

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE BIOCÊNCIAS

DEPARTAMENTO DE GENÉTICA

**CARACTERIZAÇÃO MOLECULAR EM PACIENTES COM HEMOFILIA B NO
RIO GRANDE DO SUL**

MARIANA ROST MEIRELES

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

Orientador: Francisco Mauro Salzano

Co-orientadores: Gustavo Fioravanti Vieira, Eliane Bandinelli

Porto Alegre, março de 2017

INTITUIÇÕES FINANCIADORAS

Este trabalho foi desenvolvido no Laboratório de Hemostasia e Núcleo de Bioinformática do Departamento de Genética da Universidade Federal do Rio Grande do Sul e foi financiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e pela Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS), Programa de Apoio a Núcleos de Excelência (PRONEX).

A aluna recebeu auxílio financeiro por meio de uma bolsa concedida pelo CNPq, vinculada ao Programa de Pós-Graduação em Genética e Biologia Molecular da UFRGS.

AGRADECIMENTOS

Aos meus queridos professores orientadores, que foram grandes apoiadores em todas as fases do projeto. Por todos os ensinamentos e amizade. A Eliane Bandinelli, Gustavo Fioravanti Vieira e Francisco Mauro Salzano: meu respeito e admiração a vocês.

Aos colegas que passaram pelo Laboratório de Hemostasia, que tornaram o dia-a-dia muito mais animado, e dividiram experiências comigo. Em especial às amigas Ana Paula Maciel Ornaghi e Mariana Rodrigues Botton.

À minha família, que está sempre ao meu lado, e me transmite segurança para que eu siga meus sonhos. Meus pais, meus irmãos e meu namorado.

SUMÁRIO

LISTA DE ABREVIATURAS E SÍMBOLOS.....	6
LISTA DE FIGURAS.....	7
LISTA DE TABELAS.....	8
RESUMO.....	9
ABSTRACT.....	11
CAPITULO 1 – INTRODUÇÃO E OBJETIVOS.....	13
1.1. Hemostasia.....	13
1.2. A Via Clássica da Coagulação Sanguínea.....	15
1.2.1. A via intrínseca.....	17
1.2.2. A via extrínseca.....	17
1.2.3. A via comum.....	18
1.3. O Modelo Celular da Coagulação Sanguínea.....	19
1.3.1. Iniciação.....	21
1.3.2. Amplificação.....	21
1.3.3. Propagação e Finalização.....	22
1.4. O fator IX da Coagulação.....	22
1.5. O gene F9.....	27
1.6. Hemofilia B.....	28
1.6.1. Mutações no F9 e a hemofilia B.....	29
1.7. Relação estrutura-função: Bioinformática-estrutural.....	35
1.8. Objetivos.....	36
1.8.1. Objetivo geral.....	36
1.8.2. Objetivos específicos.....	36
1.9. Detalhes do projeto.....	37

CAPITULO 2 – MOLECULAR CHARACTERIZATION OF HAEMOPHILIA B IN A SOUTHERN BRAZIL POPULATION.....	38
CAPITULO 3 – ELECTROSTATIC ANALYSIS OF HAEMOPHILIA B CAUSATIVE MUTATIONS.....	55
CAPITULO 4 – DISCUSSÃO GERAL.....	75
CAPITULO 5- CONCLUSÕES E PERSPECTIVAS.....	79
REFERÊNCIAS BIBLIOGRÁFICAS.....	80
ANEXOS.....	86

LISTA DE ABREVIATURAS E SÍMBOLOS

ADP= difosfato de adenosina

EGF= *epidermal growth factor- domains*

EP= *electrostatic potential* – potencial eletrostático

FI, FII, FIII, FIV, FV, FVI, FVII, FVIII, FIX, FX, FXI, FXII = Fator I, II, III, IV (...) e assim por diante, respectivamente

F9= Gene do fator IX

FT= fator tissular

FvW= fator Von Willebrand

FIIa, FIIIa...= o sufixo “a” indica a ativação do fator (FVIIa= fator VII ativado)

GLA= *gamma-carboxyglutamic (GLA) domain*

GPI= glicoproteína I

HMWK= *high molecular weightkininogen* –cininogênio de alto peso molecular

HA= hemofilia A

HB= hemofilia B

MHC= Complexo Principal de Histocompatibilidade

PACE=*pairedbasic amino acidcleavingenzyme*

PE= potencial eletrostático

PCR= *Polimerase chainreaction* – reação em cadeia da polimerase

SAP= *Shrimpalkalinephosphatase* – fosfatase alcalina do camarão

TTPA= tempo de tromboplastina parcial ativado

TT= tempo de trombina

UTR= *Untranslatedregions*

kDA= kilodaltons

µg= micrograma

µl= microlitro

LISTA DE FIGURAS

Figura 1. O modelo clássico da cascata da coagulação sanguínea.....	16
Figura 2. O modelo celular da coagulação sanguínea.....	20
Figura 3. Fases do FIX: do gene à ativação da proteína.....	23
Figura 4. Modificações pós-traducionais do fator IX.....	26
Figura 5. A estrutura eixo principal do FIX ativado.....	26
Figura 6. Mapa das mutações de ponto nos resíduos do FIX.....	30

LISTA DE TABELAS

Tabela 1. Frequência dos tipos de mutação no F9.....	31
Tabela 2. Distribuição das mutações descritas no banco de dados do FIX.....	32
Tabela 3. Países com casos reportados em <i>Haemophilia B MutationDatabase</i>	33 a 34
Tabela 4. Número de indivíduos relatados nos países com maior ocorrência de casos.....	34

RESUMO

A hemofilia B (HB) (OMIM 306900) consiste em uma doença da coagulação de herança recessiva ligada ao X e afeta 1 a cada 30.000 nascimentos no sexo masculino. Essa doença genética é causada por mutações no gene do fator IX (F9) da coagulação. Existem cerca de 1095 mutações diferentes descritas em HB, com a prevalência de mutações do tipo *missense* (sentido trocado). Essas variantes levam a sangramentos, especialmente após injúrias vasculares ou cirurgias. A gravidade dos sangramentos está relacionada com a magnitude da deficiência do fator. A forma como as mutações interferem na função do FIX não é claramente determinada; diferentes aspectos poderiam estar envolvidos, como modificações estruturais, perda de estabilidade e mudanças eletrostáticas.

Este estudo tem o objetivo de (a) identificar mutações no gene F9 em pacientes com hemofilia B leve, moderada e grave, vivendo no sul do Brasil, estado do Rio Grande do Sul; (b) analisar a função das mutações na determinação da doença e fazer correlações entre o genótipo e o fenótipo, buscando regiões, domínios e aminoácidos que são determinantes para a função do FIX; e (c) investigar efeitos estruturais e eletrostáticos no FIX, de mutações previamente detectadas como causa da hemofilia B.

Um total de 43 pacientes independentes diagnosticados com HB foi incluído no estudo. A atividade de coagulação do fator IX foi medida através de ensaios de coagulação, e os pacientes foram classificados conforme os testes laboratoriais em leves (n=6), moderados (n=7) e graves (n=30). O DNA genômico foi extraído e todos os éxons do F9, regiões 5'UTR e 3'UTR, junções éxon-íntron e região promotora foram amplificados por PCR. Foi realizado o seqüenciamento direto para a detecção de mutações e as seqüências dos pacientes foram comparadas com a seqüência do gene F9 normal utilizando a ferramenta CodonCodeAligner, implementado no software MEGA 5.

As variações encontradas foram comparadas com aquelas presentes no *Haemophilia B Mutation Database* (<http://www.factorix.org/>) e quando não estavam presentes nesse, nas ferramentas *1000 Genomes Browser* e *USCS Genome Browser*, para verificar se eram polimorfismos. As mutações *missense* foram analisadas utilizando o BLASTP e *UNIPROT/KB/Swiss-Prot Mammalian database*, comparando resíduos do FIX

humano com seqüências do FIX ortólogas depositadas no mesmo banco de dados. Cinco algoritmos foram empregados para verificar o impacto das mutações *missense* encontradas na estrutura da proteína: PolyPhen-2 (*Polymorphism Phenotyping-2*), SiFT (*Sorting Intolerant to Tolerant*), HOPE (*Have yOur Protein Explained*), PROVEAN (*Protein Variation Effect Analyzer*), and *Mutation Taster*.

Um total de 54 diferentes mutações *missense* foi usado em análises de estrutura, 18 delas oriundas do nosso estudo em pacientes do sul do Brasil e as outras 36 descritas no Banco de Dados do Fator IX. Foram gerados modelos estruturais para cada mutação através de modelagem por homologia por meio do *Phyre-2*. Os modelos foram utilizados em análises de potencial eletrostático e em clusterizações por diferença eletrostática, com o objetivo de fazer uma comparação dos modelos mutantes com a estrutura e características físico-químicas do Fator IX normal. Os cálculos de potencial eletrostático (PE) foram realizados no *software Delphi webserver* e aplicados nos modelos por meio da interface *Chimera*. Para a clusterização das estruturas foi utilizada a ferramenta webPIPSA.

Foram encontradas 31 variantes distintas entre as 43 famílias HB, incluindo 6 mutações que não estavam descritas no Banco de Dados do Fator IX. Nenhuma das variantes foi encontrada nos bancos de dados *1000 Genomes Browser* e *UCSC Genome Browser*. Entre as variantes encontradas existe a predominância de mutações de sentido trocado (23/31), seguida de sem sentido (4/31), processamento (2/31) e pequenas deleções (2/31).

A maioria das mutações detectadas foi considerada como deletéria pelas predições dos algoritmos. Os tipos de mutações encontradas no Rio Grande do Sul apresentam uma distribuição que não se distancia daquelas observadas em pesquisas prévias. Considerando os modelos usados nas análises de clusterização, a análise de diferenças eletrostáticas mostrou que 31 dos 54 modelos analisados são diferentes da estrutura selvagem. A distância eletrostática dessas 31 mutações em relação à estrutura normal varia conforme a mutação e poderia interferir com as funções do fator IX, causando HB.

ABSTRACT

Hemophilia B (HB) (OMIM 306900) consists of an inherited X-linked recessive bleeding disorder and affects 1 in 30,000 male live births. This genetic disease is caused by mutations in the coagulation factor IX gene (F9). There are about 1,095 different HB mutations described, with a prevalence of missense mutations. These variations lead to a bleeding tendency, especially after injuries or surgery. The severity of bleeding symptoms is connected with the magnitude of the clotting factor deficiency. The role of how mutations interfere on FIX function is not clearly determined; different aspects could be involved, as structural modifications, loss of stability and electrostatic changes.

This study has the objectives of (a) identify F9 mutations in patients with mild, moderate and severe hemophilia B living in the southern Brazilian state of Rio Grande do Sul; (b) analyze the role of the mutations in disorder determination and make phenotype-genotype correlations, looking for the regions, domains and amino acids that are determinants of FIX function; (c) investigate FIX structural and electrostatic effects of previously detected and now hemophilia B causative missense mutations.

A total of 43 independent patients affected by HB were included in this study. Coagulation activity (FIX: C) was measured using the one-stage clotting assay, and then the patients were classified according to the laboratory tests into mild (n= 6), moderate (n= 7) and severe (n= 30) hemophilias. Genomic DNA was extracted and all F9 exons, the 5'UTR and 3'UTR, intron-exon junctions and the promoter region were amplified by PCR. Direct sequencing was conducted for mutation detection. Patients' sequences were compared with the normal *F9* gene sequence (GeneBank accession number K02402) using the Codon Code Aligner tool implemented with the MEGA 5 software.

The variations detected were compared with those reported in the Haemophilia B Mutation Database (<http://www.factorix.org/>) and when not found there, in the 1000 Genomes browser and USCS Genome browser to verify if they were polymorphisms. The missense mutations were examined using BLASTP and the UNIPROTKB/Swiss-Prot Mammalian database, comparing the human FIX residues (Swiss Prot ID: P00740) with orthologous FIX sequences deposited in the same databank. Five algorithms were

employed to analyze the impact of the missense mutations found on protein structure: PolyPhen-2 (Polymorphism Phenotyping-2), SiFT (Sorting Intolerant to Tolerant, HOPE (Have yOur Protein Explained), PROVEAN (Protein Variation Effect Analyzer), and Mutation Taster.

A total of 54 different HB missense mutations were employed in the structural analysis, 18 of them detected in our study of southern Brazilian patients and the other 36 described in the factor ix database. A structural model for each missense mutation was generated by homology approaches using Phyre2. The models were employed in electrostatic potential analyses and clusterization by electrostatic differences using bioinformatics tools, aiming to compare the mutant models with the structure and physicochemical characteristics of normal Factor IX. The electrostatic potential (EP) calculations were performed using the Delphi webserver and applied on the models using Chimera Interface. For the models' clusterization the webPIPSA tool was employed.

Thirty-one non-redundant variations were found among the 43 HB families, including 6 mutations not previously described in the Factor IX mutation database. None of the variations were found on the 1000 Genomes Browser or UCSC Genome Browser. Between the variations detected, there is a predominance of missense (23/31), followed by nonsense (4/31), splicing site (2/31) and small deletion (2/31) mutations.

The majority of the detected mutations were considered deleterious by algorithm predictions. The different mutation types detected in Rio Grande do Sul presented a distribution not much different from those observed in previous surveys. Considering the models used in the clusterization analysis, the electrostatic differences evaluation revealed that 31 of the 54 mutations were considered as divergent, compared to the wild structure. The electrostatic distances of these 31 mutations compared to normal vary according to the mutation and could have interfered in Factor IX functions, causing HB.

CAPÍTULO 1 – INTRODUÇÃO E OBJETIVOS

1.1.Hemostasia

A hemostasia consiste em um processo fisiológico para evitar sangramentos hemorrágicos, mantendo o sangue fluido nos vasos sanguíneos e evitando hemorragias e trombozes, além de ser crucial para o início do reparo tecidual(Schenone et al., 2004). Esse mecanismo é altamente regulado e dinâmico. As vias da coagulação sanguínea fazem a regulação fina do que existe neste processo, evitando que haja algum tipo de colapso. Erros em qualquer etapa desta via podem comprometer a coagulação.

O processo hemostático envolve desde a formação de um coágulo (coagulação) até a sua degradação (fibrinólise)(Marcus e Safier, 1993). Após a injúria vascular, a hemostasia ocorre por meio de etapas sucessivas: vasoconstrição, quando o estímulo do vaso traumatizado determina a contração da parede; formação de um tampão plaquetário (o qual envolve a adesão, ativação e agregação das plaquetas); ativação da cascata de coagulação sanguínea, que culmina na produção de uma rede de fibrina; e a dissolução desta rede, a fibrinólise. Diversas rupturas vasculares ocorrem diariamente, e em sua maioria a oclusão pelo tampão plaquetário é um mecanismo eficaz no controle hemostático, considerando que são apenas pequenas descontinuidades no vaso.

