>A), p.W378C (c.1251G>C), p.D399H (c.1312G>C), 982-983insTGC (c.980_982dupTGC) and IVS10+1G>T (c.1505+1G>T) that are likely to be pathogenic in the GBA gene. This last alteration was found as a complex allele together with p.L461P mutation ([p.L461P + IVS10+1G>T]). We were also able to identify 24 different mutations that were previously reported by other groups. Our results denote a high mutation detection rate and contribute significantly to the panel of reported mutations. The identification of mutant alleles is crucial for improving the knowledge of GBA mutation spectrum and to the better understanding of molecular basis of GD.

Keywords: Gaucher disease; glucocerebrosidase; GBA gene; novel mutations.

Introduction

Gaucher disease (GD), the most common lysosomal storage disorder, results from an inborn deficiency of glucocerebrosidase (GC; E.C. 3.2.1.45). This enzyme is responsible for degradation of glucocerebroside (glucosyl ceramide) resulting in glucose and ceramide residues. The enzyme deficiency leads to accumulation of the undegraded substrate mainly within cells of monocyte/macrophage lineage. The metabolic overload is responsible for the main clinical manifestations of the disease. GD is an autosomal recessive disorder caused by mutations in the glucocerebrosidase (GBA) gene or, rarely, by mutations in the GC activator protein saposin C [1].

Three clinical types are distinguished on the basis of neurological involvement. Type 1 is the most common and shows a chronic non-neuronopathic course. Age of onset and severity of symptoms vary widely. Type 2 is the most severe form of the disease and shows an acute neuronopathic course, with onset within few months of age and rapidly progressive and devastating neurologic deterioration. GD type 3 shows a chronic neuronopathic course and is characterized by later onset and slow progressive neurological syndrome when compared to type 2 [1].

The gene encoding for GC (GBA; GenBank accession # J03059) is located on chromosome 1q21 and spans 7.6 kb of genomic DNA that is organized into 11 exons. In addition to the functional gene, a highly homologous pseudogene sequence (GBAP; GenBank accession # J03060) can be found 16 kb downstream [2,3]. Both gene and pseudogene are in the same orientation and share 96% exonic sequence homology [2,4]. To date, more than 250 mutations have been reported in the GBA gene. These include missense and nonsense mutations, small insertions or deletions that lead to either

frameshift or in-frame alterations, splice junction mutations, and complex alleles carrying two or more mutations in *cis* [5]. Frequency of specific mutant alleles can vary in different populations. In patients of Ashkenazi Jewish ancestry, four mutations account for nearly 90% of disease alleles. These mutations are p.N370S (c.1226A>G), p.L444P (c.1448T>C), 84insG (c.84dupG) and IVS2+1G>A (c.115+1G>A). Among non-Jewish patients, these mutations account for about 50 to 60% of GD associated mutations and a broad spectrum of mutations in the GBA gene can be responsible for the remaining mutant alleles [6,7].

Diagnosis of GD is established by the measurement of total GC activity in leukocytes from peripheral blood and/or fibroblasts from skin biopsies [1]. Molecular analyses are complementary to biochemical assay. Our group has been involved in biochemical diagnosis of Brazilian patients with GD in the last 20 years [8]. Few years later, we have also introduced molecular screening of common mutations (p.N370S, p.L444P, 84insG, and IVS2+1G>A). As expected, this approach was able to define roughly half of the mutant alleles in our sample population. Therefore, a more comprehensive scheme became essential to provide a clear picture of mutations' spectrum in our sample population.

We have then designed a strategy to analyze the entire GBA coding region and applied this approach to 54 unrelated Brazilian patients with GD that were referred to our group and were previously confirmed to be affected through biochemical analysis. We report here 7 novel mutations associated to GD as well as other rare mutations in the GBA gene.

Materials

Patients

We have included in this study 54 unrelated non-Jewish GD patients from different regions of Brazil. Gender distribution was 29 (53.7%) male patients and 25 (46.3%) females. Ages at diagnosis, when this data were available, varied from 2 months to 57 years. The inclusion criteria were (i) low GC activity in leukocytes and/or fibroblasts [8] and (ii) at least one unidentified disease causing allele following the screening of common mutations (p.N370S, p.L444P, 84insG, and IVS2+1G>A). This study was approved by the hospital ethical committee.

Sample collection and DNA isolation

Blood samples (5 mL) from each patient were collected in EDTA. Genomic DNA was isolated from peripheral blood leukocytes as described previously [9], and kept at -20°C until further analysis. DNA quantification was carried out using the fluorescence-based assay (Quant-It™ ds DNA BR Assay Kit – Invitrogen) in the Qubit™ fluorometer (Invitrogen).

Amplification of the entire GBA gene

Long-range PCR was applied to selectively amplify the functional GBA gene using primers GBALF (5' CGACTTTACAAACCTCCCTG 3') and GBALR (5' CCAGATCCTATCTGTGCTGG 3') that generate a fragment of 7765 bp. Long-range PCR reaction was performed in a final volume of 25 µL containing 12.5 ng of genomic DNA, 200 µM of each dNTP, 0.2 µM of each primer (forward and reverse), 60 mM of Tris-SO₄ (pH 9.1), 18 mM of (NH₄)₂SO₄, 1.7 mM of MgSO₄, and 1 µL of Elongase®

Enzyme Mix (Invitrogen). Amplification program was initial denaturation at 94°C for 5 minutes followed by 30 cycles of denaturation at 94°C for 30 seconds, annealing at 58°C for 30 seconds and elongation at 68°C for 8 minutes, and a final extension at 68°C for 10 minutes.

Screening of common mutations in the GBA gene

p.N370S and p.L444P mutations were screened by qualitative real time PCR using the TaqMan[®] PCR Assay (Applied Biosystems) that is based on fluorescent minor groove binding probes. Primers and probes were designed using PrimerExpress[™] software version 3.0 (Applied Biosystems) and sequence of each primer is shown in table 1. PCR reaction was performed in a final volume of 12 µL containing 0.5 µL of the long-range PCR product, 0.3 µL of specific TaqMan assay and 6 µL of 2x PCR Genotyping Master Mix (Applied Biosystems). Amplification included an initial step at 50°C for 2 minutes (activation of AmpErase UNG function), AmpliTaq[®] Gold activation at 95°C for 10 minutes, followed by 40 cycles of denaturation at 95°C for 15 seconds and annealing extension at 60°C for 1 minute. Allelic discrimination step was performed at 60°C for 1 minute. PCR products were analyzed by the Sequence Detection System Software version 1.2.1 (Applied Biosystems) in the ABI Prism 7500 Sequence Detector System (Applied Biosystems) through allelic discrimination plot.

Mutation 84insG was screened by Amplification Refractory Mutation System - PCR (ARMS-PCR). Two PCR reactions were separately performed using different specific primers to discriminate between wild and mutant alleles. Primers GAU-84GGWRT (5' GCATCATGGCTGGCAGCCTCACAGGACTGC 3') and GAU-2R (5' GCCCAGGCAACAGAGTAAGACTCTGTTTCA 3') were used for amplification of wild

allele, and GAU-84GGMTF (5' GCATCATGGCTGGCAGCCTCACAGGACTGG 3') and GAU-2R for mutant allele. Then, both reactions were performed for each sample and generate a fragment of 255 bp when the allele sequence was complementary to primer sequence. Each PCR reactions were performed in a total volume of 25 μ L, containing 200 ng of genomic DNA, 200 μ M of each dNTP, 1 μ M of each primer (forward and reverse), 20 mM of Tris (pH 8.4), 50 mM of KCl, 2.5 mM of MgCl₂ and 1.25 U of *Taq* DNA polymerase (Invitrogen). In this case, amplifications conditions were 96°C for 5 minutes, followed by 13 cycles of 96°C for 45 seconds, then 45 seconds at 70°C -0.5°C/cycle and 72°C for 45 seconds, then 22 cycles of 96°C for 45 seconds, 64°C for 45 seconds and 72°C for 45 seconds, and a final extension step at 72°C for 5 minutes. The amplified products were resolved by eletrophoresis on a 2% (w/v) agarose gel and visualized under UV light.