A coagulação sanguínea é iniciada por substâncias ativadoras do endotélio vascular, da superfície plaquetária e também por proteínas da coagulação sanguínea.Ela é regulada por três componentes básicos: a parede vascular, as plaquetas e os fatores da coagulação.

Em um vaso sanguíneo ocorre a migração axial durante a circulação. Em função disso, a maioria dos eritrócitos movimenta-se pela área central do lúmen, deixando eritrócitos livres próximos à parede dos vasos, onde estão as plaquetas. Quando ocorre dano em um vaso sanguíneo, as células sanguíneas (principalmente as plaquetas) tornam-se expostas ao colágeno e a fibroblastos do endotélio subjacente. Ocorre a vasoconstrição, diminuindo o fluxo sanguíneo no local de injúria vascular, para minimizar a perda de sangue. A etapa seguinte é a hemostasia primária, que consiste no processo que leva à formação do tampão plaquetário.

Durante a hemostasia primária, o Fator Von Willebrand (FvW) faz uma ligação com o colágeno subendotelial e o endotélio lesado. O FvW consiste em uma glicoproteína multimérica, que liga-se a uma glicoproteína presente na membrana superficial das plaquetas (GPIb). Inicialmente o FvW está em uma conformação enovelada, mas com a ligação ao colágeno subendotelial e às células endoteliais lesadas, ocorre o desenovelamento deste fator. A partir deste desenovelamento ocorre a exposição dos domínios A1 do FvW, que fazem a ligação com o GPIb. A ligação do domínio A1 com GPIb, ativa outro receptor nas plaquetas, a integrina GPIIb/IIIa. A GPIIb/IIIa ativa liga-se ao domínio C1 do FvW e leva à auto-ativação das plaquetas.

O Fator Von Willebrand faz a integração das plaquetas com o endotélio. A ausência, diminuição ou mesmo problemas funcionais do Fator Von Willebrand levam à doença de Von Willebrand, que consiste em uma disfunção onde as plaquetas apresentam dificuldade em aderir ao subendotélio.

Subsequentemente à adesão das plaquetas ao subendotélio, novas plaquetas são ativadas e se aglutinam. No processo de ativação, as plaquetas têm a sua configuração alterada: tornam-se esféricas e emitem pseudópodos. Com a ativação, os trombócitos liberam grânulos citoplasmáticos (como ADP, tromboxano e serotonina) e estes recrutam novas plaquetas para o local da lesão vascular (Marcus e Safier, 1993). Células endoteliais fazem a sinalização para controlar este processo de auto-amplificação, visando conter o tampão plaquetário estritamente na região do dano vascular (Stassen et al., 2004). Além disso, mecanismos que restringem a agregação plaquetária e formação do coágulo de fibrina à região da injúria contribuem para a manutenção do sangue em um estado fluido (Hoffbrand, Pettit e Moss, 2001). Para que a coagulação sanguínea aconteça, um polímero de fibrina (coágulo) precisa ser formado a partir de fibrinogênio, por meio de uma reação catalisada pela trombina, enzima formada a partir da protrombina, por uma reação estimulada por cálcio e tromboquinase (Riddel et al., 2007).

A interação entre as plaquetas e os fatores da coagulação sanguínea é o que dá origem a um tampão plaquetário, que acaba por estancar o sangramento no local da injúria vascular. A plaqueta adere-se a macromoléculas do tecido subendotelial na região lesada formando um tampão primário (Ruggeri e Mendolicchio, 2007). De maneira geral, as plaquetas estimulam a ativação desses fatores plasmáticos da coagulação, o que culmina na formação de um coágulo de fibrina, reforçando a agregação plaquetária (Berndt et al.,

2014). À medida que ocorre a cicatrização na área do dano, o coágulo de fibrina é desfeito, bem como a agregação plaquetária.

As plaquetas, ou trombócitos, têm sua biogênese na medula óssea, consistindo em fragmentos anucleados de megacariócitos (Zajicek, 1954). Elas são recrutadas especialmente para a coagulação sanguínea, participando fisicamente e quimicamente deste processo. São ativadas por exposição ao colágeno e aderem-se ao subendotélio lesado, onde se agregam umas às outras e formam um trombo, que é posteriormente estabilizado. As plaquetas consistem em peças chave da hemostasia primária e em um indivíduo adulto e saudável estão presentes em quantidades que vão aproximadamente de 150 a 400 bilhões de unidades por litro de sangue. Em eventos de injúria vascular, as plaquetas são altamente especializadas para reconhecer perturbações junto às células endoteliais dos vasos sanguíneos ou exposição da matriz fibrosa (Berndt et al., 2014).

1.2.A Via Clássica da Coagulação Sanguínea

Para simplificar o entendimento do processo de coagulação, foi proposto um modelo. Este modelo foi denominado de “Cascata da Coagulação Sanguínea”, conhecido também como “via clássica”, pois foi o primeiro modelo que estudou o papel dos fatores da coagulação em um sistema. Este conceito envolve uma série de etapas nas quais a ativação de cada um dos fatores da coagulação sanguínea leva à ativação de outro, e resulta na geração de trombina (Broze, 1995). O processo envolve uma via intrínseca e outra extrínseca, além de uma via comum. A ativação do fator X e a geração de fibrina podem ocorrer por ambas as vias (intrínseca e extrínseca).

Cerca de 20 fatores plasmáticos atuam na cascata da coagulação, ea maioria deles consiste em enzimas que circulam em um estado inativo (como zimogênios) na circulação e são ativados em cascata (**Figura 1**). Os fatores da coagulação foram descobertos em sua maioria entre os anos de 1940 e 1960, e foram identificados por algarismos romanos, que não ilustram a sequência das reações, mas sim a ordem pela qual os componentes da via foram sendo descobertos. O modelo de cascata só foi sugerido depois que já eram conhecidos os fatores da coagulação. A via clássica foi proposta simultaneamente por dois grupos diferentes de bioquímicos que publicaram artigos no ano de 1964. O artigo com a cascata de coagulação proposta por Macfarlane (1964) foi publicado na revista *Nature*, e

subsequente a ele, Davie e Ratnoff (1964) publicaram sobre o mesmo tema na *Science* (Riddel et al., 2007).

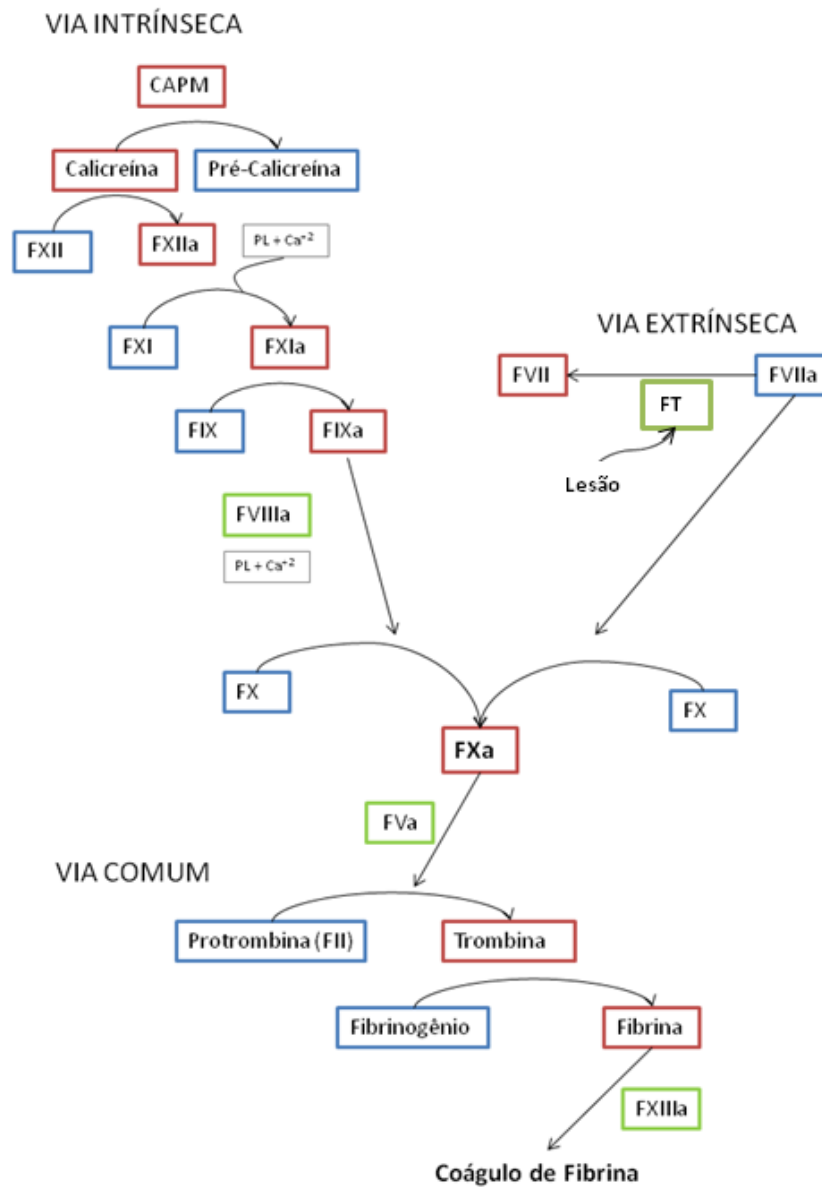


Figura1. O modelo clássico da cascata da coagulação sanguínea. Representação das vias intrínseca e extrínseca da coagulação. **Figura original.**

1.2.1. A via intrínseca

A via intrínseca consiste em uma série de reações dos fatores da coagulação. Essa rota envolve os Fatores XII, XI, VIII, IX, X e V, protrombina e fibrinogênio. Ela se inicia quando a proteína plasmática FXII (Fator Hageman) é ativada (FXIIa) pelo contato com uma superfície carregada negativamente, e culmina na geração de fibrina a partir de fibrinogênio e formação de uma rede de fibrina (Broze, 1995). Como a via depende de um fator circulante (Fator XII), ela ocorre de forma mais lenta.

O FXII é ancorado à superfície com o auxílio de um cofator, o cininogênio de alto peso molecular (HMWK- *high molecular weightkininogen*). Quando uma quantidade significadiva de FXIIa se acumula, este faz a conversão de pré-caliceína em caliceína, usando o HMWK como âncora e estacaliceína estimula a ativação do próprio FXII. O FXII ativado cliva o FXI em FXIa, que, por sua vez ativa o FIX (Di Scipio et al., 1978). O FIX ativado, em conjunto com o FXa e a trombina clivam o FVIII para sua forma ativa, que será um co-fator para a próxima reação.

O FIXa forma o complexo tenase ao se juntar como FVIIIa, Ca^{2+} e fosfolipídeos carregados negativamente (constituintes das membranas celulares). Este complexo (Autin et al., 2005; Bloem et al., 2013) ativa o FX e este liga-se ao FVa e tem atividade de protrombinase, convertendo a protrombina em trombina, que age gerando fibrina a partir de fibrinogênio (Riddel et al., 2007).

1.2.2. A via extrínseca

O termo via extrínseca é usado para descrever a resposta *in vivo* à lesão tissular. Nesta via, o fator X é ativado diretamente pela interação entre o FVII, FT e Ca^{2+} . A rota envolve o FT (um fator não circulante presente nas superfícies celulares), FVII, FX, FV, protrombina e fibrinogênio. Ocorre a formação de um complexo enzimático envolvendo o FT e o FVIIa (localizado fora do vaso sanguíneo) que ativa FX, e este, por sua vez, liga-se ao fator V (um cofator) e é ancorado na presença de íons de cálcio à superfície da membrana, gerando o complexo protrombinase (Schenone et al., 2004). Este complexo

converte protrombina em trombina que, por sua vez, leva à formação de um coágulo pela conversão de fibrinogênio em fibrina (Butenas e Mann, 2002).

1.2.3. Via comum

Quando o FX foi ativado, seja pela via intrínseca, extrínseca, ou ambas, a protrombina é ativada e passa a ser chamada de trombina. Essa reação ocorre na presença de Ca^{+2} , FV e fosfolípídeos. A trombina pode catalisar a formação de mais trombina a partir de protrombina, bem como a formação dos co-fatores FVa e FVIIIa, resultando em uma amplificação da coagulação. A trombina leva à formação de fibrina a partir de fibrinogênio, que também ativa o fator XIII, o qual estabiliza o coágulo. O processo termina quando ocorre a aglutinação de plaquetas e células como os eritrócitos, formando um coágulo que bloqueia o extravasamento de sangue (Broze, 1995; Schenone et al., 2004).

Quanto às proteínas plasmáticas relacionadas com a coagulação, estas podem ser classificadas em grupos. Há um primeiro conjunto onde estariam as proteínas que promovem a formação da rede de fibrina: cofatores como o FVII e o FV, proteínas estruturais (fibrinogênio) e pró-enzimas (FXI, protrombina). Outro grupo apresenta proteínas reguladoras, que modulam e localizam a formação do coágulo à região lesada: são os inibidores de protease (ex: antitrombina III) e enzimas que inativam componentes pró-coagulantes (ex: proteína C ativada). Há também um último conjunto envolvendo proteínas fibrinolíticas, que dissolvem a rede de fibrina; são enzimas e inibidores de proteases (Salzano, 1990).

Cabe ressaltar que a via intrínseca e extrínseca se entrecruzam fazendo a complementação uma da outra (Bolton-Maggs e Pasi, 2003) (**Figura 1**). Estudos recentes indicam que o modelo da cascata da coagulação sanguínea, pode ser considerado como ultrapassado se observado sob a luz da fisiologia da coagulação, pois apresenta barreiras para sua aplicação *in vivo* (Becker, 2005). Porém, se considerarmos testes laboratoriais de coagulação realizados *in vitro*, o sistema baseado nas duas vias possui uma grande utilidade. O diagnóstico de alterações hemostáticas utiliza essas duas vias da cascata da coagulação, para inferir o TP (tempo de protrombina) e o TTPA (tempo de tromboplastina parcial ativado) por meio de testes, que avaliam respectivamente a via extrínseca e intrínseca (Davie e Ratnoff, 1964).

1.3.O Modelo Celular da Coagulação Sanguínea

Muitos trabalhos têm demonstrado mecanismos pelos quais as células influenciam o processo de coagulação. É conhecido que os fatores protéicos da coagulação sanguínea direcionam e controlam o processo, com as células servindo primariamente como provedoras de fosfatidilserina, e formando a superfície na qual os complexos procoagulantes se ancoram (Hoffman, 2003).

Outros trabalhos propõem modelos nos quais a coagulação sanguínea é regulada pelas propriedades da superfície celular. Estes enfatizam a importância de receptores celulares específicos para as proteínas da coagulação, onde células com o mesmo conteúdo de fosfatidilserina podem executar diferentes funções no processo de hemostasia, de acordo com os receptores de superfície (Becker, 2005).

Em 2001, foi sugerido o Modelo Celular de Coagulação Sanguínea para melhor representar o que ocorre *in vivo* durante o reparo de uma lesão endotelial. Este novo modelo postula que o processo não ocorre como uma cascata, mas sim em estágios sobrepostos de iniciação, amplificação e propagação (Hoffman e Monroe, 2001) (**Figura 2**).

Essas três etapas do modelo ocorrem na superfície de diferentes células. A iniciação acontece na superfície de fibroblastos com o fator tissular (TF) (Wilcox et al., 1989). A partir de um estímulo em favor da coagulação, os fatores Xa, IXa e a trombina são formados, com a finalidade de iniciar o processo de coagulação. O estímulo inicial (ativação) é direcionado para as plaquetas e elas vão tornar-se ativadas pelo processo de amplificação. As plaquetas são aderidas, ativadas e passam a acumular co-fatores ativados em suas superfícies. A etapa de propagação ocorre pela superfície ativada dessas plaquetas, quando os co-fatores ativados anteriormente se combinam com proteases ativas na superfície plaquetária. A atividade desses complexos pró-coagulantes leva a uma alta na concentração de trombina, resultando na polimerização de fibrina.

Existem muitos recursos que objetivam evitar uma coagulação imprópria, entre eles: (a) a inativação e a propagação ocorrem em diferentes superfícies celulares; e (b) os

inibidores plasmáticos de proteases afetam as reações nas superfícies celulares inibindo proteases ativas difundidas no plasma.

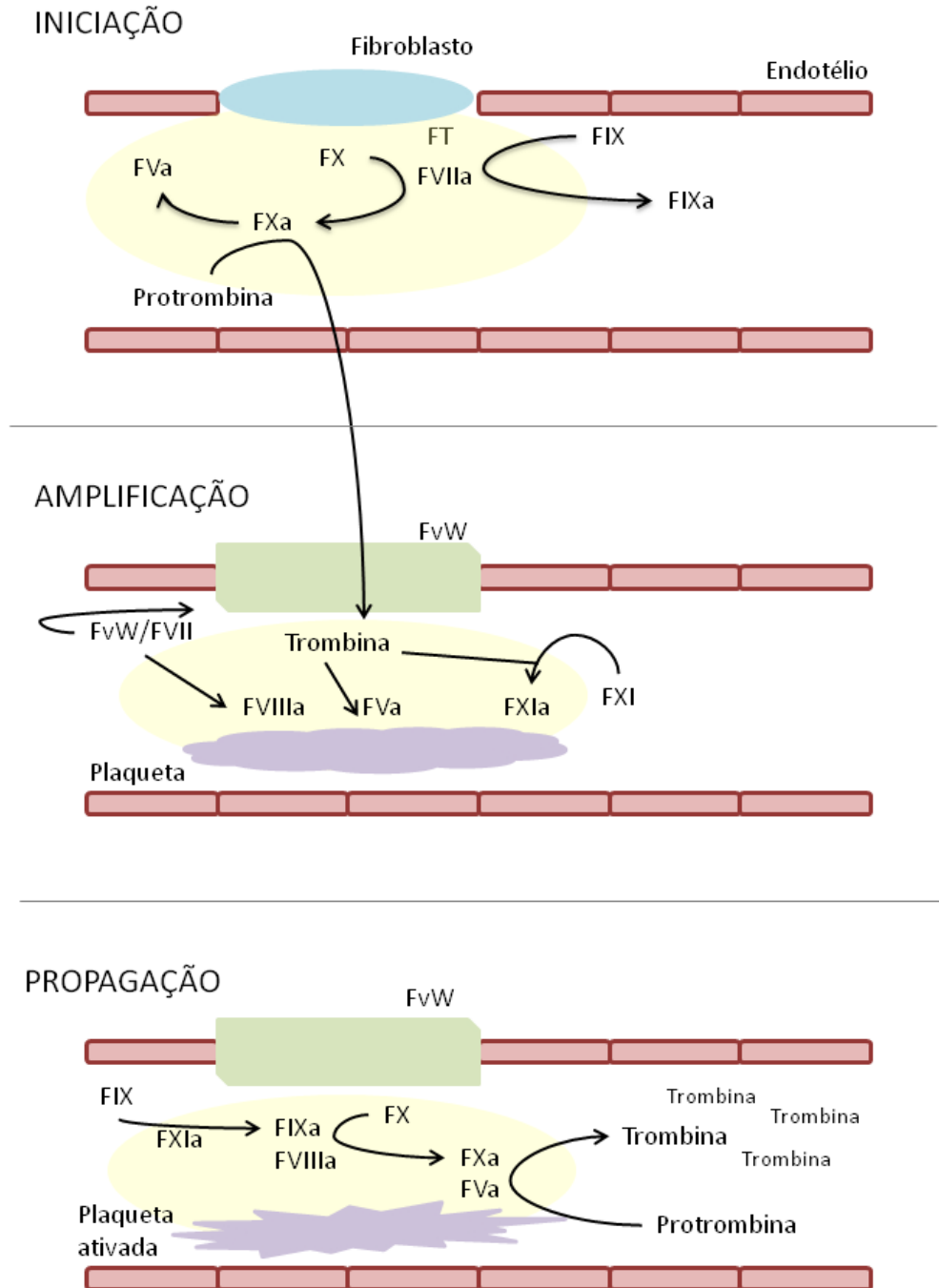


Figura 2. O modelo celular da coagulação sanguínea. Representação das etapas de iniciação, amplificação e propagação. **Figura original.**

1.3.1. Iniciação

A coagulação é iniciada pelo fator tissular (FT), que consiste em uma proteína integral de membrana, não tendo relação estrutural com outras proteínas da coagulação, mesmo estando presente na estrutura da membrana onde esses fatores são sintetizados. A iniciação e a propagação ocorrem em superfícies de celulares distintas, o que possibilita uma regulação mais fina.

Quando ocorrem rompimentos na parede, o plasma entra em contato com o FT que está ligado a células extra vasculares. O FVII acopla-se ao FT e é rapidamente ativado por proteases, que podem ou não estar associadas ao processo de coagulação (McMichael, 2012). A aproximação do complexo FVIIa/TF à superfície ativada das plaquetas é um passo chave no início da coagulação. Tal complexo leva à ativação de pequenas quantidades de FX e FIX. O FXa pode ativar o FV plasmático e algumas outras proteases celulares não envolvidas diretamente na coagulação sanguínea (Hoffman e Monroe, 2001).

1.3.2. Amplificação

A amplificação da coagulação acontece na superfície das plaquetas (Hoffman e Monroe, 2001; Roberts et al., 2004). As plaquetas são ativadas pela trombina e passam a liberar o Fator V. O complexo Fator VIII/ FvW acopla-se aos receptores que foram expostos nas plaquetas, sendo então clivado e levando à ativação do FVIII e liberação do FvW (Federici, 2003). Enquanto o FVIII se mantém ligado às plaquetas, o FvW promove a adesão plaquetária no subendotélio, simultaneamente ao processo de coagulação. A trombina também é responsável por ativar o FXI na superfície plaquetária (Huntington, 2005), que por sua vez ativa a maior quantidade de FIX, também nas plaquetas (Gailani, 2000). Deve ser também mencionado que uma pequena parcela de FIX já havia sido ativada durante a iniciação do processo.

1.3.3. Propagação e finalização

Com todos os elementos ativados, ocorre a ligação dos co-fatores VIIIa e Va nas plaquetas, e em seguida, a ligação das suas enzimas correspondentes, Fator IX ativado (FIXa) e FXa. O complexo tenase consiste em um sistema de ativação do FX, onde o FVIIIa está acoplado com o FIXa. O FVIII age como um catalisador que aumenta a ativação do FX pelo FIXa.

O FXa, juntamente com seu co-fator (FVa) forma o complexo protrombinase na superfície fosfolipídica das plaquetas ativadas. Esse complexo gera trombina a partir de protrombina (Versteeg et al., 2013). A trombina cliva os dímeros de fibrinogênio nas extremidades terminais das cadeias α e β , liberando os fibrinopeptídeos A e B que formam um monômero de fibrina capaz de auto-associação, levando à formação de um coágulo de fibrina. Ocorre a estabilização do coágulo por ativação do fator XIII, na presença de íons de cálcio (Hoffbrand, Pettit e Moss, 2001); (Butenas e Mann, 2002).

1.4. O fator IX da coagulação

O Fator IX é sintetizado nos hepatócitos como uma proteína precursora, com 461 aminoácidos e 57 kDa, que sofre processamentos pós-traducionais, como a remoção da sequência sinal e do pró-peptídeo, que são clivados em reações separadas antes da secreção para a circulação. Falhas na remoção dessas seqüências podem resultar em proteínas não funcionais (Green et al., 1989).

Na circulação, o fator IX está presente com uma cadeia de 415 aminoácidos e massa molecular de 55 kDa. A concentração do FIX em um plasma não deficiente é de aproximadamente 5 $\mu\text{g/mL}$. Na circulação, a proteína madura (zimogênio da serino-protease), tem sua ativação proteolítica via complexo FVIIa/FT ou Fator XI ativado (FXIa) de acordo com o modelo celular da coagulação sanguínea, resultando na sua forma ativa (FIXa) (Schmidt e Bajaj, 2003).

Uma representação esquemática da relação gene – proteína é apresentada na **Figura 3**. O primeiro éxon do F9 codifica uma seqüência sinal (pré-peptídeo) que leva à liberação do FIX do hepatócito (Bowen, 2002). O pré-peptídeo direciona a proteína do retículo

endoplasmático para o complexo de Golgi. Esse peptídeo sinal, quando traduzido, leva à ligação do ribossomo com a superfície do retículo endoplasmático, onde a tradução é continuada após a clivagem deste pré-peptídeo por uma peptidase sinal, uma enzima conhecida como PACE/furina (*pairedbasic amino acidcleavingenzyme*) que comanda a retomada da tradução e produção de uma proteína totalmente funcional (Bristol et al., 1996).

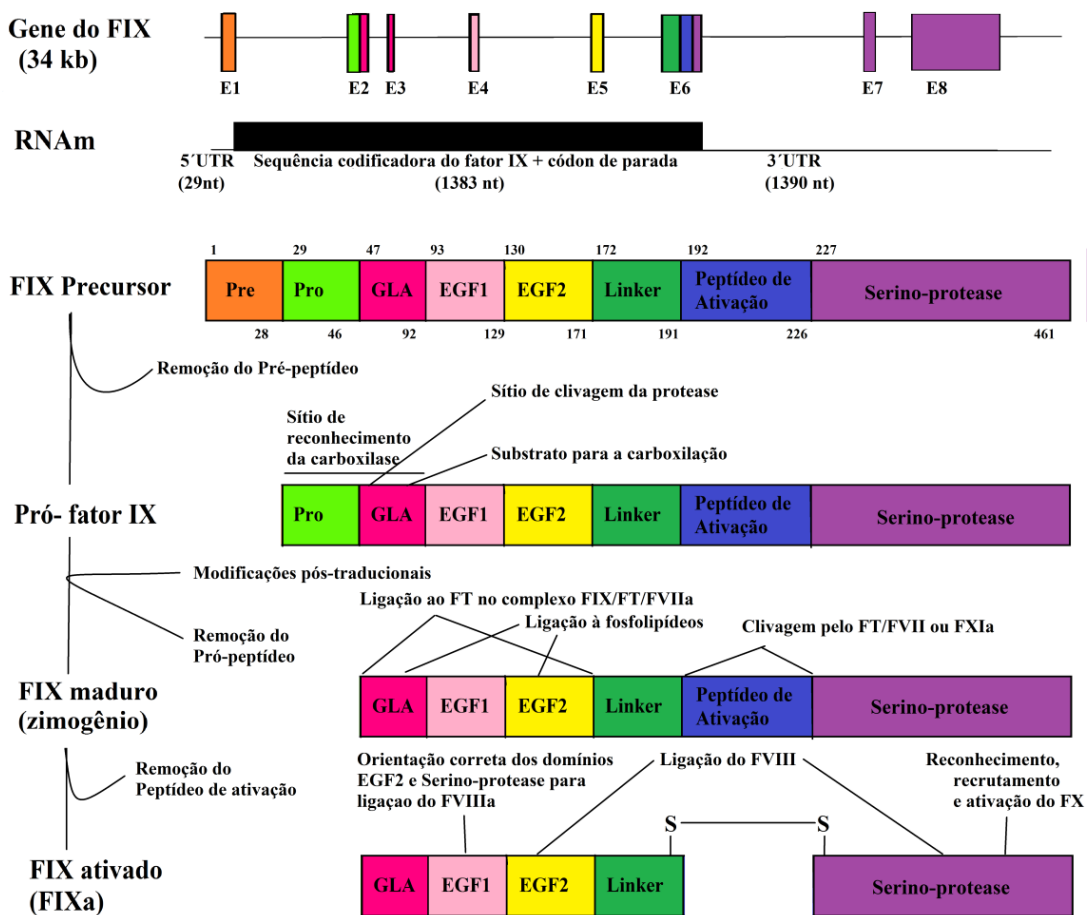


Figura 3. Fases do FIX: do gene à ativação. Representação das etapas de transcrição, tradução e modificações pós-traducionais que o FIX sofre até tornar-se ativado. **Figura original.**

O éxon 2 codifica toda a seqüência do pró-peptídeo e a maior porção do domínio GLA (*Vitamin K-dependent carboxylation/gamma-carboxylglutamic*). O pró-peptídeo juntamente com o domínio GLA, forma o sítio de reconhecimento da γ carboxilase, uma enzima dependente de vitamina K, responsável pela modificação pós-traducional dos 12 primeiros resíduos de ácido glutâmico (*Glu*) do domínio GLA no peptídeo maduro. Essa carboxilação converte o ácido glutâmico em ácido γ carboxilglutâmico, que é necessário para a interação do fator IX com fosfolípidos e células endoteliais (Bentley et al., 1986; Chavali et al., 2011).

Após a γ carboxilação dos resíduos, que acontece antes da secreção do FIX, ocorre a separação do pró-peptídeo com relação ao FIX maduro, por meio de clivagem realizada por uma peptidase específica.

O domínio GLA liga-se à superfície das plaquetas por meio do Ca^{2+} , iniciando a ativação (Green et al., 1989). Ele compreende uma sequência da tirosina (*Tyr*) 47 ao ácido glutâmico (*Glu*) 92 e sua principal função é ancorar e concentrar o fator da coagulação na membrana ativada via pontes de Ca^{2+} (Zögg e Brandstetter, 2009). O éxon 3 codifica a parte final do domínio GLA (Harlos et al., 1987).

O quarto e o quinto éxons codificam os domínios EGF (*epidermal growth factor* – fator de crescimento epidérmico). Existem dois domínios EGF que juntos abrangem do aminoácido 93, um ácido aspártico, ao resíduo na posição 191, uma arginina. A região amino-terminal dos domínios EGF liga-se ao Ca^{2+} , criando uma interface que permite interações de ligação (Handford et al., 1990). A capacidade de incorporação do FIX ao complexo tenase (FVIIIa/ FIXa) e sua ligação ao FVIIIa e FX são determinadas nestes domínios EGF (Chavali et al., 2011; Venkateswarlu, 2014).