Mutation IVS2+1G>A was screened by PCR-RFLP strategy using primers GAU-84GGWRT and GAU-2R (sequences shown above) that generate a fragment of 255 bp. PCR reaction was performed in a total volume of 25 μ L, containing 200 ng of genomic DNA, 200 μ M of each dNTP, 1 μ M of each primer (forward and reverse), 20 mM of Tris (pH 8.4), 50 mM of KCl, 2.5 mM of MgCl₂ and 1.25 U of *Taq* DNA polymerase (Invitrogen). Amplification protocol was initial denaturation at 96°C for 5 minutes, 35 cycles of denaturation at 96°C for 45 seconds, annealing at 68°C for 45 seconds and extension at 72°C for 45 seconds, followed by final extension at 72°C for 10 minutes. Digestion reaction was performed in a total volume of 15 μ L, containing 10 μ L of PCR product in 20 mM of Tris-acetate, 50 mM of potassium acetate, 10 mM of magnesium acetate, 1 mM of dithiothreitol (pH 7.9) and 1 U of *Hph*I (New England Biolabs). Reaction was then placed at 37°C and left overnight. The digested product was then resolved by eletrophoresis on a 3% (w/v) agarose gel and visualized under UV light. As IVS2+1G>A

mutation destroys an *HphI* site, two smaller fragments (213 and 42 bp) are seen when in the presence of mutant allele.

PCR amplification and DNA sequencing

Exons 1 to 11 and flanking intron regions were amplified by nested PCR using long-range PCR product as template. GBA coding region were subdivided into 10 different amplicons, given that exons 10 and 11 were analyzed together. Briefly, PCR reaction was performed in a total volume of 25 μL , containing 0.5 μL of the long-range PCR product, 200 μM of each dNTP, 0.2 μM of each primer (forward and reverse), 20 mM of Tris (pH 8.4), 50 mM of KCl, 2.5 mM of MgCl_2 and 1.25 U of *Taq* DNA polymerase (Invitrogen). Amplification protocol was initial denaturation at 94°C for 10 minutes, 30 cycles of denaturation at 94°C for 30 seconds, annealing at 57-62°C for 30 seconds and extension at 72°C for 30 seconds, followed by final extension at 72°C for 10 minutes. An aliquot of PCR products was verified by electrophoresis on a 1.5% (w/v) agarose gel and visualized under UV light. Specific annealing temperature and primers' sequences are available on table 2.

Amplicons were purified using 2.5 U of Exonuclease I (USB) and 0.25 U of Shrimp Alkaline Phosphatase (USB). Direct sequencing was performed using BigDye Terminator Cycle Sequencing kit version 3.1 (Applied Biosystems) following the manufacturer's instructions and sequences were analyzed by DNA Sequencing Analysis Software version 5.2 (Applied Biosystems) in the ABI 3130xl genetic analyzer (Applied Biosystems). All identified mutations were confirmed by sequencing an independent DNA sample with forward and reverse primers.

GBA cDNA nucleotide positions are numbered from the first ATG initiator and the amino acid sequence is that of the mature protein after cleavage of the leader sequence. Following the current recommendation that all variants be described at the most basic level, mutations discussed in the text are referred by traditional name of the allele to aid in recognition following by cDNA nomenclature included within brackets [10].

Evaluation of novel mutations

Novel mutations detected in GBA coding region were assessed using two web-based tools PolyPhen (<http://genetics.bwh.harvard.edu/pph>) and SNPs3D (<http://www.snps3d.org>). Briefly, PolyPhen computes the impact of an amino acid substitution on a human protein by analyzing its structural and functional characteristics. The program generates a position-specific independent counts (PSIC) score for the two amino acids variants. If the difference between the calculated scores for these two amino acid variants (Δ PSIC) is ≤ 0.5 the impact of the amino acid substitution on the protein's function is likely to be benign [11].

SNPs3D combines two support vector machine (SVM) methods to assess the functional effect of a missense mutation. It computes the impact of an amino acid change on the structure and stability of the protein. In addition, it incorporates sequence conservation information by comparing the mutated amino acid position to ortholog sequences. The output is a SVM score. A positive SVM score classifies the amino acid change as non-deleterious, whilst a negative score indicates a deleterious effect on the protein's function. It has been noted that accuracy is significantly higher for SVM scores > 0.5 and < -0.5 [12].

GC sequences from 12 different species were searched using protein database from the National Center for Biotechnology Information (NCBI). Multiple alignment of protein sequences was performed by the ClustalW version 2.0 using FASTA format.

Results

This study led to the identification of 7 novel sequence variations, which 5 of them are likely to be disease causing alterations in the GBA gene including a novel complex allele. In addition to these changes, we were also able to find other 24 rare mutations. The molecular description and location of these 31 different variants are summarized in Table 3.

Novel alterations comprise 5 missense changes [p.S125N (c.491G>A), p.F213L (c.756T>G), p.P245T (c.850C>A), p.W378C (c.1251G>C), and p.D399H (c.1312G>C)], an in-frame mutation [982-983insTGC (c.980_982dupTGC)], and a splicing mutation as a complex allele [p.L461P + IVS10+1G>T]. These findings are described below.

First novel sequence alteration is caused by a G to A change at position 491 of cDNA located at exon 5 of the GBA gene and leads to a serine to asparagine substitution at residue 125 of protein (p.S125N). This mutation was found in a male type 1 GD patient from South region of Brazil that was diagnosed at 56 years of age and carries the p.N370S mutation on the other mutant allele.

In another type 1 GD patient we have found a T to G transversion at base 756 of cDNA in exon 6 of the gene that predicts a phenylalanine to leucine substitution at residue 213 of protein (p.F213L). The p.N370S mutation was also found in this compound heterozygous female type 1 GD patient from Southeast of Brazil whose diagnosis was established when she was 26 years old. No additional nucleotide change was detected in her whole GBA coding region.

We have defined two sequence variations in exon 7. First is a C to A change at position 850 of cDNA predicting a proline to threonine substitution at amino acid 245 of protein (p.P245T). This mutation was defined in a male type 3 GD patient from Southeast region of Brazil who is also heterozygous for *RecNciI* (c.1448T>C + c.1483G>C + c.1497G>C). No additional nucleotide change was detected in his whole GBA coding region. The other alteration in exon 7 is an in-frame insertion of three nucleotides (TGC) that predict a new leucine residue between positions 327 and 328 of the mature protein (Fig. 1). This mutation is referred to as 982-983insTGC and was found in a male type 1 patient also from Southeast of Brazil that was diagnosed at 31 years of age. This patient carries the p.N370S mutation in the other mutant allele.

We have also found two novel G to C nucleotide changes in exon 9. First transversion is placed at base 1251 of GBA cDNA and causes tryptophan to cysteine change at residue 378 of mature protein (p.W378C). This alteration was found in 6 unrelated alleles from 6 type 1 GD patients deriving from North or Northeast regions of Brazil. One of these patients carries the p.G377S (c.1246G>A) mutation in the other mutant chromosome while the remaining 5 patients are compound heterozygotes [p.N370S] + [p.W378C]. The other G to C transversion is located at position 1312 and is responsible for aspartate to histidine substitution at residue 399 of protein (p.D399H). This mutation was found in a type 1 GD patient diagnosed at 45 years of age that is also heterozygote for p.N370S mutation.

At last, another novel mutation is due to a G to T transversion in the splice donor site at first nucleotide of intron 10 (Fig. 1). This mutation, named IVS10+1G>T, destroys the conserved splice donor sequence and is probably responsible for a novel exon skipping event in this gene. This alteration was found in a female type 1 patient from South region

of Brazil that was diagnosed at 3 years of age. In addition to this finding, p.N370S and p.L461P (c.1499T>C) were also detected in this patient. As DNA from her parents were also available, we were able to determine that p.N370S mutation was inherited from paternal origin while both p.L461P and IVS10+1G>T were transmitted by her mother.

In silico evaluation of the 5 novel point sequence variations using PolyPhen and SNPs3D predicts deleterious effects on protein for p.S125N, p.W378C and p.D399H mutations (Table 4). These programs showed discordant results for p.F213L and p.P245T alterations. We have then considered that these substitutions cause non-deleterious effects to the protein in order to maintain a more conservative approach up to have data from the multiple alignment (Table 4).

Considering previously reported mutations, the most frequent detected among patients included in this study was the p.N370S mutation that was identified in 26 alleles (24.1%) as an individual mutation, and coexisting with the non-pseudogene-derived p.M361I (c.1200G>A) mutation in another one. The p.G377S was the second most frequent mutation accounting for 12 mutant alleles (11.1%), followed by p.I489T (c.1583T>C) mutation that was found in 7 alleles (6.5%). Mutations p.W378C and p.V398I (c.1309G>A) were present in 6 alleles (5.6%) each one. RecNciI was found in 3 alleles (2.8%) while p.L444P mutation was detected in 2 alleles (1.9%), and [p.E326K + p.L444P] was represented once (0.9%).