O domínio EGF do tipo B (EGF1) contém inúmeros resíduos de aminoácidos, incluindo *Asp93*, *Asp 95*, *Asp 110* e *Gln96*, que estão envolvidos na ligação de alta afinidade com o cálcio. O resíduo *Asp110* sofre uma modificação pós-traducional por β -hidroxilação do ácido aspártico (Stenflo et al., 1987; Handford et al., 1990) enquanto que os aminoácidos *Ser 99* e *Ser 107* sofrem glicosilação. Já o domínio EGF do tipo A (o segundo deles), consiste em uma região que possui afinidade para interagir com o FVIIIa (fator VIII ativado), que funciona como co-fator do FIX. Além disso, este domínio contribui com a ligação do FIX às plaquetas (Chavali et al., 2011).

O peptídeo de ativação e uma parte N-terminal do domínio serino-protease são codificados pelo sexto éxon. Essa região leva ao reconhecimento e recrutamento do complexo FVIIa/FT ou FXI que determinam a clivagem do peptídeo de ativação. O peptídeo clivado possui aproximadamente 35 resíduos de aminoácidos, da alanina 192 até a arginina 226, que são removidos na ativação do zimogênio daserino-protease (Green et al., 1989).

Este domínio é rico em glutamato e aspartato e possui três modificações pós-traducionais descritas como principais: a sulfatação do resíduo 201, uma tirosina, N-glicosilação nos resíduos 203, 213 e 218, todos de asparagina, e a fosforilação do resíduo 204 de serina (Makino et al., 2000). Com essas alterações, o zimogênio fica preservado de ativações e clivagens antecipadas e torna-se solúvel.

A ativação proteolítica ocorre pela clivagem em duas regiões (Arg¹⁹¹-Ala¹⁹² e Arg²²⁶-Val²²⁷), correspondentes às margens do peptídeo de ativação que abrange do resíduo na posição 192 até a posição 226. Com a remoção do peptídeo de ativação, o FIXa forma duas cadeias unidas por uma ponte dissulfeto (Schmidt e Bajaj, 2003): uma cadeia leve que abrange do resíduo 47 ao 191, e uma cadeia pesada que vai do resíduo 227 ao 461. A ponte dissulfeto que liga a cadeia leve à cadeia pesada situa-se entre o resíduo 178 do domínio EGF2 e 335 do domínio serino-protease.

Os éxons 7 e 8 codificam o domínio serino-protease, que abriga uma tríade: *His 267*, *Asp315*, *Ser411*, altamente conservada em todas as serino-proteases (Lillicrap, 1998; Mahajan et al., 2004). Outros resíduos neste domínio devem contribuir como sítios de ligação do FIX com o FX e com o seu co-fator, o FVIIIa (Green et al., 1989). Esse domínio está envolvido na clivagem proteolítica do FX em FIXa. Essas regiões que estão envolvidas como substrato ou parte enzimática da catálise do FIX são as que possuem altas taxas de mutações deletérias (Chavali et al., 2011).

Após as modificações pós-traducionais (Arruda et al., 2001) (**Figura 4**) e ativação, o FIX apresenta quatro domínios, bem como outras proteínas da cascata da coagulação que são dependentes de vitamina K, como o FX (Orlova et al., 2012). Na cadeia leve estão presentes o domínio Gla, EGF1 e EGF2 (fator de crescimento epidérmico 1 e 2), enquanto que na cadeia pesada está presente o domínio serino-protease (Schmidt e Bajaj, 2003).

A **Figura 5** mostra um modelo da estrutura tridimensional do FIX ativado, com o domínios GLA, EGF1, EGF2 e serino-protease.

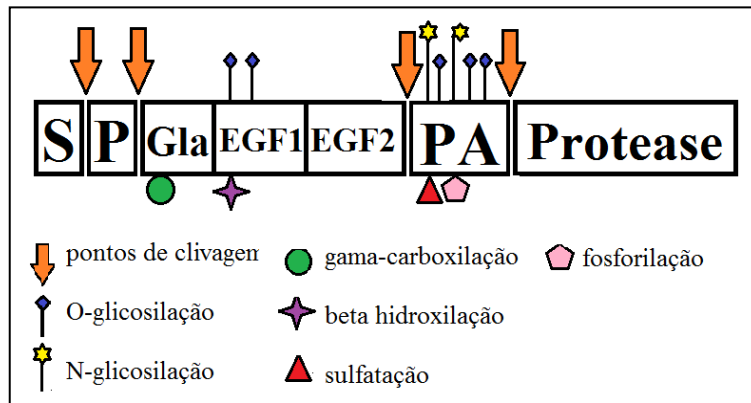


Figura 4. Modificações pós-traducionais do fator IX. Representação esquemática da estrutura proteica do FIX, com suas modificações pós-traducionais. Os domínios estão esquematicamente representados como S → Peptídeo sinal ou Pré-peptídeo; P → Pró-peptídeo; Gla → domínio Gla; EGF1 → fator de crescimento epidérmico; EGF2 → fator de crescimento epidérmico 2; PA → Peptídeo de ativação; Protease → domínio serino-protease. **Figura original.**

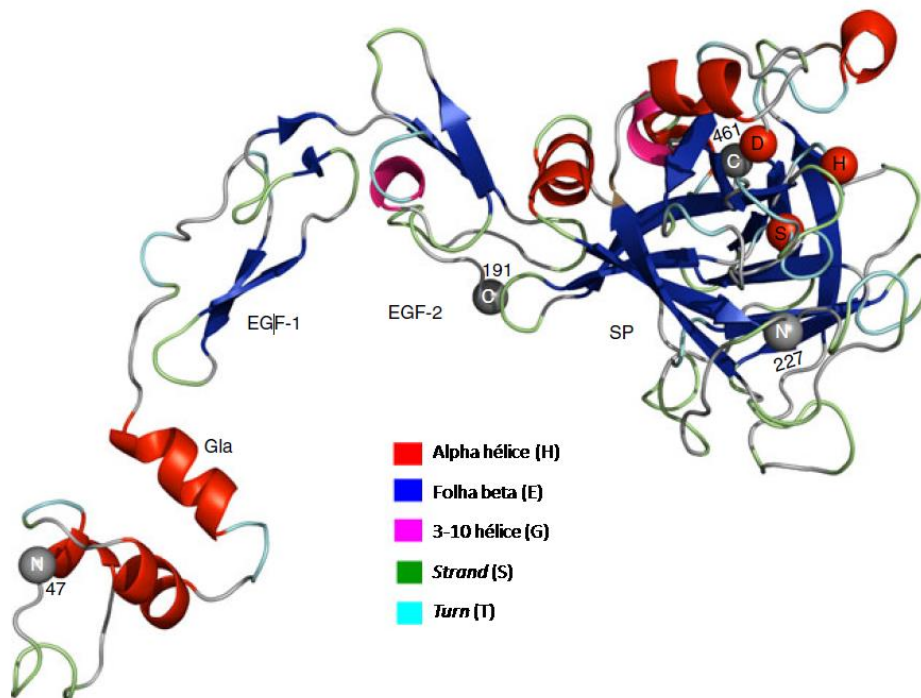


Figura 5. A estrutura do eixo principal do FIX ativado. Representação dos domínios GLA, EGF1, EGF2 e Serino-protease, presentes no FIX ativado, após a clivagem do peptídeo de ativação. A imagem mostra regiões de alfa hélice (vermelho), folhas beta (azul escuro), filamentos (verde), voltas (azul claro). Em cinza, está a tríade catalítica His267-Asp315-Ser411. **Adaptado de Rallapalli et al (2013).**

1.5.O gene F9

O gene do FIX está localizado no braço longo do cromossomo X, na banda Xq27.1 (Camerino et al., 1984). Ele foi clonado por diversos grupos entre 1982 e 1983. O *F9* possui 34kb e está organizado em 8 éxons que codificam os domínios proteicos do FIX (Chance et al., 1983; Anson et al., 1984).A região reguladora 5' do gene não tem a sequência consenso TATA Box e nem o elemento de iniciação. Embora não haja consenso sobre o assunto, o maior sítio de início da transcrição parece ocorrer 29 nucleotídeos a montante (*upstream*) do início de tradução (Lillicrap, 1998).

Se comparado ao *F8*, o gene do fator IX é consideravelmente menor e estruturalmente mais simples. O transcrito possui um comprimento de 2.802 bases(Yoshitake et al., 1985) e compreende uma pequena região 5' não traduzida (29 bases), uma fase de leitura somada ao códon de parada (1.383 bases), e uma região 3' não traduzida com 1.390 bases (Anson et al., 1984). Esta região inclui a sequência de clivagem AATAAA, 15 nucleotídeos à montante da cauda de poli-A.

O quadro de leitura aberta (*open reading frame*) codifica os domínios de Pré-peptídeo e de Pró-peptídeo. O Pré-peptídeo, ou sequência sinal, direciona o fator IX para a secreção, enquanto o Pró-peptídeo oferece um domínio de ligação para a carboxilase dependente de vitamina K. Esta enzima carboxila resíduos de ácido glutâmico no domínio GLA adjacente. O restante representa o zimogênio do FIX, o qual entra na circulação sanguínea após a remoção do Pré-peptídeo e do Pró-peptídeo (Anson et al., 1984). Os domínios do zimogênio são identificados de acordo com a sua estrutura e função.

1.6.Hemofilia B

As coagulopatias hereditárias são distúrbios que levam a alterações na hemostasia, devido a alguma deficiência qualitativa e/ou quantitativa de um ou mais fatores da coagulação sanguínea. Em sua maioria, essas doenças genéticas têm herança autossômica recessiva, poucas sendo autossômicas dominantes ou recessivas ligadas ao X. Entre essas coagulopatias, aquelas de maior prevalência são a hemofilia A, a doença de Von Willebrand e a hemofilia B(Peyvandi et al., 2006).

A hemofilia B, também conhecida como doença de Christmas, é uma coagulopatia de herança recessiva ligada ao cromossomo X (Brunetti-Pierrri et al., 2009; Rezende et al., 2009). Esta doença hemorrágica é caracterizada por anormalidades no gene (F9) que codifica o fator IX da coagulação sanguínea(Bertina e Veltkamp, 1978). Esta deficiência leva à dificuldade de produzir adequadamente um coágulo estável após dano tecidual (Rezende et al., 2009; Dunn, 2010). Sua frequência é de um a cada 30.000- 40.000 nascimentos no sexo masculino, com uma prevalência cerca de cinco vezes menor que a da hemofilia A, que representa cerca de 80% dos casos(Casaña et al., 2009). Em 70% dos casos a transmissão da hemofilia ocorre através de mães portadoras para indivíduos do sexo masculino. Nos 30% restantes, a doença é originada a partir de mutações *de novo*, esporádicas. Em função do tipo de transmissão, filhas de homens com hemofilia serão portadoras obrigatórias.

A existência de duas formas de hemofilia foi sugerida a partir de um estudo de ligação entre a cegueira para cores e a hemofilia. A partir desta abordagem, Rosemary Biggs desenvolveu estudos em Oxford e confirmou a existência de dois tipos de hemofilia, A e B. O paciente numero 1 do estudo tinha o sobrenome “Christmas” e com isso a condição de hemofilia B foi conhecida também por “doença de Christmas” (Lillicrap, 1998).

Os fatores determinantes para a manifestação da hemofilia A (FVIII) e da hemofilia B (FIX)participam da mesma reação na via intrínseca da hemostasia e as manifestações clínicas dessas doenças genéticas são muito semelhantes (Lillicrap, 1998). A deficiência ou disfunção de qualquer um desses fatores compromete a ativação do fator X e os passos seguintes da cascata da coagulação. Assim, a base bioquímica que leva às hemofilias é uma

insuficiência do complexo tenase. As duas condições são clinicamente similares por perturbarem o mesmo passo essencial no processo de geração de fibrina (Bowen, 2002).

A gravidade dos sangramentos hemorrágicos está relacionada à magnitude da deficiência do fator da coagulação. De acordo com a atividade residual do FIX, a hemofilia B é classificada em grave, moderada ou leve (White et al., 2001). Indivíduos com um quadro grave apresentam a atividade do fator IX menor ou igual a 1% do normal; enquanto que aqueles com deficiência moderada possuem atividade do fator IX entre 1-5%, e aqueles indivíduos com um quadro leve de hemofilia B tem a atividade entre 5-40% (Chavali et al., 2011).

O nível de atividade do fator influencia no quadro clínico dos pacientes. Quanto mais graves, maior a propensão a sangramentos espontâneos, o que geralmente não ocorre em pacientes leves, os quais costumam ter apenas sangramentos pós-traumáticos (Lillicrap, 1998). Foi verificado, porém, que 60-70% dos indivíduos portadores de hemofilia B desenvolvem formas graves ou moderadas (Nerich et al., 2008), embora essas proporções não estejam ainda bem estabelecidas.

O conhecimento das bases moleculares que levam à hemofilia B aumentou nas últimas décadas a partir da clonagem e caracterização do gene do fator IX em 1982. Com isso, novas pesquisas surgiram e estão contribuindo para descrever a condição. Segundo a literatura, a gravidade do fenótipo tem relação com o tipo e a posição das mutações. As mutações que levam à hemofilia B incluem mutações de ponto (sem sentido e de sentido trocado), deleções, inserções e defeitos de processamento (Nakamura et al., 2015).

1.6.1. Mutações no F9 e a hemofilia B

A hemofilia B é causada por variantes (mutações) no gene F9. Existe uma grande heterogeneidade clínica nesta desordem devido ao fato de existirem cerca de 1095 mutações diferentes que levam a esta condição. Diferentes mutações estão presentes em todos os éxons do fator IX. Conseqüentemente existem alterações de aminoácidos que estão registradas como causas da hemofilia B em todos os domínios do Fator IX.

Existem variantes presentes em 336 (73%) dos 461 resíduos da proteína (**Figura 6**). De acordo com o banco de dados do FIX, existem 1113 mutações que resultam na hemofilia B, nas formas grave, moderada e leve. Entre essas 1113 variantes, a maioria corresponde a mutações de ponto (812/1113= 73,0%) e deleções (182/1113= 16,3%). E a menor porcentagem se apresenta em duplicações (0,4%) (**Tabela 1**) mostra os tipos de mutações e suas freqüências na hemofilia B. Quanto às mutações de ponto, 112 (13,8%) ocorrem em íntrons ou regiões *UTR* (*untranslatedregions*- regiões não traduzidas); 86 (10,6%) são mutações sem sentido (*nonsense*); 10 correspondem a mutações silenciosas (1,2%); e a maioria (604/812= 74,4%) são mutações de sentido trocado (*missense*).

MQRVNMIMAE	SPGLITICLL	GYLLSAECTV	FLDHENANKI	LNRPKRYNSG	50
KLEEFVQGNL	ERECMEEKCS	FEEAREVFEN	TERTTEFWKQ	YVDGDQCESN	100
PCLNGGSCKD	DINSYECWCP	FGFEGKNCEL	DVTCNIKNGR	CEQFCKNSAD	150
NKVVCSCTEG	YRLAENQKSC	EPAVPPFCGR	VSVSQTSKLT	RAETVFPDVD	200
YVNSTEAEETI	LDNITQSTQS	FNDFTRVVG	EDAKPGQFPW	QVVLNGKVDA	250
FCGGSIVNEK	WIVTAAHCVE	TGVKITVVG	EHNIEETEHT	EQKRNVIRII	300
PHHNYNAAIN	KYNHDIALLE	LDEPLVLNSY	VTPICIADKE	YTNIFLKFGS	350
GYVSGWGRVF	HKGRSALVLQ	YLRVPLVDRA	TCLRSTKFTI	YNNMFCAGFH	400
EGGRDSCQGD	SGGPHVTEVE	GTSFLTGIIS	WGEECAMK GK	YGIYTKVSRY	450
VNWIKEKTKL	T				

IDENTIFICAÇÃO DOS DOMÍNIOS: GLA -- EGF1 -- EGF2 -- Serine Protease

Figura 6: Mapa das mutações de ponto nos resíduos do FIX. Esquema ilustrando os 461 resíduos do FIX: em vermelho, resíduos onde estão descritas mutações de ponto. As tarjas representam domínios presentes no FIX ativado: em rosa, o domínio GLA; em laranja, EGF1; em azul, EGF2; e em amarelo; Serino-protease. **Adaptado de Rallapali et al. (2013)**

Tabela1.Frequência dos tipos de mutação no F9.

Tipo de mutação	Número de ocorrências	Frequência com relação ao total de 1113 (%)
Complexas	5	0,50%
Pontuais	812	73%
Deleções	182	16,30%
Duplicações	4	0,40%
Indels	17	1,50%
Inserções	39	3,50%
Polimorfismos	54	4,80%

Fonte: (Rallapalli et al., 2013)

De maneira geral, as mutações parecem ter uma distribuição uniforme entre os domínios EGF1, EGF2 e Serino-protease, porém são menos frequentes no domínio GLA. Algumas mutações estão presentes nos íntrons (137=12%), e cerca de 2% das mutações afeta tanto éxons quanto íntrons. Mutações que alteram a fase de leitura (*frameshift*) estão presentes em 198 casos; destas 141 ocorrem como deleções; 38 como inserções; 11 como indels; 3 como duplicações e 5 são complexas. Mutações silenciosas (sinônimas) podem ser neutras tanto do ponto de vista clínico, quanto em termos de funcionalidade da proteína; 10 diferentes mutações silenciosas foram encontradas em pacientes com fenótipos de leve a grave. A **Tabela 2** mostra a distribuição de 1093 mutações no F9. As 20 mutações que ocorrem simultaneamente em éxons e íntrons (correspondem a 1,8% do total de 1113 mutações) não estão representadas na tabela.

Dos 24 resíduos de cisteína presentes na molécula do fator IX, 22 estão envolvidos na formação de pontes dissulfeto. Existem 11 pontes cisteína-cisteína colaborando com a estabilidade da estrutura do FIX. A primeira ponte ocorre no domínio GLA, entre os resíduos 64 e 69; outras três ocorrem no domínio EGF1, entre os resíduos 97-108, 102-117, e 119-128; três em EGF2 (134-145, 141-155, e 157-170); e no domínio Serino-protease também ocorrem 3 pontes (252-268, 382-396, e 407-435); ainda existe uma ponte entre a cadeia leve e a cadeia pesada, que envolve os resíduos 178 e 335.

Tabela 2. Distribuição das mutações descritas no banco de dados do FIX.

Região	Mutações		Total
	N	%	N
5'UTR	28	2,5	33
3'UTR	5	0,5	
Éxon 1	38	3,4	923
Éxon 2	119	10,7	
Éxon 3	15	1,3	
Éxon 4	88	8	
Éxon 5	96	8,6	
Éxon 6	88	7,9	
Éxon 7	67	6	
Éxon 8	412	37	
Íntron 1	27	2	137
Íntron 2	25	2,2	
Íntron 3	14	1,3	
Íntron 4	29	2,6	
Íntron 5	13	1,2	
Íntron 6	20	1,8	
Íntron 7	9	0,8	
Total	1093	98,2	1093













Fonte: (Rallapalli et al., 2013).

A **Tabela 3** mostra a distribuição das mutações descritas no banco de dados da Hemofilia B, considerando centros e laboratórios entre os países. O país com o maior número de mutações no banco de dados foi o EUA (658), seguido da Alemanha (435), Reino Unido (387), França (218), China (174), Itália (173) e Índia (155). Os demais países têm o número de mutações registradas no banco de dados inferior a 100, sendo que 11 países têm esse número menor que 20, incluindo o Brasil que reportou com apenas 12

variantes. As informações sobre mutações no banco vão de encontro com a realidade no caso do Brasil. Ele é o terceiro país em números de casos de indivíduos com hemofilia B (**Tabela 4**), parecendo estar mal representado em termos de análise de mutações (n=12).

Tabela 3. Países com casos registrados na *Haemophilia B Mutation Database*.

Países	Número de mutações descritas	Porcentagem (%)
Sem registro de país	996	26,8
Argentina	55	1,4
Austrália	2	<1,0
Brasil	12	0,3
Canadá	42	1,1
China	174	4,6
Dinamarca	32	0,8
França	218	5,8
Alemanha	435	11,7
Índia	155	4,1
Irlanda	16	0,4
Itália	173	4,6
Japão	34	0,9
Jordânia	10	0,2
Malásia	15	0,4
México	9	0,2
Holanda	15	0,4

Nova Zelândia		32	0,8
Paquistão		15	0,4
Portugal		13	0,3
Coréia do Sul		33	0,8
Espanha		67	1,8
Suécia		37	1,0
Suíça		1	<1,0
Taiwan		20	0,5
Tunísia		17	0,4
Turquia		40	1,0
Reino Unido		387	10,4
Estados Unidos		658	17,7
Total		3713	

Fonte: (Rallapalli et al., 2013)

Tabela 4. Número de indivíduos relatados nos países com maior ocorrência de casos de hemofilia

Países	Hemofilia B	Hemofilia A
Estados Unidos	4421	14175
Índia	2127	14508
Brasil	1945	9912
China	1787	11837
Reino Unido	1450	6390
França	1267	5581
Irã	1071	4944
Japão	1064	4986
Egito	1062	4358
Rússia	992	5801

Fonte: (World Federation of Hemophilia, 2016)

A partir de uma pesquisa bibliográfica, foi constatado que até o presente momento apenas um estudo foi publicado sobre a hemofilia B no Brasil. Neste estudo (Figueiredo, 1993) foram descritas pela primeira vez duas grandes deleções e uma mutação sem sentido. As deleções envolvem a perda do gene inteiro e deleção do éxon 1 ao éxon3, e a mutação sem sentido corresponde a p.Arg298*. Outros pesquisadores brasileiros também estudaram mutações no gene do FIX, porém não publicaram os resultados. O estudo mais recente descrevendo uma população na América Latina é de 2013 e descreve pela primeira vez a população com hemofilia B na Argentina (Radic et al., 2013). Foram analisadas 55 famílias, onde foram detectadas 47 mutações, 10 delas encontradas pela primeira vez.

1.7.Relação estrutura-função: Bioinformática estrutural

A genética humana enfrenta um grande desafio: a análise de implicações funcionais de mutações. Considerando mutações de sentido trocado, muitas vezes elas são alterações altamente patogênicas, e outras vezes podem se mostrar inócuas, os chamados polimorfismos. A maioria das variantes que ocorrem no gene F9 são mutações de sentidotrocado. Para as doenças genéticas onde uma variante pode determinar a perda da funcionalidade normal ou mesmo interferir em uma via, como é o caso da hemofilia B, estudos para encontrar meios de entender os mecanismos que levam à condição, relacionando genótipo-fenótipo e estrutura-função, tornam-se essenciais. Análises *in silico* vêm apresentando contribuições importantes nos campos do impacto das mutações na estrutura das moléculas.

Compreendendo o efeito que uma mutação tem na proteína, é possível melhorar análises de sua função e de interações com outras estruturas, bem como estabelecer relações de causa e efeito entre a mutação, estrutura alterada e quadro clínico. Análises cristalográficas e o progresso que acontece no desenvolvimento de *softwares* que realizam modelagem molecular e predições de estrutura têm contribuído para elucidar os efeitos que as mutações exercem na proteína (Mertens et al., 1999). Metodologias de modelagem por homologia a partir de estruturas já cristalografadas, podem trazer novos entendimentos sobre mutações já descritas.

Além de métodos *in silico* que envolvem a modelagem estrutural, também existem ferramentas que funcionam para fazer a predição do dano que pode ser gerado por mutações. Muitas dessas ferramentas são alimentadas por bancos de dados, que são curados e atualizados, para uma melhor acurácia. O número de ferramentas que predizem impactos de mutações em estruturas tem aumentado bastante, e por utilizarem algoritmos diferentes levam a interpretações diferentes. Portanto, é importante que sejam utilizadas ferramentas variadas para realizar a predição dos impactos das mutações.

1.8. Objetivos

1.8.1. Objetivo geral

O presente estudo visa a detecção de mutações no DNA de pacientes com hemofilia B do Rio Grande do Sul, analisando o papel destas mutações na determinação da condição, verificando alterações na estrutura e no potencial eletrostático do Fator IX. As mutações geralmente estão associadas a um determinado fenótipo clínico, o que torna importante a sua caracterização para compreender mecanismos de patogênese através da identificação de regiões funcionalmente importantes e estruturas essenciais.

1.8.2. Objetivos específicos

- a) Relacionar os tipos de alterações encontradas no gene F9 com a prevalência de mutações já descritas.
- b) Realizar associações entre as mutações encontradas e o quadro clínico apresentado pelos pacientes.
- c) Relacionar estrutura e função das variantes por meio da Bioinformática Estrutural.
- d) Melhorar a qualidade e a quantidade de informações sobre a hemofilia B no Rio Grande do Sul, visando o atendimento mais adequado aos pacientes e contribuindo no aconselhamento genético destes e suas famílias.

1.9.Detalhes do Projeto

Este estudo faz parte de um amplo projeto com mais de meio século, que visa à investigação de coagulopatias hereditárias na população do Rio Grande do Sul; seus artigos mais recentes são: Gorziza et al.(2013) e Rosset et al.(2014).

Abordando especificamente o tema da hemofilia B, houve trabalhos prévios (Pantoja, 2014) que não foram publicados como artigos científicos, mas colaboraram com a atual pesquisa.

CAPÍTULO 2 – MOLECULAR CHARACTERIZATION OF HAEMOPHILIA B PATIENTS IN SOUTHERN BRAZIL

Artigo submetido para publicação à revista *Haemophilia* e já revisto em uma primeira rodada de avaliação.

Molecular characterization of haemophilia B patients in southern Brazil

M. R. MEIRELES, A. G. PANTOJA, A. P. ORNAGHI, G. F. VIEIRA, F. M. SALZANO
and E. BANDINELLI

*Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande
do Sul, Porto Alegre, RS, Brazil*

Running title: HAEMOPHILIA B IN BRAZIL

Keywords: haemophilia B, factor IX, molecular studies, phenotype-genotype correlations,
missense mutations, protein structure

Numberofword: 2,794.

Correspondence: Francisco M. Salzano, Departamento de Genética, Instituto de
Biociências, UFRGS, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil.
Tel.: +55 51 3308-6747; fax: 55 51 3308-9823
E-mail: francisco.salzano@ufrgs.br

Summary

Introduction: The Haemophilia B (HB) gene (*F9*) is highly variable, with 1,095 mutations registered in its database. Therefore, it is unfortunate that with only one exception, no recent study of its molecular variability has been performed in Latin America.

Subjects and methods: A total of 43 HB patients with severe (70%), moderate (16%) and mild (14%) clinical pictures living in southern Brazil were studied by molecular methods, and the results analyzed with bioinformatic tools.

Results: Thirty-one different mutations, 25 already described but six not reported in the HB database were observed. The majority (74%) were of the missense, but nonsense (13%), frameshift and splicing (both 6.5% each) types were also found. The missense changes were investigated in relation to their phylogenetic conservation, and their predicted effects carefully considered in structural-functional analyses.

Conclusions: While the different mutation types presented a distribution not much different from those observed in previous surveys, the information given is of fundamental theoretical and clinical importance for the characterization of this under investigated condition in Latin America.

Haemophilia B (HB) (OMIM 306900) is a genetic disorder caused by mutations at the Factor IX (FIX) gene (*F9*) of the coagulation cascade. *F9* is a highly variable gene; not less than 1,095 different HB mutations are reported in the Haemophilia B Mutation Database (www.factorix.org/). Therefore, it is of high practical and theoretical value to verify whether the reported mutations in a given population are the same as those occurring elsewhere, or if there are new variants, which can result in different clinical pictures. It is then unfortunate that, with the exception of an Argentinian sample [1], no recent study on HB molecular *F9* variability had been performed in Latin America. We, therefore, decided to investigate: (a) the types of mutations occurring in a sample of HB southern Brazilians; (b) verify their distribution on the FIX domains; (c) relate the observed missense mutations with those reported elsewhere; (d) test the conservation of these regions in nine other mammalian species; and (e) make phenotype/genotype correlations taking into consideration the molecular localization of the variants. The present work is part of a longterm project that is being developed for half-a-century; relatively recent contributions are [2, 3].

A total of 43 independent HB patients were included in this study. They were ascertained in HEMOCENTRO-RS, a central hemotherapy institution of the State of Rio Grande do Sul, Brazil, and have ages varying from 3 to 64 years (mean: 28.7). Based on morphological traits they were classified as European-derived (91%) and African-derived (9%), proportion similar to that of the general population; while their clinical conditions were found to be severe (30), moderate (8), and mild (6). All of them signed an Informed Consent Form and the project was approved through the Brazilian Ethics Platform (PlataformaBrasil, CAAE 45164015.3.0000.5347).

HB diagnosis was performed through the FIX coagulation activity (FIX:C) using the one-stage clotting assay [4]. Genomic DNA was extracted from peripheral blood cells according to a nonenzymatic method [5]. The entire *F9* gene coding region, as well as their promoter and exon/intron boundaries, were amplified by the polymerase chain reaction (PCR) using eight primer pairs (Supplementary Information, Table S1). The amplicons were purified using the Exo-Sap kit (Amersham Biosciences, Les Ulis, France) and sent to Macrogen Inc. (Seoul, South Korea) where Sanger sequencing was performed. Nucleotide variations were re-tested in two independent samples to confirm the results.