Majority of patients' genotype was defined by this study as shown in table 5. Genotypes that were found in more than one case were as follows: [p.N370S] + [p.W378C] in 5 cases; [p.G377S] + [p.G377S] in 4 cases; [p.V398I] + [p.V398I], [p.N370S] +

[p.I489T], and [p.R120W] + [p.N370S] in 3 cases each; and [793delC] + [p.N370S] in 2 cases. All remaining genotypes were represented by one case only.

Discussion

Protocol described here allowed the identification of 7 novel sequence variations in the GBA gene of patients with GD. These include 5 missense mutations (p.S125N, p.F213L, p.P245T, p.W378C, and p.D399H) in the GBA gene.

The p.S125N mutation implies the replacement of a polar uncharged amino acid with a hydroxyl group by an uncharged residue with amide group containing a long side chain. Ser125 is conserved in several species from human to worms, arguing in favor of a relevant role in the enzyme function (Fig. 2). This mutation destroys a *PvuII* restriction site that can be used for mutation screening.

The p.F213L mutation replaces a bulky aromatic amino acid by a residue containing an open aliphatic chain, which possibly changes the protein conformation. Although *in silico* evaluation predicted non-deleterious effect on GC protein, this phenylalanine residue is conserved in several species, such as rat, mouse, chimpanzee, orangutan, dog, pig, ox, and horse (Fig. 2). In addition, as stated previously, no additional mutation was identified in this GD patient.

The other mutation, p.P245T, denotes the replacement of a cyclic aliphatic structure by an open polar uncharged chain and would be predicted to have an effect on protein structure or activity. As mention above for p.F213L mutation, also for p.P245T, *in silico* evaluation predicted non-deleterious effect on mature protein, but this proline residue is conserved in several species, such as rat, mouse, chimpanzee, orangutan, dog, pig, ox, and horse (Fig. 2). Besides, no additional mutation was identified in this GD patient.

The mutation in exon 9, p.W378C, involves the replacement of an amino acid with an indole side chain by an amino acid with a sulfhydryl group, which possibly changes the protein conformation. A deleterious effect of this mutation was also indicated by *in silico* analysis. An alignment of GC from 12 different species (human, mouse, rat, chimpanzee, orangutan, dog, pig, ox, horse, tick, *C. elegans*, and honeybee) demonstrates that this tryptophan residue is highly conserved in these species, indicating a great degree of structural and functional relevance (Fig. 2). This mutation creates an *HpyCH4V* restriction site that can be used for mutation screening.

The p.D399H implies the replacement of an acidic amino acid by a basic one, which is a relevant change to affect structure and/or function of normal protein. As described above for p.W378C, Asp399 is highly conserved in 12 species and *in silico* analysis denotes deleterious effect of this mutation (Fig. 2). This mutation destroys a *HindIII* restriction site that can be used for mutation screening.

The additional variants are likely to have a pathological effect due to their nature: an insertion and a single base change in the first position of a splice donor site. Despite the insertion reported is a three bases in-frame mutation, the addition of a new residue in the mature protein is very likely to disturb normal folding and function of protein produced. And the point mutation located at first base of a splice donor site is also likely to disrupt normal splicing event at this position, given rise to exon skipping, alternative splicing or other different processing events.

Most novel sequence alterations detected by this study were found in patients that are compound heterozygous with the p.N370S mutation. In many cases, presence of p.N370S allele in any genotype could protect patients of carrying neurological involvement

[13]. In patients evaluated in this study whose genotype is a combination of p.N370S with a novel sequence variation no neurological involvement was reported. Therefore, independently of novel changes severity, patients' phenotypes indicate that p.N370S mutation is also acting as a protective factor in these cases. Similar observations have to be mentioned concerning combination of p.N370S with 982-983insTGC in a GD patient, who do not show neurological impairment, which further support the protective role of p.N370S mutation.

This study confirmed data previously reported in other Brazilian GD patients that p.G377S mutation is also frequently found among mutant alleles [14]. This data indicate the inclusion of this alteration in a preliminary screening in Brazil. This mutation was also reported to be the third most frequent among GD patients in Portugal and Spain [15]. The finding that homoallelic patients for p.G377S were essentially asymptomatic or had mild disease (type 1) suggests that this lesion is also neuroprotective [16]. However, others reports have also detected type 3 patients carrying one copy of p.G377S [14,17]. These observations suggest the possibility of an allele-dose effect, whereby two copies of p.G377S would lead to an enzymatic residual activity sufficient to prevent neurological involvement, but one copy of p.G377S with a second severe mutation would lead to type 3 GD. Following this concept the patient reported with [p.G377S] + [p.E326K + p.L444P] genotype is likely to develop neurological symptoms.

There are still mutant alleles that were not identified by the approach applied in this study. This fact is probably due to limitations inherent of the techniques employed here. For example, any allele-dose effect would not be picked out by approach described as well as any sequence alteration in the promoter region of the gene, in non-translated or in non-

coding regions that were not cover by PCR amplification scheme. We are currently testing further approaches to overcome these limitations and unmask the remaining mutant alleles.

The vast majority of mutations identified have not been found in homoallelic form, especially those new alterations. Ultimately, it may be more accurate to envision the associated phenotypes as a continuum spectrum with the major distinction being the presence and degree of neurological involvement [4,5].

Finally, the identification of mutant alleles is crucial for improving the knowledge of GBA mutation spectrum worldwide contributing to the better understanding of molecular basis of the disease. Furthermore, such information may also contribute to the establishment of genotype-phenotype correlations as well as for more specific genetic counseling and/or to offer a customized molecular analysis for families at risk.

Acknowledgements

The authors are grateful to patients and their families who kindly agreed to participate in this study. The authors would also like to thank Dr K. Michelin-Tirelli and Thais Santa Rita for their help with this work. MS, RG, and MLSP were supported by CNPq, Brazil. The work was also supported by Brazilian Funding Agencies (CNPq, FAPERGS and FIPE-HCPA) and Genzyme do Brazil.

References

- [1] E. Beutler, and G.A. Grabowski, Gaucher disease. in: C.R. Scriver, A.L. Beaudet, W.S. Sly, and D. Valle, (Eds.), *The metabolic and molecular bases of inherited disease*, McGraw-Hill, New York, 2001, pp. 3635-3668.
- [2] M. Horowitz, S. Wilder, Z. Horowitz, O. Reiner, T. Gelbart, and E. Beutler, The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics* 4 (1989) 87-96.
- [3] S.L. Winfield, N. Tayebi, B.M. Martin, E.I. Ginns, and E. Sidransky, Identification of three additional genes contiguous to the glucocerebrosidase locus on chromosome 1q21: implications for Gaucher disease. *Genome Res.* 7 (1997) 1020-6.
- [4] E. Sidransky, Gaucher disease: complexity in a "simple" disorder. *Mol. Genet. Metab.* 83 (2004) 6-15.
- [5] K.S. Hruska, M.E. LaMarca, C.R. Scott, and E. Sidransky, Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum. Mutat.* 29 (2008) 567-83.
- [6] G.A. Grabowski, and M. Horowitz, Gaucher's disease: molecular, genetic and enzymological aspects. *Baillieres Clin. Haematol.* 10 (1997) 635-56.
- [7] P. Alfonso, S. Aznarez, M. Giralt, M. Pocovi, and P. Giraldo, Mutation analysis and genotype/phenotype relationships of Gaucher disease patients in Spain. *J. Hum. Genet.* 52 (2007) 391-6.
- [8] K. Michelin, A. Wajner, F.T. de Souza, A.S. de Mello, M.G. Burin, M.L. Pereira, R.F. Pires, R. Giugliani, and J.C. Coelho, Application of a comprehensive protocol for the identification of Gaucher disease in Brazil. *Am. J. Med. Genet. A* 136 (2005) 58-62.
- [9] S.A. Miller, D.D. Dykes, and H.F. Polesky, A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.* 16 (1988) 1215.
- [10] J.T. den Dunnen, and S.E. Antonarakis, Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum. Mutat.* 15 (2000) 7-12.
- [11] V. Ramensky, P. Bork, and S. Sunyaev, Human non-synonymous SNPs: server and survey. *Nucleic Acids Res.* 30 (2002) 3894-900.
- [12] P. Yue, and J. Moul, Identification and analysis of deleterious human SNPs. *J. Mol. Biol.* 356 (2006) 1263-74.
- [13] P. Alfonso, A. Cenarro, J.I. Perez-Calvo, M. Giralt, P. Giraldo, and M. Pocovi, Mutation prevalence among 51 unrelated Spanish patients with Gaucher disease: identification of 11 novel mutations. *Blood Cells Mol. Dis.* 27 (2001) 882-91.
- [14] R. Rozenberg, F.T. Araujo, D.C. Fox, P. Aranda, A. Nonino, C. Micheletti, A.M. Martins, R. Cravo, E. Sobreira, and L.V. Pereira, High frequency of mutation G377S in Brazilian type 3 Gaucher disease patients. *Braz. J. Med. Biol. Res.* 39 (2006) 1171-9.
- [15] O. Amaral, E. Pinto, M. Fortuna, L. Lacerda, and M.C. Sa Miranda, Type 1 Gaucher disease: identification of N396T and prevalence of glucocerebrosidase mutations in the Portuguese. *Hum. Mutat.* 8 (1996) 280-1.
- [16] O. Amaral, A. Marcao, M. Sa Miranda, R.J. Desnick, and M.E. Grace, Gaucher disease: expression and characterization of mild and severe acid beta-glucosidase mutations in Portuguese type 1 patients. *Eur. J. Hum. Genet.* 8 (2000) 95-102.