Patients' sequences were compared with the normal *F9* gene sequence (GeneBank

accession number K02402) using the Codon Code Aligner tool implemented with the MEGA 5 software. The variations detected were compared with those reported in the Haemophilia B Mutation Database (<http://www.factorix.org/>) and when not found there, in the 1000 Genomes browser and USCS Genome browser to verify if they were polymorphisms. The Human Mutation Database and the Center for Disease Control and Prevention (CDC) catalogues were also checked in this regard.

Missense mutations were examined using the BLASTP and the UNIPROTKB/Swiss-Prot Mammalian database, comparing the human FIX residues (Swiss Prot ID: P00740) with orthologous FIX sequences deposited in the same databank.

Five algorithms were employed to analyze the impact of novel and the already described missense mutations found on protein structure: PolyPhen-2 (Polymorphism Phenotyping-2, <http://genetics.bwh.harvard.edu/pph2/>) [6], SIFT (Sorting Intolerant to Tolerant, <http://sift.jcvi.org/>) [7], HOPE (<http://www.cmbi.ru.nl/hope/>) [8], PROVEAN (Protein Variation Effect Analyzer, <http://provean.jcvi.org/>) [9], and Mutation Taster (<http://www.mutationtaster.org>).

Protein models for each new missense mutation were generated on Phyre2, a web portal that allows protein analysis, prediction, and modeling, using an expert mode, and one-to-one threading model. The PDB (Protein Data Bank [10, 11]) structure 1NL0 was used for homology modeling of mutations in the EGF1 and EGF2 domains, and PDB 1 pfx (a porcine model) for the Serine Protease domains. The Chimera Interface to Modeler was employed for visualization of the generated structures.

Table 1 lists the 31 mutations found in our study. Six of them not listed in www.factorix.org were observed by us in seven patients, (five with the severe and one each with the moderate and mild clinical types). The majority (4/6:67%) are of the missense type, the remaining two being respectively a frameshift and a splicing mutations. Two of the five occurring in the coding regions (40%) were present in the serine protease domain. In cases 3 and 4 the mutations occurred together with the p.Th194Ala polymorphism.

It is well-known that sites of functional importance show strong phylogenetic conservation. Therefore, as an additional test for the significance of the missense HB mutations we found, we compared the *Homo sapiens* mutation sites with the corresponding amino acids in eight mammalian species. The results are shown in Table S2. As indicated

there, there is strict conservation in 18 of the 23 sites considered, and in the remaining five sites the exceptions are restricted to one amino acid only in four species, and to two amino acids in one of the eight species examined.

We also made an intensive investigation about the predictive effects of the missense mutations detected, examining the results obtained in five algorithms (Table S3). They generally agreed indicating the damaging effects of the mutations considered, but some exceptions were found. They can be explained by the different methods employed by the algorithms to establish their scores, and whether they are designed for the analysis of missense mutations only, or are more general (like PROVEAN and Mutation Taster). The results obtained with HOPE, however, indicated that all mutations would lead to structural changes, such as spacing between units, incorrect folding, or binding changes, as well as charge and hydrophobicity modifications. These changes would be especially important when they are near sites of interaction with other coagulation substances [Factors VIIa, VIII, X, XIa, Tissue Factor (TF)].

Figure 1 presents in a graphical way the structural changes that would result from the four missense mutations found by ourselves that were not reported in www.factorix.org. The p.Cys157Arg, substitution, found in a severely affected patient, leads to a larger residue than that of the normal protein, changing charges and hydrophobicity, the latter altering the interactions with the molecule's core and surface. The loss of a cysteine, on the other hand, could also result in changes in one of the EGF2 disulfide bonds. Mutation p.Thr386Lys, also detected in one patient with a severe clinical picture, conditions a larger residue and hydrophobicity changes with reflections in the phenylalanine and glutamic acid residues at positions 388 and 433, respectively. The mutant residue is also positively charged, while the normal one is neutral. The position of this site, at the core of the protein, probably also affects the molecule's folding. The p.Val153Ala change, found in a moderately affected patient, leads to an opposite effect, of size reduction that can also affect function. Finally, p.Asp95Tyr, found in a patient with a mild clinical pattern, results in a larger-sized residue that is more hydrophobic. Apparently, one salt bridge was lost due to the mutation. EGF1, where this mutation is found, is a substract for binding to TF and to the FIX/TF/FVIIIa complex; these relationships could be altered by the change.

Four nonsense mutations were found by us, all of them previously listed in

www.factorix.org. The variant p.Arg75* occurred in three independent patients with the severe phenotype, one of them with inhibitor development.. The premature end of the translation process in the GLA domain affects all others. The three other mutations (p.Arg298*, p.Tyr371*, and p.Arg384*), present in the serine protease domain, were found, the first two in severely affected patients, and p.Arg384* in two patients with respectively moderate and mild phenotypes. The mutation taster algorithm furnished a probability of 100% for them as being causes of HB.

The two frameshift variations (p.Arg358Serfs*, and p.Trp453fs*) occur in the serine protease domain and were detected in two independent patients with the severe phenotype. The mutation Trp453* was not previously listed in www.factorix.org and PROVEAN classified them as a deleterious mutation.

Two splicing site mutations (c.88+5G>A, and c.88+51G>A) were found, the first reported and the second not reported in the above-indicated site. Variation c.88+5G>A was observed in one moderately affected patients, and c.88+51G>A in another with a severe phenotype. According to the Mutation Taster algorithm, both should cause disease with probabilities respectively of 100% and 99.9%.

In terms of epidemiology, only three of the mutations reported here (p.Arg43Gln, p.Thr342Met, and p. Arg298*) have been previously reported in Latin America [1]. Therefore, our results add an important reference to the genetic characterization of the populations of this region.

Our results will be useful to HB patients and their families through: (a) a refinement of the diagnostic process; and (b) appropriate genetic counseling of *F9* female carriers. We thank the patients and staff of HEMOCENTRO-RS for agreeing to participate in the study and for logistical support. Financing was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico.

Disclosures

The authors stated that they had no interests which might be perceived as posing a conflict or bias.

References

- 1 Radic CP, Rossetti LC, Abelleiro MM *et al.* Assessment of the F9 genotype-specificFIX inhibitor risks and characterization of 10 novel severe F9 defects in the first molecular series of Argentinian patients with haemophilia B. *ThrombHaemost*2013; 109: 24-32.
- 2 Gorziza RP, Vieira IA, Kappel DB *et al.* Genetic changes in severe haemophilia A: new contribution to the aetiology of a complex disease. *BloodCoagulFibrinol*2013; 24: 164-9.
- 3 Rosset C, Vieira IA, Sinigaglia M *et al.* Detection of new mutations and molecular pathology of mild and moderate haemophilia A patients from southern Brazil. *Haemophilia*2013; 19: 773-81.
- 4 Austen DEG, Rhynes IL. A Laboratory Manual of Blood Coagulation. Oxford: Blackwell Scientific, 1975: 109.
- 5 Lahiri DK, Nurnberger JI. A rapid non-enzymatic method for the preparation of HMWDNA from blood for RFLP studies. *Nucleic Acids Res* 1991; 19: 5444.
- 6 Adzhubei IA, Schmidt S, Peshkin L *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* 2010; 7: 248-9.
- 7 Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc*2009; 4: 1073-81.
- 8 Venselaar H, TeBeek TA, Kuipers RK *et al.* Protein structure analysis of mutations causing inheritable diseases. An e-Science approach with life scientist friendly interfaces. *BMC Bioinformatics* 2010; 11: 548.
- 9 Choi Y, Sims GE, Murphy S *et al.* Predicting the functional effect of amino acid substitutions and indels. *PLoS One* 2012; 7: e46688.
- 10 Berman HM, Kleywegt GJ, Nakamura H *et al.* The Protein Data Bank archive as an open data resource. *J Comput Aided Mol Des* 2014; 28: 1009-14.
- 11 Rose PW, Prlić A, Altunkaya A, *et al.* The RCSB protein data bank: integrative viewof protein, gene and 3D structural information. *Nucleic Acids Res* 2016; doi: 10.093/nat/gkw1000.

Figure Legend

Fig. 1. Localization of the novel factor IX (FIX) protein missense mutations found showing the structures of the normal (left) and mutated forms. Arg157 and Lys386, depicted in green, occurred in two severely affected patients; Tyr95, shown in blue, in one patient with a mild, and Ala153 (red) in one moderately affected phenotypes.

Table 1. Previously described mutations (<http://www.factorix.org/>) detected in this study.

Case	N° of patients*	Clinical severity	Nucleotide change [†]	Amino acid change [‡]	Amino acid change [§]	Mutation type	Exon/Intron	FX domain
1	1	moderate	c.88+5G>A			Splicing site	IVS1	
2	3	severe	c.128G>A	p.Arg43Gln	p.Arg(-4)Gln	Missense	Exon 2	Pro-Peptide
3	1	severe	c.128G>T	p.Arg43Leu	p.Arg(-4)Leu	Missense	Exon 2	Pro-Peptide
4	1	severe	c.127C>T	p.Arg43Trp	p.Arg(-4)Trp	Missense	Exon 2	Pro-Peptide
5	1	moderate	c.164T>G	p.Phe55Cys	p.Phe9Cys	Missense	Exon 2	GLA
6	1	severe	c.191G>A	p.Cys64Tyr	p.Cys18Arg	Missense	Exon 2	GLA
7	3	severe	c.223C>T	p.Arg75*	p.Arg29*	Nonsense	Exon 2	GLA
8	3	mild	c.224G>A	p.Arg75Gln	p.Arg29Gln	Missense	Exon 2	GLA
9	1	severe	c.400T>C	p.Cys134Arg	p.Cys88Arg	Missense	Exon 5	EGF-2
10	1	severe	c.427C>A	p.Gln143Lys	p.Gln97Lys	Missense	Exon 5	EGF-2
11	2	severe	c.509G>A	p.Cys170Tyr	p.Cys124Tyr	Missense	Exon 5	EGF-2
12	1	moderate	c.572G>A	p.Arg191His	p.Arg145His	Missense	Exon 6	Activation Peptide
13	2	severe	c.677G>A	p.Arg226Gln	p.Arg180Gln	Missense	Exon 6	Activation Peptide
14	1	severe	c.706G>T	p.Gly236Val	p.Gly190Val	Missense	Exon 6	Serine Protease
15	3	severe	c.892C>T	p.Arg298*	p.Arg252*	Nonsense	Exon 8	Serine Protease
16	1	mild	c.905A>G	p.His302Arg	p.His256Arg	Missense	Exon 8	Serine Protease
17	2	moderate	c.1025C>T	p.Thr342Met	p.Thr296Met	Missense	Exon 8	Serine Protease
18	1	severe	c.1072_1073delAG	p.Arg358Serifs*15	p.Arg312Serifs*	Frameshift	Exon 8	Serine Protease
19	1	mild	c.1097C>G	p.Ala366Gly	p.Ala320Gly	Missense	Exon 8	Serine Protease
20	1	severe	c.1113C>A	p.Tyr371*	p.Tyr325*	Nonsense	Exon 8	Serine Protease
21	1	severe	c.1120G>T	p.Val374Phe	p.Val328Phe	Missense	Exon 8	Serine Protease
22	1	severe	c.1150C>T	p.Arg384*	p.Arg338*	Nonsense	Exon 8	Serine Protease
23	1	moderate	c.1169T>C	p.Ile390Thr	p.Ile344Thr	Missense	Exon 8	Serine Protease
24	1	severe	c.1237G>A	p.Gly413Arg	p.Gly367Arg	Missense	Exon 8	Serine Protease
25	1	severe	c.1295G>T	p.Gly432Val	p.Gly386Val	Missense	Exon 8	Serine Protease

Table 1. (Cont.)

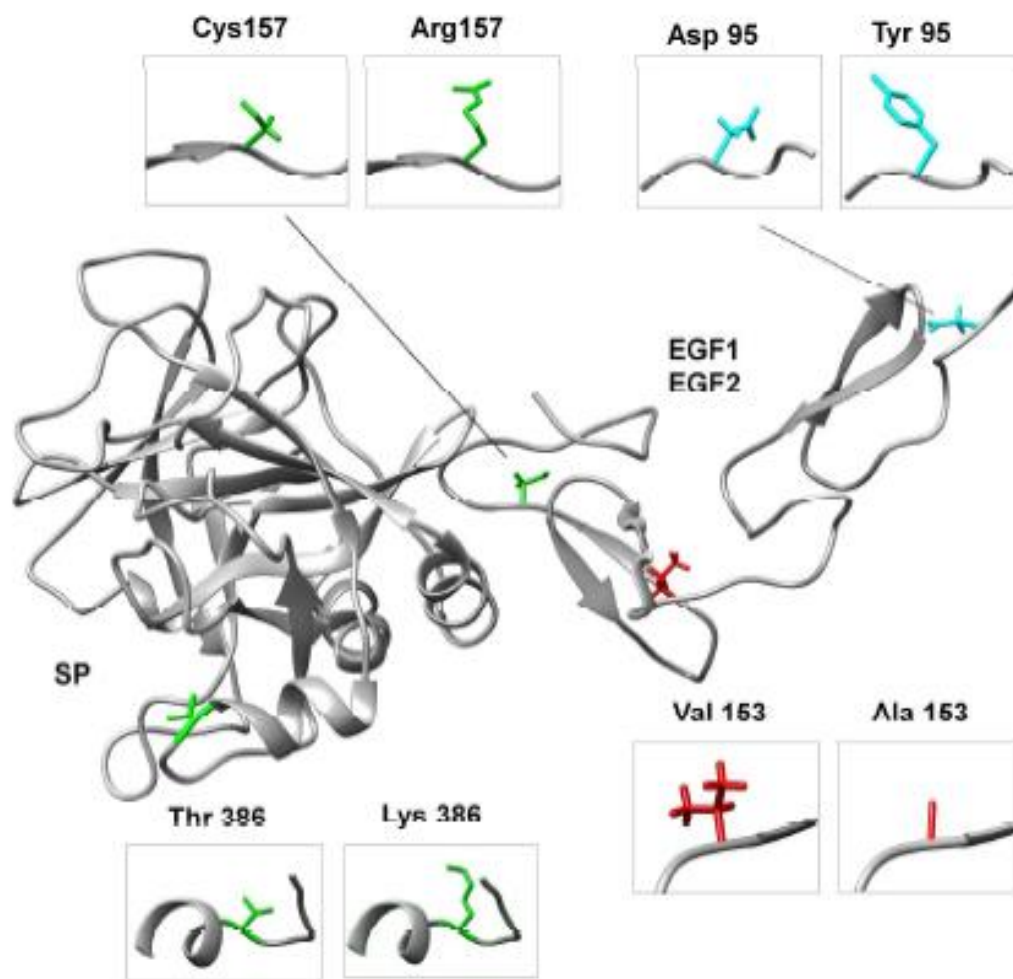
Case	N° of patients*	Clinical severity	Nucleotide change [†]	Amino acid change*	Amino acid change [‡]	Mutation type	Exon/Intron	FLX domain
New mutations								
1	1	severe	c.88+51G>A			Splicing site	IVS1	
2	1	mild	c.283G>T	p.Asp95Tyr	p.Asp49Tyr	Missense	Exon 4	EGF-1
3	1	moderate	c.458U>C	p.Val153Ala	p.Val107Ala	Missense	Exon 5	EGF-2
4	2	severe	c.469T>C	p.Cys157Arg	p.Cys111Arg	Missense	Exon 5	EGF-2
5	1	severe	c.1157C>A	p.Thr386Lys	p.Thr340Lys	Missense	Exon 8	Serine Protease
6	1	severe	c.1359_1359delG	p.Trp453fs*	p.Trp407fs*	Frameshift	Exon 8	Serine Protease

*Number of independent patients carrying the same mutation.

†cDNA mutation numbering is based on cDNA sequence, with nucleotide +1 corresponding to the A of the ATG translation initiation codon.

‡Sequence changes at protein level following the Nomenclature for Description of Genetic Variations approved by the Human Genome Variation Society.

§Sequence changes at protein level according to the Legacy nomenclature.



Localization of the novel factor IX (FIX) protein missense mutations found showing the structures of the normal (left) and mutated forms. Arg157 and Lys386, depicted in green, occurred in two severely affected patients; Tyr95, shown in blue, in one patient with a mild, and Ala153 (red) in one moderately affected phenotypes. More information about these substitutions are present in the text.

896x896mm (87 x 87 DPI)

SUPPORTING INFORMATION

Table S1. Information about the primer fragments used in this research.

Exon	Primer ¹	Sequence	Alignment temperature (°C)	Fragment size (no. of nucleotides)
1	Forward	5'- AGC TAC AGG CTG GAG ACA ATA -3'	60	668
	Reverse	5'- TTT AGT GAA GAA GAC AGC ATC AG -3'		
2	Forward	5'- TTG GCT TTC AGA TTA TTT GGA T -3'	62	329
	Reverse	5'- AAA GAT TTT CTA TTC TAT GCT CTG C -3'		
3	Forward	5'- ATA GGA AAT CAA TAC CAA AAC ACT T -3'	62	227
	Reverse	5'- TAT GGG TTA GAG GGT TGG ACT -3'		
4	Forward	5'- AAT CAG ACT CCC ATC CCA AT -3'	62	271
	Reverse	5'- TCC AGT TTC AAC TTG TTT CAG A -3'		
5	Forward	5'- GAC CCA TAC ATG AGT CAG TAG TT -3'	62	315
	Reverse	5'- AAG GAA GCA GAT TCA AGT AGG -3'		
6	Forward	5'- TAC TGA TGG GCC TGC TTC TC -3'	60	398
	Reverse	5'- AAT AGC CTC AGT CTC CCA CCT -3'		
7	Forward	5'- CTA TTC CTG TAA CCA GCA CAC A -3'	58	343
	Reverse	5'- CGA CGT GGG TTC TGA AAT TA -3'		
8	Forward	5'- GTC AGT GGT CCC AAG TAG TC -3'	58	705
	Reverse	5'- GGA AAG TGA TTA GTT AGT GAG AG -3'		

¹The primer sense of amplification.

SUPPORTING INFORMATION

Table S2. Phylogenetic conservation and missense mutations described in this study; in bold mutations not listed in www.factorix.org/

Mutation	Conservation*												
	HS	Mut	PT	BT	OA	SS	CF	FC	OC	MM	RN		
p.Arg43Gln	R	Q	R	R	na [†]	na [†]	R	R	na [†]	R	R		
p.Arg43Leu	R	L	R	R	na [†]	na [†]	R	R	na [†]	R	R		
p.Arg43Trp	R	W	R	R	na [†]	na [†]	R	R	na [†]	R	R		
p.Phe55Cys	F	C	F	F	na [†]	F	F	F	na [†]	F	F		
p.Cys64Tyr	C	Y	C	C	na [†]	C	C	C	na [†]	C	C		
p.Arg75Gln	R	Q	R	R	na [†]	R	R	R	na [†]	R	R		
p.Asp95Tyr	D	Y	D	D	na [†]	D	D	D	na [†]	D	D		
p.Cys134Arg	C	R	C	C	na [†]	C	C	C	na [†]	C	C		
p.Gln143Lys	Q	K	Q	Q	na [†]	Q	Q	Q	na [†]	Q	Q		
p.Val153Ala	V	A	V	V	na [†]	V	V	V	na [†]	V	I		
p.Cys157Arg	C	R	C	C	na [†]	C	C	C	na [†]	C	C		
p.Cys170Tyr	C	Y	C	C	na [†]	C	C	C	na [†]	C	C		
p.Arg191His	R	H	R	R	R	R	R	R	R	R	R		
p.Arg226Gln	R	Q	R	R	R	R	R	R	R	R	R		
p.Gly236Val	G	V	G	G	G	G	G	G	G	G	G		
p.His302Arg	H	R	R	R	Y	R	R	R	Y	R	R		
p.Thr342Met	T	M	T	T	T	T	S	T	T	T	T		
p.Ala366Gly	A	G	A	A	A	A	A	A	A	A	A		
p.Val374Phe	V	F	V	V	V	V	V	V	V	V	V		
p.Thr386Lys	T	K	T	T	T	T	T	T	T	T	T		
p.Ile390Thr	I	T	I	V	I	I	I	I	I	I	I		
p.Gly413Arg	G	R	G	G	G	G	G	G	G	G	G		
p.Gly432Val	G	V	G	G	G	G	G	G	G	G	G		

*Conservation of the amino acid positions in the following species: *Homo sapiens* (HS), HS mutated residue (Mut), *Pan troglodytes* (PT), *Bos taurus* (BT), *Ovis aries* (OA), *Sus scrofa* (SS), *Canis familiaris* (CF), *Felis catus* (FC), *Oryctolagus cuniculus* (OC), *Mus musculus* (MM), and *Rattus norvegicus* (RN).

[†]na: No alignment possible.

SUPPORTING INFORMATION

Table S3. Characteristics of mutations described in this study.

Mutation	Polyphen-2		PROVEAN		SIFT	Mutation Taster		Putative deleterious effect (HOPE and present bioinformatic analysis)
	Score	prediction	Score	Prediction (cutoff= -2.5)		Score	prediction	
p.Arg43Gln	1		-3.222	Deleterious	0		0.739	Alters charge and size; disordered region; possible loss of interactions
p.Arg43Leu	1		-5.572	Deleterious	0		0.724	Alters charge and hydrophobicity possible loss of interactions
p.Arg43Trp	1		-6.345	Deleterious	0		0.744	Alters charge, size and hydrophobicity.
p.Phe55Cys	0.995		-3.574	Deleterious	0		0.930	Loss of hydrogen bonds, alters the folding protein. Alters size and hydrophobicity.
p.Cys64Tyr	1		-9.319	Deleterious	0		0.999	Alters size and hydrophobicity. Loss of a disulphide bond, effect on the stability of the domain.
p.Arg75Gln	1		-3.081	Deleterious	0		0.999	Alters charge, size and hydrophobicity.
p.Asp95Tyr	1		-6.125	Deleterious	0		0.999	Disturbs the interaction with metal-ion Ca ⁺² . Alters charge, size and hydrophobicity. Loss of a salt bridge. Disturbs contact with residues in another domain.
p.Cys134Arg	1		-7.521	Deleterious	0		0.999	Loss of a disulphide bond.
p.Gln143Lys	0.991		-2.309	Neutral	0.02		0.731	Alters charge and size; possible repulsion of ligands or other residues with the same charge.
p.Val153Ala	0.001		-0.397	Neutral	0.2		0.739	Alters size.
p.Cys157Arg	1		-10.201	Deleterious	0		0.999	Alters charge, size and hydrophobicity. Loss of a disulphide bond and hydrophobic interactions on the core and surface.
p.Cys170Tyr	1		-9.320	Deleterious	0		0.999	Loss of a disulphide bond.
p.Arg191His	0.461		-2.178	Neutral	0.06		0.999	Alters charge, size. Residues in the vicinity are binding sites. Affects structure and binding site.

SUPPORTING INFORMATION

Table S3. (Cont.) Characteristics of mutations described in this study.

Mutation	Polyphen-2		PROVEAN		SIFT	Mutation Taster	Putative deleterious effect (HOPE and present bioinformatic analysis)
	Score	prediction	Score	Prediction (cutoff=-2.5)			
p.Thr194Ala	0.003		-0.922	Neutral	0.54	0.999	Alters size and hydrofobicity.
p.Arg226Gln	0.998		-3.356	Deleterious	0	0.999	Alters size and charge.
p.Gly236Val	1		-8.154	Deleterious	0	0.999	Alters size. Possible modification in the center of FIX.
p.His302Arg	1		-6.791	Deleterious	0.01	0.999	Alters size and charge. This residue is connected with Asp304 through a hydrogen bond.
p.Thr342Met	1		-2.992	Deleterious	0	0.999	Alters size and hydrofobicity. Possible damage on the hydrogen bond made with the 338 position.
p.Ala366Gly	0.095		-2.788	Deleterious	0.01	0.999	Alters size and hydrofobicity. Possible loss of rigidity.
p.Val374Phe	0.999		-4.537	Deleterious	0	0.999	Alters size. Possible modification in the center of FIX.
p.Thr386Lys	0.995		-3.108	Deleterious	0	0.999	Alters charge, size and hydrofobicity. Affects a hydrogen bond formation with residues 388 and 433.
p.Ile390Thr	0.977		-4.649	Deleterious	0.01	0.999	Alters size and hydrofobicity. Possible empty space on the core of FIX and loss of hydrofobic interactions.
p.Gly413Arg	1		-7.954	Deleterious	0	0.999	Alters size, charge and hydrofobicity. Possible loss of flexibility.
p.Gly432Val	1		-8.737	Deleterious	0	0.999	Alters size and hydrofobicity. The interaction with FIX ligand can be affected. Possible loss of flexibility.

*Sequence changes at protein level following the Nomenclature for Description of Genetic Variations approved by the Human Genome Variation Society.

CAPÍTULO 3—ELECTROSTATIC ANALYSIS OF HEMOPHILIA B CAUSATIVE MUTATIONS

Artigo em preparação para ser submetido à revista *Journal of Thrombosis and Haemostasis*.

Electrostatic analysis of Hemophilia B causative mutations

M. R. MEIRELES, G. F. VIEIRA, F. M. SALZANO and E. BANDINELLI

Departamento de Genética, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil

Running title: Hemophilia B structural analysis

Word counts: Abstract: 212; Text:

Correspondence: Francisco M. Salzano, Departamento de Genética, Instituto de Biociências, UFRGS, Caixa Postal 15053, 91501-970 Porto Alegre, RS, Brazil.
Tel.: +55 51 3308-6747; fax: 55 51 3308-9823
E-mail: francisco.salzano@ufrgs.br

Essentials

Haemophilia B is understudied as compared to haemophilia A.

3D models were generated by homology tools and electrostatic patterns clusterized.

Electrostatic changes are mainly determined by the mutations positions in the molecule.

Electrostatic and clinical patterns do not necessarily coincide.

Summary

Background: Factor IX (FIX) is an important protein in the coagulation process, its lack or deficiency leading to haemophilia B (HB). This condition has been much less studied than haemophilia A, especially in Latin America. We are trying to fill this gap through a previous study which described 43 patients from southern Brazil, who carried 31 different mutations, investigated with molecular methods; and with this report, which considered in detail their missense mutations. *Objectives:* To examine the electrostatic patterns (EP) of the missense mutations found, and their clinical effects, comparing them with mutations described elsewhere affecting the same residues. *Methods:* 3D models were generated by homology approaches, and their electrostatic patterns analyzed with clusterization tools. *Results:* (a) Different mutations in the same site generally clustered together in relation to EP; (b) These patterns are not related to the mutations clinical effects; and (c) The mutations and their effects do not occur preferentially considering the molecule's gamma-carboxy glutamic (GLA), epidermal growth factors (EGF1 and EGF2), and serine protease (SPROT) domains. *Conclusions:* Molecular structures, due to their properties, condition different clinical effects. We have considered here what happens in terms of electrostatic changes. These changes are mainly determined by the mutations positions in the molecule, but electrostatic and clinical patterns do not necessarily coincide.

Keywords: electrostatic potentials, haemophilia B, missense mutations, phenotype-genotype correlations, protein structure

Introduction

Haemophilia B (HB) is a coagulation disorder due to mutations in the *F9* gene, located in chromosome Xq27.1 and that has eight exons and seven introns. The corresponding protein, Factor IX (FIX), has a mature peptide of 415 amino acids (aa) with three basic domains: gamma-carboxy glutamic (GLA) (47 aa); epidermal growth factors 1 and 2 (EGF1 and EGF2) (98 aa), and serine protease (SPROT) (235 aa). GLA+EGF1+EGF2 together comprise the molecule's light chain, while SPROT is the catalytic heavy chain domain.

HB has been much less studied than hemophilia A, and this is especially notable in Latin America. To fill this gap we recently performed an investigation about its molecular characterization in southern Brazil [1]. As was verified in other series, the majority of the mutations found was of the missense type, and the present report examines in detail these mutations, comparing them with those described by other authors in the same sites, using appropriate bioinformatic tools. The results in relation to electrostatic patterns (EP) have been considered, with the objective of answering the following questions: (a) How different missense mutations affecting the same protein residue are reflected in the HB molecule? (b) Are the similarities/differences found related to the mutations' clinical effects? and (c) Their presence in different domains of the resulting protein influences the patterns found?

Methods

A total of 54 missense mutations was included in the analysis, considering their clinical effects (severe, moderate or mild HB) through the information given in [1] and in the Factor IX database (www.factorix.org).

For each domain present in the mature FIX protein, a 3D model was generated by homology approaches based in Protein Data Bank[2] (PDB) structures as follows: 1pfx for the light chain (GLA, EGF1 and EGF2 domains) and 1rfn for the heavy chain (SPROT, Serine Protease domain). The molecular modeling tools used for the construction of the chains were I-tasser (Iterative Threading ASSEmbleyRefinement) [3, 4] and Phyre2 [5], in

one-to-one threading mode. These mode options were based on the MOD-FOLD [6] and PDB sum[7] evaluation results.

Phyre2 was also used to generate 3D mutated models by homology approaches. This tool was applied in its expert mode, one-to-one threading job, with the modeling of the amino acids mutated sequences based in the wild template. The Delphi web server[8] was applied for the evaluation of the electrostatic potential (EP) calculations in each model and the Chimera [9] software to obtain model visualization and EP applications.

To obtain heat maps and epograms of electrostatic distances between structures, the webPIPSA (Protein Interaction Property Similarity Analysis) was employed. The webPIPSA derived information was compared based on manual implementation of electrostatic potential Delphi calculations in the mutated models using Chimera.