[17] E. Beutler, T. Gelbart, and C.R. Scott, Hematologically important mutations: Gaucher disease. *Blood Cells Mol. Dis.* 35 (2005) 355-64.

Tables

Table 1. Sequences of primers and probes used to identify common mutations.

Mutation	Primer sequence (5' > 3')	Probe sequence (5' > 3')
p.N370S	Forward	Normal
	GCCTTTGTCTCTTTGCCTTTGTC	TTACCCTAGAACCTCCTG - VIC
	Reverse	Mutant
	CCAGCCGACCACATGGTA	ACCCTAGAGCCTCCTG - FAM
p.L444P	Forward	Normal
	CTGAGGGCTCCCAGAGAGT	CTGCGTCCAGGTCGT - VIC
	Reverse	Mutant
	GCCATCGGGATGCATCAGT	TGCGTCCGGGTCGT - FAM

Table 2. Primers used for amplification and sequencing of the GBA gene.

Exons and flanking regions	Primer sequence (5' > 3')		T _{annealing} (°C)	Lenght (bp)
	Forward	Reverse		
1	CCTAGTGCCTATAGCTAAGG	CTGGATTCAAAGAGAGTCTG	57	236
2	GTCCTAATGAATGTGGGAGACC	CTTACTGGAAGGCTACCAAAGG	61	286
3	G TTCAGTCTCTCCTAGCAGATG	GGAAACTCCATGGTGATCAC	61	353
4	GTCCTCCTAGAGGTAAATGGTG	GCAGAGTGAGATTCTGCCTC	61	316
5	GATAAGCAGAGTCCCATACTCTC	CTGTACAAGCAGACCTACCCTAC	62	281
6	CTAATGGCTGAACCGGATG	GGAAGTGGAAGTAGGTTGAGG	62	346
7	CAAAGTGCTGGGATTACAGG	CTCTAAGTTTGGGAGCCAGTC	62	404
8	CTAGTTGCATTCTTCCCGTC	GCTTCTGTCAGTCTTTGGTG	63	407
9	CTCCCACATGTGACCCTTAC	CTCGTGGTGTAGAGTGATGTAAG	61	329
10 and 11	GTGGGTGACTTCTTAGATGAGG	CTTTAGTCACAGACAGCGTGTG	62	473

Table 3. Alleles defined by this study. Novel sequence variations are shown in bold.

Allele ^a	cDNA nucleotide substitution ^b	Exon	Protein ^c	# of alleles
Substitutions				
p.R48Q	c.260G>A	3	p.Arg87Gln	1
p.R120W ^d	c.475C>T	5	p.Arg159Trp	3
p.M123T	c.485T>C	5	p.Met162Thr	1
p.S125N	c.491G>A	5	p.Ser164Asn	1
p.R131C	c.508C>T	5	p.Arg170Cys	1
p.W179X	c.653G>A	6	p.Trp218X	1
p.N188S ^d	c.680A>G	6	p.Asn227Ser	1
p.G202R ^d	c.721G>A	6	p.Gly241Arg	1
p.F213L	c.756T>G	6	p.Phe252Leu	1
p.P245T	c.850C>A	7	p.Pro284Thr	1
p.H311R	c.1049A>G	8	p.His350Arg	1
p.Y313H	c.1054T>C	8	p.Tyr352His	1
p.E349K	c.1162G>A	8	p.Glu388Lys	1
p.R353W	c.1174C>T	8	p.Arg392Trp	1
p.N370S	c.1226A>G	9	p.Asn409Ser	26
p.G377S	c.1246G>A	9	p.Gly416Ser	12
p.W378C	c.1251G>C	9	p.Trp417Cys	6
p.N396T	c.1304A>T	9	p.Asn435Thr	4
p.V398I	c.1309G>A	9	p.Val437Ile	6
p.D399H	c.1312G>C	9	p.Asp438His	1
p.L444P ^d	c.1448T>C	10	p.Leu483Pro	2
p.I489T	c.1583T>C	11	p.Ile528Thr	7
p.R496H	c.1604G>A	11	p.Arg535His	1
Insertions				
84insG	c.84dupG	2	p.Leu29AlafsX18	1
982-983insTGC	c.980_982dupTGC	7	p.Leu327_Pro328insLeu	1
Deletions				
413delC	c.413delC	4	p.Pro138LeufsX62	1
793delC	c.793delC	7	p.Gln265SerfsX5	2
Complex alleles				
p.E326K + p.L444P ^d	c.1093G>A	8	p.Glu365Lys	1
	c.1448T>C	10	p.Leu483Pro	
p.M361I + p.N370S	c.1200G>A	8	p.Met400Ile	1
	c.1226A>G	9	p.Asn409Ser	

	c.1448T>C	10	p.Leu483Pro	
RecNciI	c.1483G>C	10	p.Ala495Pro	3
	c.1497G>C	10	p.Val499Val	
p.L461P ^d + IVS10+1G>T	c.1499T>C	10	p.Leu500Pro	1
	c.1505+1G>T		Not applicable	

^a Nomenclature is given according to www.hgmv.org/mutnomen; ^b Nucleotides are numbered from A of first ATG; ^c Amino acid designations ("p.") are based on the primary GBA translation product, including the 39-residue signal peptide; ^d These alterations correspond to the normal pseudogene sequence.

Table 4. Novel missense sequence variations and evaluation of functional impact using web-based tools PolyPhen and SNPs3D.

Novel missense mutation	PolyPhen	SNP3D	Predict effect on GC
	Δ PSIC	SVM score	
p.S125N (c.491G>A)	1.755	- 2.60	deleterious
p.F213L (c.756T>G)	0.976	0.54	non-deleterious
p.P245T (c.850C>A)	1.639	0.77	non-deleterious
p.W378C (c.1251G>C)	4.114	- 3.26	deleterious
p.D399H (c.1312G>C)	2.321	- 1.98	deleterious

Δ PSIC \leq 0.5 = benign; SVM scores $>$ + 0.5 = benign.

Table 5. Genotypes defined by this study.

Genotype	# of patients
[p.N370S] + [p.L461P + IVS10+1G>T]	1
[p.Y313H] + [p.L444P]	1
[793delC] + [p.N370S]	2
[p.N370S] + [p.W378C]	5
[p.G377S] + [p.I489T]	1
[p.R353W] + [RecNciI]	1
[p.V398I] + [p.V398I]	3
[p.F213L] + [p.N370S]	1
[p.W179X] + [p.M361I + p.N370S]	1
[p.N370S] + [p.I489T]	3
[p.N188S] + [RecNciI]	1
[p.M123T] + [p.E349K]	1
[p.G377S] + [p.W378C]	1
[p.S125N] + [p.N370S]	1
[p.P245T] + [RecNciI]	1
[p.G377S] + [p.G377S]	4
[p.R120W] + [p.N370S]	3
[p.N396T] + [p.N396T]	1
[p.N370S] + [p.D399H]	1
[413delC] + [p.N370S]	1
[p.I489T] + [p.I489T]	1
[p.N370S] + [p.G377S]	1
[p.N396T] + [p.I489T]	1
[p.N396T] + [p.L444P]	1
[p.G202R] + [p.N370S]	1
[p.R131C] + [p.N370S]	1
[84insG] + [p.R496H]	1
[p.H311R] + [p.N370S]	1
[982-983insTGC] + [p.N370S]	1
[p.G377S] + [p.E326K + p.L444P]	1
[p.R48Q] + [?]	1
[p.N370S] + [?]	3
[?] + [?]	6
Total	54

? denotes unidentified mutation.

Figure captions

Figure 1. Direct DNA sequencing of the GBA gene. A) Direct sequencing of part of exon 10 from the forward primer. First arrow shows T to C substitution (c.1499T>C) that leads to p.L461P mutation, while second arrow shows G to T substitution responsible for IVS10+1G>T alteration. B) Direct sequencing of part of exon 7 from the forward primer. Dashed line denotes the location where 3 nucleotides (TGC) were inserted, leading to 982-983insTGC mutation.

Figure 2. Multiple alignment of GC in several species. The protein sequences shown in this alignment are (A) rat (Gene ID 684536); (B) mouse (Gene ID 14466); (C) human (Gene ID 2629); (D) chimpanzee (Gene ID 449571); (E) orangutan (Gene ID 100174563); (F) dog (Gene ID 612206); (G) pig (Gene ID 449572); (H) ox (Gene ID 537087); (I) horse (Gene ID 100063514); (J) tick (Gene ID 8051161); (K) *C. elegans* (Gene ID 178535); (L) honeybee (Gene ID 409708). The asterisks denote the residues where missense mutations characterized in this study are located. *1 denotes the p.S125N mutation; *2 denotes the p.F213L mutation; *3 denotes the p.P245T mutation; *4 denotes the p.W378C mutation; and *5 denotes the p.D399H mutation. The boxes show novel missense mutations with conserved residues of GC in several species.

Figures

Figure 1

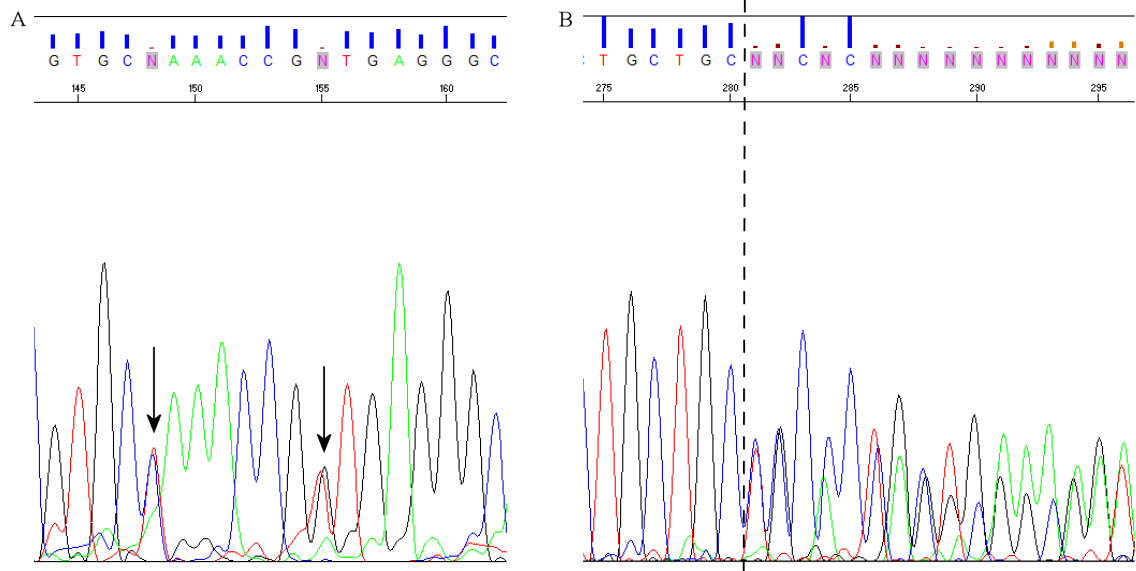
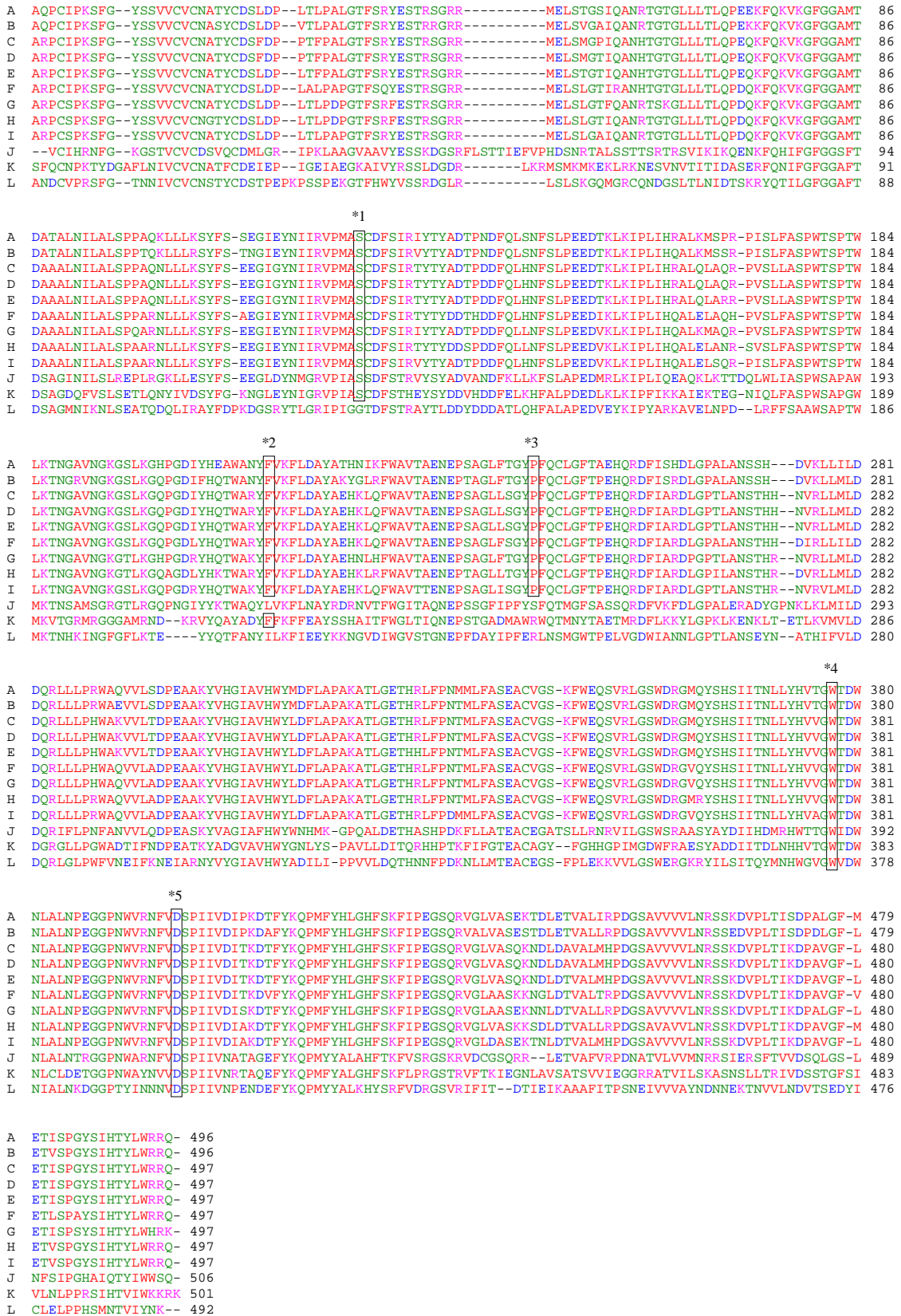


Figure 2



4. DISCUSSÃO

O gene GBA dos pacientes foi amplificado através da técnica de PCR longo, seguido de *nested* PCR. Os fragmentos obtidos foram analisados através de sequenciamento direto. Esta análise permitiu a identificação de 31 mutações diferentes, sendo que 7 delas (p.S125N, p.F213L, p.P245T, p.W378C, p.D399H, 982-983insTGC e p.L461P + IVS10+1G>T) foram descritas pela primeira vez neste trabalho. Cinco destas alterações (p.S125N, p.F213L, p.P245T, p.W378C e p.D399H) são mutações de sentido trocado (*missense*) no gene GBA.

A mutação p.S125N causa a substituição de um aminoácido polar com grupamento hidroxila por um resíduo não-carregado com grupamento amida contendo uma cadeia lateral mais longa. Ser125 é um resíduo conservado em várias espécies de humanos a vermes, destacando seu papel relevante na função da enzima. Esta mutação destrói um sítio de restrição *PvuII*, que pode ser usado para triagem desta mutação.

A mutação p.F213L substitui um aminoácido aromático por um resíduo contendo uma cadeia alifática aberta, que provavelmente altera a conformação da proteína. Embora a avaliação *in silico* tenha previsto um efeito não-deletério na proteína GC, este resíduo de fenilalanina é conservado em diversas espécies, tais como rato, camundongo, chimpanzé, orangotango, cão, porco, boi e cavalo. Além disso, como mencionado anteriormente, nenhuma outra mutação foi identificada nesse paciente com DG.

A outra mutação, p.P245T, determina a substituição de uma estrutura alifática cíclica por uma polar não-carregada aberta e que possivelmente tem efeito na estrutura ou na atividade da GC. Assim como mencionado para a mutação p.F213L, a análise *in silico*

da mutação p.P245T predisse um efeito não-deletério na proteína madura, embora este resíduo de prolina seja conservado em várias espécies, dentre elas rato, camundongo, chimpanzé, orangotango, cão, porco, boi e cavalo. Além disso, nenhuma mutação adicional foi identificada nesse paciente com DG.

A mutação no éxon 9 do gene GBA, p.W378C, causa a substituição de um aminoácido com uma cadeia lateral indol por um aminoácido com um grupamento sulfidríla, que possivelmente altera a conformação da proteína. A análise *in silico* indicou um efeito deletério para esta mutação. Um alinhamento de 12 espécies diferentes da GC (homem, rato, camundongo, chimpanzé, orangotango, cão, porco, boi, cavalo, carrapato, *C. elegans* e abelha) mostrou que o resíduo Trp378 é altamente conservado nestas espécies, indicando elevado grau de relevância estrutural e funcional. Esta mutação cria um sítio de restrição *HpyCH4V*, que pode ser usado para triagem desta mutação.

A mutação p.D399H determina a substituição de um aminoácido ácido por um básico na proteína, ocasionando uma mudança importante que pode afetar a estrutura e/ou função da GC. Como descrito para a p.W378C, o resíduo Asp399 também é altamente conservado em 12 espécies e a análise *in silico* indicou efeito deletério desta alteração. Esta mutação destrói um sítio de restrição *HindII*, que pode ser usado para triagem desta mutação.

As demais variantes provavelmente têm efeito patológico devido a sua natureza: uma inserção e uma mudança de um nucleotídeo na primeira posição do sítio doador de *splicing*. A inserção identificada é uma mutação *in-frame* de 3 nucleotídeos, que, ao adicionar um novo resíduo na proteína madura, provavelmente perturba o enovelamento normal e a função da proteína produzida. A mutação de ponto localizada no primeiro

nucleotídeo do sítio doador de *splicing* provavelmente também altere o evento de *splicing* normal nesta posição.

A maioria das novas variações de sequência identificadas neste estudo foi encontrada em pacientes heterozigotos com a mutação p.N370S. Em muitos casos, a presença do alelo p.N370S no genótipo protege os pacientes do desenvolvimento das manifestações neurológicas da doença (Alfonso *et al.*, 2001). Nos pacientes analisados neste estudo, aqueles que tinham genótipos p.N370S em combinação com uma variação de sequência nova não apresentaram comprometimento neurológico. Portanto, independentemente da gravidade da mutação nova, os fenótipos dos pacientes indicam que a mutação p.N370S está agindo como um fator de proteção nestes casos. Algumas considerações podem ser mencionadas em relação à combinação de p.N370S com 982-983insTGC em um paciente com DG, que não apresentou comprometimento neurológico apoiando o papel protetor da mutação p.N370S.

Este estudo confirmou dados anteriores, que estudaram outros pacientes brasileiros com DG, de que a mutação p.G377S é frequentemente encontrada entre os alelos mutantes (Rozenberg *et al.*, 2006a). Esses dados indicam que esta alteração deve ser incluída no teste de triagem preliminar no Brasil. Esta mutação também foi descrita como a terceira mais frequente entre pacientes com DG em Portugal e Espanha (Amaral *et al.*, 1996). Pacientes homozigotos para a mutação p.G377S são assintomáticos ou desenvolvem a forma leve da doença (tipo 1), sugerindo que esta alteração também é neuroprotetora (Amaral *et al.*, 2000). Entretanto, outros estudos identificaram pacientes com DG tipo 3 heterozigotos para mutação p.G377S (Rozenberg *et al.*, 2006a). Estas observações sugerem a possibilidade de um efeito dependente do número de alelos afetados, de modo que duas cópias de p.G377S produziram atividade residual suficiente para evitar o prejuízo ao SNC,

enquanto uma cópia de p.G377S em associação com uma mutação grave levaria à DG tipo 3. O paciente que apresenta o genótipo [p.G377S] + [p.E326K + p.L444P] provavelmente desenvolverá sintomas neurológicos.

Alguns alelos mutantes não foram identificados pela abordagem aplicada neste estudo. Este fato é devido, provavelmente, a limitações das técnicas empregadas. Por exemplo, um efeito dose-alelo não será obtido pela metodologia empregada, bem como, uma alteração na sequência da região promotora do gene, em regiões não-traduzidas ou em regiões não-codificantes que não foram amplificadas através da PCR. Atualmente, estamos desenvolvendo outras abordagens para superar estas dificuldades e, assim, identificar os alelos mutantes que permanecem desconhecidos.

A grande maioria das mutações identificadas não foi encontrada em pacientes homocigotos, especialmente as alterações novas. O mais indicado é considerar os fenótipos como uma continuidade de sintomas, distinguindo apenas em relação à presença e o grau de envolvimento neurológico (Sidransky, 2004; Hruska *et al.*, 2008).

A identificação dos alelos mutantes é importante para o conhecimento do espectro de mutações no nosso país e para aumentar o conhecimento das bases moleculares da doença. Além disso, essas informações podem também contribuir para a melhor compreensão de correlações genótipo-fenótipo, assim como, para o aconselhamento genético e/ou para a oferta de análises moleculares individualizadas para famílias em risco.

5. CONCLUSÕES

- Um protocolo, baseado na combinação de PCR longo e *nested* PCR, foi desenvolvido e padronizado para amplificação específica de toda a região codificante do gene GBA em 10 fragmentos.
- O sequenciamento direto das regiões amplificadas das amostras de DNA de 54 pacientes com DG foi realizado e permitiu a definição de alterações de sequência em 92 alelos.
- Sete mutações novas (p.S125N, p.F213L, p.P245T, p.W378C, p.D399H, 982-983insTGC e IVS10+1G>T) e outras 24 alterações descritas previamente por outros grupos foram identificadas no grupo de pacientes estudado.
- O efeito patogênico e/ou deletério das variações de sequência novas foi avaliado por análises *in silico* (programas de predição de efeito funcional e alinhamento de sequências) e todas elas apresentam resultados potencialmente compatíveis com alterações responsáveis por causar alteração na atividade enzimática.
- Os resultados obtidos contribuem para o aconselhamento genético e/ou para a oferta de análises moleculares individualizadas para famílias em risco.

6. REFERÊNCIAS

- Aerts JM, Hollak CE, Boot RG, Groener JE, Maas M (2006). Substrate reduction therapy of glycosphingolipid storage disorders. *J Inher Metab Dis*; 29(2-3): 449-456.
- Alattia JR, Shaw JE, Yip CM, Privé GG (2007). Molecular imaging of membrane interfaces reveals mode of β -glucosidase activation by saposin C. *PNAS*; 104(44): 17394-17399.
- Alfonso P, Cenarro A, Pérez-Calvo JI, Giralto M, Giraldo P, Pocoví M (2001). Mutation prevalence among 51 unrelated Spanish patients with Gaucher disease: Identification of 11 novel mutations. *Blood Cells Mol Dis*; 27(5): 882-891.
- Alfonso P, Aznarez S, Giralto M, Pocovi M, Giraldo P (2007). Mutation analysis and genotype/phenotype relationships of Gaucher disease patients in Spain. *J Hum Genet*; 52(5): 391-396.
- Alterini R, Rigacci L, Stefanacci S (1996). Pseudo-Gaucher cells in the bone marrow of a patient with centrocytic nodular non-Hodgkin's lymphoma. *Haematologica*; 81(3): 282-283.
- Amaral O, Pinto E, Fortuna M, Lacerda L, Sá Miranda MC (1996). Type 1 Gaucher disease: identification of N396T and prevalence of glucocerebrosidase mutations in the Portuguese. *Hum Mutat*; 8(3): 280-281.

- Amaral O, Marcão A, Sá Miranda M, Desnick RJ, Grace ME (2000). Gaucher disease: expression and characterization of mild and severe acid β -glucosidase mutations in Portuguese type 1 patients. *Eur J Hum Genet*; 8(2): 95-102.
- Beck M (2010). Therapy for lysosomal storage disorders. *IUBMD Life*; 62(1): 33-40.
- Berg-Fussman A, Grace ME, Ioannou Y, Grabowski GA (1993). Human acid beta-glucosidase. N-glycosylation site occupancy and the effect of glycosylation on enzymatic activity. *J Biol Chem*; 268(20): 14861-14866.
- Beutler E, Grabowski GA (2001). Gaucher disease. In: Scriver CR, Beaudet AL, Sly WS, Valle D, Childs B, Kinzler KW, Vogelstein B (eds): *The metabolic and molecular bases of inherited disease*. 8th ed. New York: McGraw-Hill; p. 3635-3665.
- Beutler E (2006). Gaucher disease: multiple lessons from a single gene disorder. *Acta Paediatr Suppl*; 95(451): 103-109.
- Blech-Hermoni YN, Ziegler SG, Hruska KS, Stubblefield BK, LaMarca ME, Portnoy ME, NISC Comparative Sequencing Program, Green ED, Sidransky E (2009). In silico and functional studies of the regulation of the glucocerebrosidase gene. *Mol Genet Metab*; 99(3): 275-282.
- Chen M, Wang J (2008). Gaucher disease: review of the literature. *Arch Pathol Lab Med*; 132(5): 851-853.
- Coelho JC, Giugliani R (2000). Fibroblasts of skin fragments as a tool for the investigation of genetic diseases: technical recommendations. *Genet Mol Biol*; 23(2): 269-271.

- Cormand B, Harboe TL, Gort L, Campoy C, Blanco M, Chamoles N, Chabás A, Vilageliu L, Grinberg D (1998). Mutation analysis of Gaucher disease patients from Argentina: high prevalence of the RecNciI mutation. *Am J Med Genet*; 80(4): 343-351.
- den Dunnen JT, Antonarakis SE (2000). Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat*; 15(1): 7-12.
- Dvir H, Harel M, McCarthy AA, Toker L, Silman I, Futerman AH, Sussman JL (2003). X-ray structure of human acid- β -glucosidase, the defective enzyme in Gaucher disease. *EMBO reports*; 4(7): 704-709.
- Emre S, Gürakan F, Yüce A, Rolf A, Scott R, Özen H (2008). Molecular analysis of Turkish Gaucher disease patients: Identification of novel mutations in glucocerebrosidase (GBA) gene. *Eur J Med Genet*; 51(4): 315-321.
- Germain DP (2004). Gaucher's disease: a paradigm for interventional genetics. *Clin Genet*; 65(2): 77-86.
- Grabowski GA, Horowitz M (1997). Gaucher's disease: molecular, genetic and enzymological aspects. *Baillière's Clin Haematol*; 10(4): 635-656.
- Grabowski GA (2004). Gaucher disease: lessons from a decade of therapy. *J Pediatr*; 144: S15-S19.
- Grabowski GA (2008). Phenotype, diagnosis, and treatment of Gaucher's disease. *Lancet*; 372(9645): 1263-1271.
- Harmanci O, Bayraktar Y (2008). Gaucher disease: new developments in treatment and etiology. *World J Gastroenterol*; 14(25): 3968-3973.

- Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E (1989). The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics*; 4(1): 87-96.
- Hruska KS, LaMarca ME, Scott CR, Sidransky E (2008). Gaucher disease: mutation and polymorphism spectrum in the glucocerebrosidase gene (GBA). *Hum Mutat*; 29(5): 567-583.
- Hughes DA (2009). Enzyme, substrate, and myeloma in Gaucher disease. *Am J Hematol*; 84(4): 199-201.
- Jmoudiak M, Futerman AH (2005). Gaucher disease: pathological mechanisms and modern management. *Br J Haematol*; 129(2): 178-188.
- KEGG: Kyoto Encyclopedia of Genes and Genomes. Disponível em: <<http://www.genome.jp/kegg/>>. Acesso em janeiro de 2010.
- Kolter T, Sandhoff K (2009). Lysosomal degradation of membrane lipids. *FEBS Lett*; doi: 10.1016/j.febslet.2009.10.021.
- Koprivica V, Stone DL, Park JK, Callahan M, Frisch A, Cohen IJ, Tayebi N, Sidransky E (2000). Analysis and classification of 304 mutant alleles in patients with type 1 and type 3 Gaucher disease. *Am J Hum Genet*; 66(6): 1777-1786.
- Liou B, Kazimierczuk A, Zhang M, Scott CR, Hedge RS, Grabowski GA (2006). Analyses of variant acid- β -glucosidases: effects of Gaucher disease mutations. *J Biol Chem*; 281(7): 4242-4253.

- Martins AM, Valadares ER, Porta G, Coelho J, Semionato Filho J, Pianovski MAD, Kerstenetzky MS, Montoril MF, Aranda PC, Pires RF, Mota RMV, Bortolheiro TC (2009). Recommendations on diagnosis, treatment, and monitoring for Gaucher disease. *J Pediatr*; 155(4 Suppl): S10-S18.
- Michelin K, Wajner A, Souza FTS, Mello AS, Burin MG, Pereira MLS, Pires RF, Giugliani R, Coelho JC (2005). Application of a comprehensive protocol for the identification of Gaucher disease in Brazil. *Am J Med Genet A*; 136(1): 58-62.
- Rozenberg R, Araújo FT, Fox DC, Aranda P, Nonino A, Micheletti C, Martins AM, Cravo R, Sobreira E, Pereira LV (2006a). High frequency of mutation G377S in Brazilian type 3 Gaucher disease patients. *Braz J Med Biol Res*; 39(9): 1171-1179.
- Rozenberg R, Fox DC, Sobreira E, Pereira LV (2006b). Detection of 12 new mutations in Gaucher disease Brazilian patients. *Blood Cells Mol Dis*; 37(3): 204-209.
- Sawkar AR, D'Haese W, Kelly JW (2006). Therapeutic strategies to ameliorate lysosomal storage disorders - a focus on Gaucher disease. *Cell Mol Life Sci*; 63(10): 1179-1192.
- Sidransky E (2004). Gaucher disease: complexity in a "simple" disorder. *Mol Genet Metab*; 83(1-2): 6-15.
- Sun Y, Liou B, Ran H, Skelton MR, Williams MT, Vorhees CV, Kitatani K, Hannun YA, Witte DP, Xu YH, Grabowski GA (2010). Neuronopathic Gaucher disease in the mouse: viable combined selective saposin C deficiency and mutant glucocerebrosidase (V394L) mice with glucosylsphingosine and glucosylceramide

accumulation and progressive neurological deficits. *Hum Mol Genet*; doi:
10.1093/hmg/ddp580.

Zimran A, Elstein D (2003). Gaucher disease and the clinical experience with substrate reduction therapy. *Phil Trans R Soc Lond B*; 358(1433): 961-966.

7. ANEXO

Instruções para submissão de manuscrito para *Blood Cells, Molecules and Diseases*

Guide for Authors

Subject Matter

Blood Cells, Molecules, and Diseases (BCMD) publishes original articles and reviews in the broad area of hematology, including both basic and clinical studies and encompassing areas such as hemostasis, thrombosis, immunology, metabolism, genetics, molecular biology, therapy, and morphology. Letters to the Editor are also welcome. Contributing Editors are encouraged to write or solicit a commentary on an article accepted for publication.

Manuscript Checklist

Size

Size limits are flexible within the approximate range of 1000 to 10,000 words. With figures, tables, and references appropriate to the subject matter, articles will, in general, not be longer than 20 journal pages. Letters to the Editor and Commentaries are limited to fewer than 1000 words.

Format

Please provide the following on the title page of original articles and reviews:

- Title (concise and informative)
- Author names and affiliation addresses (where the work was done)
- Corresponding author. Designate one author who will handle all correspondence at all stages of the process.

Followed by (in this order):

- The names and addresses of three potential referees.
- Abstract
- Keywords (maximum of 5)
- Introduction
- Materials
- Results
- Discussion
- Acknowledgments
- References
- Tables
- Figure captions
- Figures, schemes, illustrations (as separate files)

Manuscript Submission

Manuscript must be submitted through one of the contributing editors. Authors should contact one of the contributing editors to ascertain whether the editor is in a position to consider the paper. It is useful to do this by email or fax, including the summary of the paper so that the editor is apprised of the subject matter of the paper. Please access this link for contributing editor contact information:

http://www.elsevier.com/wps/find/journaleditorialboard.cws_home/622796/editorialboard#editorialboard

Because all papers are transmitted electronically (as well as set into print), all manuscripts must be submitted in electronic form, preferably MS Word for the text. Figures should also be sent in electronic format in a high-resolution format. They may be sent to the corresponding editor on disk by mail and it is helpful to include a hardcopy of the manuscript and figures.

Submitted articles receive only a light edit so that we will be able to implement immediate distribution. If the references or other parts of the paper are not in the correct format (see above) or if the paper is not written in good English, it will be returned to the author for correction.

Color illustrations are encouraged and will be included without additional charge in the electronic edition of the journal, but will be published as black-and-white versions in the printed journal unless the author pays for the cost of color reproduction. If the author wishes the printed edition to contain black-and-white versions of colored illustrations, the black-and-white versions must be supplied along with the color.

Should authors be asked by the Editor to revise their manuscripts, the revision will be due within 12 weeks. After this time, the manuscript will be considered a new submission.

Further considerations:

Upon acceptance of an article, authors will be asked to transfer copyright (for more information on copyright, see <http://authors.elsevier.com>). This transfer will ensure the widest possible dissemination of information. A letter will be sent to the corresponding author confirming receipt of manuscript. A form facilitating transfer of copyright will be provided after acceptance.

If material from other copyrighted works is included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases: contact Elsevier Global Rights

Department, PO Box 800, Oxford OX5 1DX, UK, phone: (+44) 1865 843830, fax: (+44) 1865 853333, email: permission@elsevier.com.

General Information:

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, without the written consent of the Publisher.

US National Institutes of Health (NIH) voluntary posting ("Public Access") policy
Elsevier facilitates author posting in connection with the voluntary posting request of the NIH (referred to as the NIH "Public Access Policy"; see <http://www.nih.gov/about/publicaccess/index.htm>) by posting the peer-reviewed author's manuscript directly to PubMed Central on request from the author, after formal publication. Upon notification from Elsevier of acceptance, we will ask you to confirm via e-mail (by e-mailing us at NIHauthorrequest@elsevier.com) that your work has received NIH funding (with the NIH award number, as well as the name and e-mail address of the Prime Investigator) and that you intend to respond to the NIH request. Upon such confirmation, Elsevier will submit to PubMed Central on your behalf a version of your manuscript that will include peer-review comments, for posting 12 months after the formal publication date. This will ensure that you will have responded fully to the NIH request policy. There will be no need for you to post your manuscript directly to PubMed Central, and any such posting is prohibited.

Authors' rights

As an author you (or your employer or institution) may do the following:

- make copies (print or electronic) of the article for your own personal use, including for your own classroom teaching use
- make copies and distribute such copies (including through e-mail) of the article to research colleagues, for the personal use by such colleagues (but not commercially or systematically, e.g., via an e-mail list or list server)
- post a pre-print version of the article on Internet websites including electronic pre-print servers, and retain indefinitely such version on such servers or sites
- post a revised personal version of the final text of the article (to reflect changes made in the peer review and editing process) on your personal or institutional website or server, with a link to the journal homepage (on <http://www.elsevier.com>)

- present the article at a meeting or conference and distribute copies of the article to the delegates attending such a meeting
- for your employer, if the article is a 'work for hire', made within the scope of your employment, your employer may use all or part of the information in the article for other intra-company use (e.g., training)
- retain patent and trademark rights and rights to any processes or procedure described in the article
- include the article in full or in part in a thesis or dissertation (provided that this is not to be published commercially)
- use the article or any part thereof in a printed compilation of your works, such as collected writings or lecture notes (subsequent to publication of your article in the journal)
- prepare other derivative works, extend the article into book-length form, or otherwise re-use portions or excerpts in other works, with full acknowledgment of its original publication in the journal

Preparation of text

General

Language polishing

Authors who require information about language editing and copyediting services pre- and post-submission may visit:

<http://www.elsevier.com/wps/find/authorshome.authors/languagepolishing>

or contact:

authorsupport@elsevier.com for more information.

Please note Elsevier neither endorses nor takes responsibility for any products, goods or services offered by outside vendors through our services or in any advertising. For more information please refer to our Terms & Conditions:

⇒ http://www.elsevier.com/wps/find/termsconditions.cws_home/termsconditions

Nomenclature and units

Follow internationally accepted rules and conventions: use the international system of units (SI). If other quantities are mentioned, give their equivalent in SI. For further information you are urged to consult:

IUPAC: Nomenclature of Organic Chemistry: ⇒ <http://www.iupac.org/>

IUPAC: Nomenclature of Inorganic Chemistry: ⇒ <http://www.iupac.org/>

References

Please ensure that every reference cited in the text is also present in the reference list (and vice versa). Any references cited in the abstract must be given in full. Unpublished results and personal communications are not recommended in the reference list, but may be mentioned in the text. If these references are included in the reference list, they should follow the standard reference style of the journal and should include a substitution of the publication date with either "Unpublished results" or "Personal communication." Citation of a reference as "in press" implies that the item has been accepted for publication.

Text

References should be numbered (numbers in square brackets) sequentially as they appear in the text. The actual authors can be referred to, but the reference number(s) must always be given.

Example: "..... as demonstrated [3,6]. Barnaby and Jones [8] obtained a different result"
Reference List: List all authors when there are four or fewer, when there are five or more, list the first three, then "et al."

Examples:

Reference to a journal publication:

[1] J. van der Geer, J.A.J. Hanraads, R.A. Lupton, et al. The art of writing a scientific article, *J. Sci. Commun.* 163 (2000) 51-59.

Reference to a book:

[2] W. Strunk Jr., E.B. White, *The Elements of Style*, third ed., Macmillan, New York, 1979.

Reference to a chapter in an edited book:

[3] G.R. Mettam, L.B. Adams, How to prepare an electronic version of your article, in: B.S. Jones, R.Z. Smith (Eds.), *Introduction to the Electronic Age*, E-Publishing Inc., New York, 1999, pp. 281-304.

Citing and listing of Web references

As a minimum, the full URL should be given. Any further information, if known (author names, dates, reference to a source publication, etc.), should also be given. Web references can be listed separately (e.g., after the reference list) under a different heading if desired, or can be included in the reference list.

Use of the Digital Object Identifier

The digital object identifier (DOI) may be used to cite and link to electronic documents. The DOI consists of a unique alphanumeric character string which is assigned to a document by the publisher upon the initial electronic publication. The assigned DOI never changes. Therefore, it is an ideal medium for citing a document, particularly 'Articles in press' because they have not yet received their full bibliographic information. The correct format for citing a DOI is:

doi:10.1016/j.bcnd.2003.10.071

When you use the DOI to create URL hyperlinks to documents on the web, they are guaranteed never to change.

Captions

Ensure that each illustration has a caption. Supply captions separately, not attached to the figure. A caption should comprise a brief title (**not** on the figure itself) and a description of the illustration. Keep text in the illustrations themselves to a minimum but explain all symbols and abbreviations used.

Preparation of electronic illustrations

General points

- Please use uniform lettering and sizing of your original artwork.
- Save text in illustrations as "graphics" or enclose the font.
- Only use the following fonts in your illustrations: Arial, Courier, Helvetica, Times, Symbol.
- Number the illustrations according to their sequence in the text.
- Use a logical naming convention for your artwork files.
- Provide all illustrations as separate files and as hardcopy printouts on separate sheets.
- Provide captions to illustrations separately.
- Produce images near to the desired size of the printed version.

A detailed guide on electronic artwork is available on our website at <http://authors.elsevier.com/artwork>

Please make sure that artwork files are in an acceptable format (TIFF, EPS or MS Office files) and with the correct resolution. If, together with your accepted article, you submit usable color figures then Elsevier will ensure, at no additional charge, that these figures will appear in color on the Web (e.g., ScienceDirect and other sites) regardless of whether or not these illustrations are reproduced in color in the printed version. For color reproduction in print, you will receive information regarding the costs from Elsevier after receipt of your accepted article. Please indicate your preference for color in print or on the Web only. For further information on the preparation of electronic artwork, please see <http://authors.elsevier.com/artwork>.

Please note: Because of technical complications which can arise by converting color figures to "gray scale" (for the printed version should you not opt for color in print), please submit in addition usable black and white versions of all the color illustrations.

Preparation of supplementary data

Elsevier accepts electronic supplementary material to support and enhance your scientific research. Supplementary files offer the Author additional possibilities to publish supporting applications, movies, animation sequences, high-resolution images, background datasets, sound clips and more. Supplementary files supplied will be published online alongside the electronic version of your article in Elsevier Web products, including ScienceDirect: <http://www.sciencedirect.com>. In order to ensure that your submitted material is directly usable, please ensure that data is provided in one of our recommended file formats. Authors should submit the material in electronic format together with the article and supply a concise and descriptive caption for each file. For more detailed instructions please visit our artwork instruction pages at <http://authors.elsevier.com>.

Proofs

One set of page proofs in PDF format will be sent by e-mail to the corresponding author, to be checked for typesetting/editing. No changes or additions to the manuscript will be allowed at this stage. Proofreading is solely the author's responsibility.