Results and Discussion

Table 1 lists the 54 mutations, indicating the corresponding molecular domains and clinical effects. As indicated there, seven of them were present in the GLA, 19 in the EGF1 and EGF2, and 28 in the Serine Protease domains. For the 51 that had their phenotypic effects ascertained 28 (55%) determined severe, 13 (25%) moderate, and 10 (20%) mild effects. Eighteen of them were found by us in our previous study [1].

Results in relation to the GLA domain are presented in Table 2, and in Figs. 1 and Supplementary Information S1. Three clusters were obtained, one including the wild type (WS¹) and four mutations (two observed by us in our study [1]. The cluster was formed by mutations in two sites (55 and 64) and the clinical effects of these mutations were of the three types, indicating no influence of the changes in electrostatic potentials in them. The other mutation found by us on residue 75, together with another in the same site, showed an electrostatic distance of 0.6 from the normal molecule, but these changes have not resulted in a severe disease.

Six electrostatic clusters were obtained in relation to the 19 mutations affecting the EGF1 and EGF2 domains (Table 2; Figs. 2 and S2). Five of them (including the mutation detected by us, Cys170Tyr) clustered with WS¹, and therefore, despite their severe clinical effects, did not differ from normal in EP.

The other five clusters showed a distinct (1.5) distance from the normal EP, and the four other mutations observed in our study [1] clustered in four of them. No consistent effect of type of residue or clinical effect was detected in these clusters.

The 28 mutations affecting the Serine Protease domains formed six clusters. Sixteen mutations (57%) clustered with WS¹, while the others were about evenly distributed in the five other clusters with distinct (0.4-0.8) electrostatic distances from WS¹. Again, no distinct relationship could be found between EP with protein residues or mutation effects.

To verify whether the 51 mutations would occur preferentially in a given protein domain, we tested their distribution using Fisher's exact test. The results indicated no preferential location (p:0.387).

Despite these negative findings, we did detect a specific structural pattern of electrostatic changes in the EGF1 and EGF2 domains (Fig. S4). While the mutation patterns generally lead to heterogeneous EPs, the lower right portion of the molecule shows a homogeneous surface picture independently of the mutated residues. This general degree of change may affect the protein's binding with its cofactors.

We are now in a position to answer the questions posed in the introduction: (a) different mutations in the same residue generally cluster together in relation to electrostatic patterns [in a total of 53 comparisons, mutations in the same site cluster together in 37 instances (70%), a value significantly different (p<0.01)] from that of no difference in cluster position]; (b) these patterns are not related to clinical effects [for instance the three mutations considered in residue 55 clustered together, despite having completely different effects (mild, moderate, severe)]; and (c) as indicated earlier, the mutations investigated do not occur preferentially in a given domain, and their distribution in clinical types do not show statistically significant differences, despite the fact that the severe form is most prevalent in our sample, in the EGF1+EGF2 domain (75%), against 43% and 47%, respectively, in the GLA and SPROT domains.

Addendum

Authors contributions: M. R. Meireles performed the laboratory and bioinformatic tests and analyzed the data; G. F. Vieira coordinated and supervised the bioinformatic

analyses; F. M. Salzano supervised all stages of the work; E. Bandinelli supervised the laboratory tests; and all authors contributed to the writing of the paper.

Acknowledgments

We thank the patients and staff of HEMOCENTRO-RS for agreeing to participate in the study and for logistical support. Financing for this work was provided by Conselho Nacional de Desenvolvimento Científico e Tecnológico and Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (Apoio a Núcleos de Excelência Program).

References

1. Meireles MR, Pantoja AG, Ornaghi APM, Vieira GF, Salzano FM, Bandinelli E. Molecular characterization of haemophilia B patients in southern Brazil (submitted).
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000;**28**:235-42.
3. Zhang Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008; **9**:40.
4. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. *Nat Methods* 2015;**12**:7-8.
5. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc* 2015;**10**:845-58.
6. McGuffin LJ, Buenavista MT, Roche DB. The ModFOLD4 server for the quality assessment of 3D protein models. *Nucleic Acids Res* 2013;**41**:W368-72.
7. Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, Thornton JM. PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem Sci* 1997;**22**:488-90.
8. Smith N, Witham S, Sarkar S, Zhang J, Li L, Li C, Alexov E. DelPhi web server v2: incorporating atomic-style geometrical figures into the computational protocol. *Bioinformatics* 2012;**28**:1655-57.
9. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera – a visualization system for exploratory research and analysis. *J Comput Chem* 2004;**25**:1605-12.

Figure legends

Fig. 1. On the left is the clusterization of the GLA domain mutation models with the wild structure (WS¹). The epogram and scale indicate electrostatic distances among models and were generated in the webPIPSA tool. Symbols ▲ indicate severe, ¹ moderate, and * mild clinical effects. On the right are representative models of each cluster. The models were elaborated with the Phyre2 web portal and electrostatic potentials calculated with the Delphi webserver, both implemented with the Chimera Interface. Mutations detected by our group in a previous study [1] are in bold.

Fig. 2. On the left is the clusterization of the EGF1 and EGF2 domains mutation models with the wild structure (WS¹). The epogram and scale indicate electrostatic distances among models and were generated in the webPIPSA tool. Symbols ▲ indicate severe, ¹ moderate, and * mild clinical effects. Three mutations have no information² about phenotype. On the right are representative models of each cluster. The models were elaborated with the Phyre2 web portal and electrostatic potentials calculated with the Delphi webserver, both implemented with the Chimera Interface. Mutations detected by our group in a previous study [1] are in bold.

Fig. 3. On the left is the clusterization of the Serine protease domain mutation models with the wild structure (WS¹). The epogram and scale indicate electrostatic distances among models and were generated in the webPIPSA tool. Symbols ▲ indicate severe, ¹ moderate, and * mild clinical effects. On the right are representative models of each cluster. The models were elaborated with the Phyre2 web portal and electrostatic potentials calculated with Delphi webserver, both implemented with the Chimera Interface. Mutations detected by our group in a previous study [1] are in bold.

Table 1 Mutations analyzed in the present study with their molecular locations and clinical effects

Mutation ¹	Exon	Domain	Clinical severity ²	Mutation ¹	Exon	Domain	Clinical severity ²
p.Phe55Cys	2	GLA	moderate	p.Gly236Asp	6	Serine Protease	moderate
p.Phe55Ile	2	GLA	mild	p.Gly236Cys	6	Serine Protease	moderate
p.Phe55Ser	2	GLA	severe	p.His302Arg	8	Serine Protease	mild
p.Cys64Tyr	2	GLA	severe	p.Thr342Met	8	Serine Protease	moderate
p.Cys64Arg	2	GLA	severe	p.Thr342Ala	8	Serine Protease	mild
p.Arg75Gln	2	GLA	mild	p.Thr342Arg	8	Serine Protease	severe
p.Arg75Pro	2	GLA	moderate	p.Thr342Lys	8	Serine Protease	mild
p.Asp95Tyr	4	EGF1	mild	p.Ala366Gly	8	Serine Protease	mild
p.Asp95Gly	4	EGF1	moderate	p.Ala366Asp	8	Serine Protease	moderate
p.Asp95Val	4	EGF1	ni³	p.Ala366Pro	8	Serine Protease	mild
p.Cys134Arg	5	EGF2	severe	p.Ala366Val	8	Serine Protease	mild
p.Cys134Phe	5	EGF2	ni³	p.Val374Phe	8	Serine Protease	severe
p.Cys134Ser	5	EGF2	severe	p.Val374Ile	8	Serine Protease	moderate
p.Gln143Lys	5	EGF2	severe	p.Thr386Lys	8	Serine Protease	severe
p.Gln143Arg	5	EGF2	ni³	p.Thr386Arg	8	Serine Protease	severe
p.Gln143Glu	5	EGF2	severe	p.Thr386Ile	8	Serine Protease	moderate
p.Gln143Pro	5	EGF2	severe	p.Ile390Thr	8	Serine Protease	moderate
p.Val153Ala	5	EGF2	moderate	p.Ile390Asn	8	Serine Protease	severe
p.Val153Met	5	EGF2	mild	p.Ile390Phe	8	Serine Protease	severe
p.Cys157Arg	5	EGF2	severe	p.Ile390Ser	8	Serine Protease	moderate
p.Cys157Ser	5	EGF2	severe	p.Gly413Arg	8	Serine Protease	severe
p.Cys157Tyr	5	EGF2	severe	p.Gly413Ala	8	Serine Protease	moderate
p.Cys170Tyr	5	EGF2	severe	p.Gly413Glu	8	Serine Protease	severe
p.Cys170Arg	5	EGF2	severe	p.Gly432Val	8	Serine Protease	severe
p.Cys170Phe	5	EGF2	severe	p.Gly432Ala	8	Serine Protease	severe
p.Cys170Ser	5	EGF2	severe	p.Gly432Asp	8	Serine Protease	severe
p.Gly236Val	6	Serine Protease	severe	p.Gly432Ser	8	Serine Protease	severe

¹Sequence changes at protein level following the Nomenclature for Description of Genetic Variations approved by the Human Genome Variation Society. Mutations detected by us in a previous study in southern Brazil [1] are in bold.

²Severity was ascertained based on our own [1] and factorix.org information.

³ni: no information on clinical effects given in factorix.org.

Table 2 Mutations included in the study classified according to domain and clusterization by electrostatic distances

Domain	Cluster ¹	N° of mutated models	Amino acid changes ² /cluster	Approximate distance ³ from the wild structure
GLA	1	4	p.Phe55Cys , p.Phe55Ile, p.Phe55Ser, p.Cys64Tyr	0.0
	2	1	p.Cys64Arg	0.3
	3	2	p.Arg75Gln , p.Arg75Pro	0.6
EGF1 and EGF2	1	5	p.Cys134Ser, p.Val153Ala , p.Val153Met, p.Cys157Ser, p.Cys157Tyr	1.5
	2	1	p.Gln143Glu	1.5
	3	2	p.Cys134Arg , p.Cys170Arg	1.5
	4	3	p.Gln143Arg, p.Gln143Lys , p.Cys157Arg	1.5
	5	3	p.Asp95Gly, p.Asp95Tyr , p.Asp95Val	1.5
	6	5	p.Cys134Phe, p.Gln143Pro, p.Cys170Phe, p.Cys170Ser, p.Cys170Tyr	0.0
Serine Protease	1	3	p.Thr342Arg, p.Thr342Lys, p.Gly413Arg	0.8
	2	2	p.Thr386Arg, p.Thr386Lys	0.8
	3	4	p.Hys302Arg , p.Ala366Asp, p.Ala366Val, p.Ile390Thr	0.8
	4	2	p.Gly413Glu, p.Gly432Asp	0.4
	5	1	p.Gly236Asp	0.4
	6	16	p.Gly236Cys, p.Gly236Val , p.Thr342Ala, p.Thr342Met , p.Ala366Gly , p.Ala366Pro, p.Val374Ile, p.Val374Phe , p.Thr386Ile, p.Ile390Asn, p.Ile390Phe, p.Ile390Ser, p.Gly413Ala, p.Gly432Ala, p.Gly432Ser, p.Gly432Val	0.0

¹Clusterization was obtained using the webPIPSA tool, based on a pair-to-pair comparison. Mutations detected by us in a previous study in southern Brazil [1] are in bold.

²Sequence changes at the protein level following the Nomenclature for Description of Genetic Variations approved by the Human Genome Variation Society.

³Distances were calculated with webPIPSA by a pair-to-pair comparison and matrix generation ($D_{a,b} = \sqrt{2-SI_{a,b}}$).

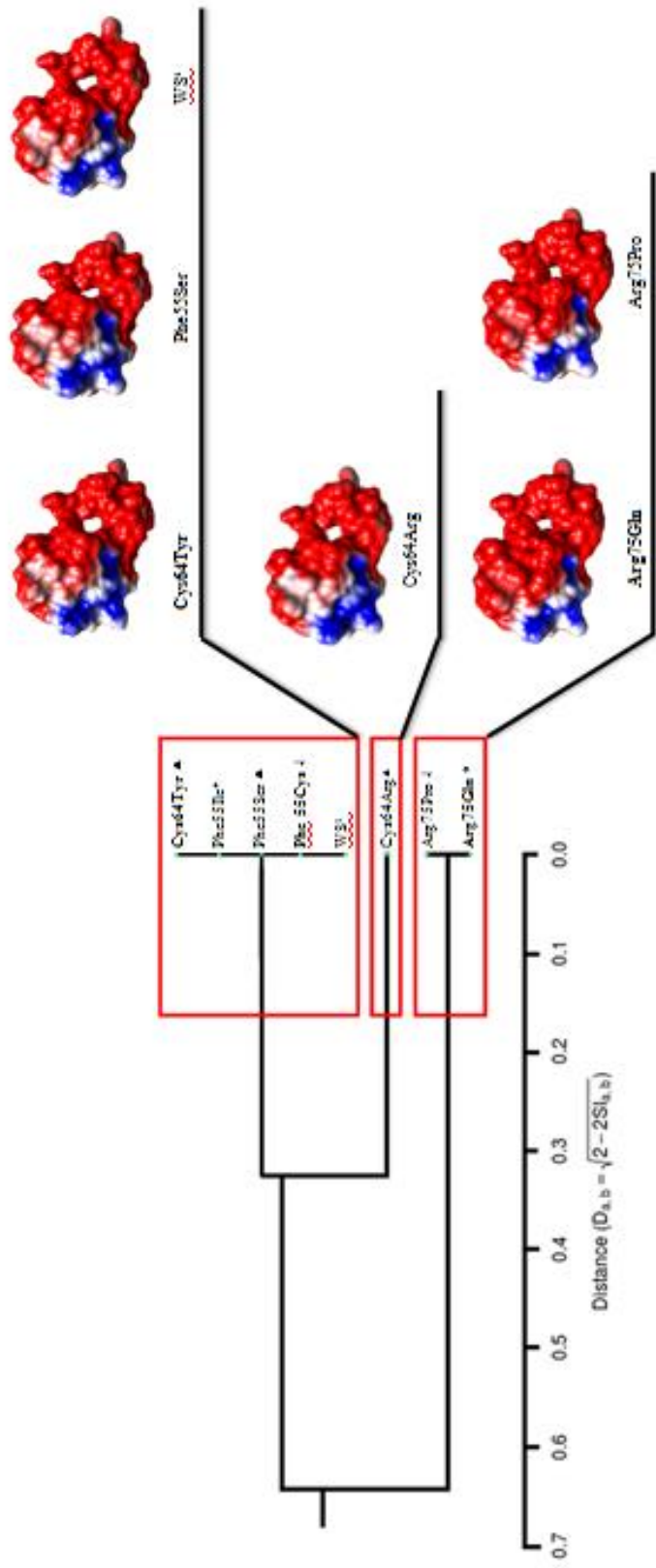


Figure 1

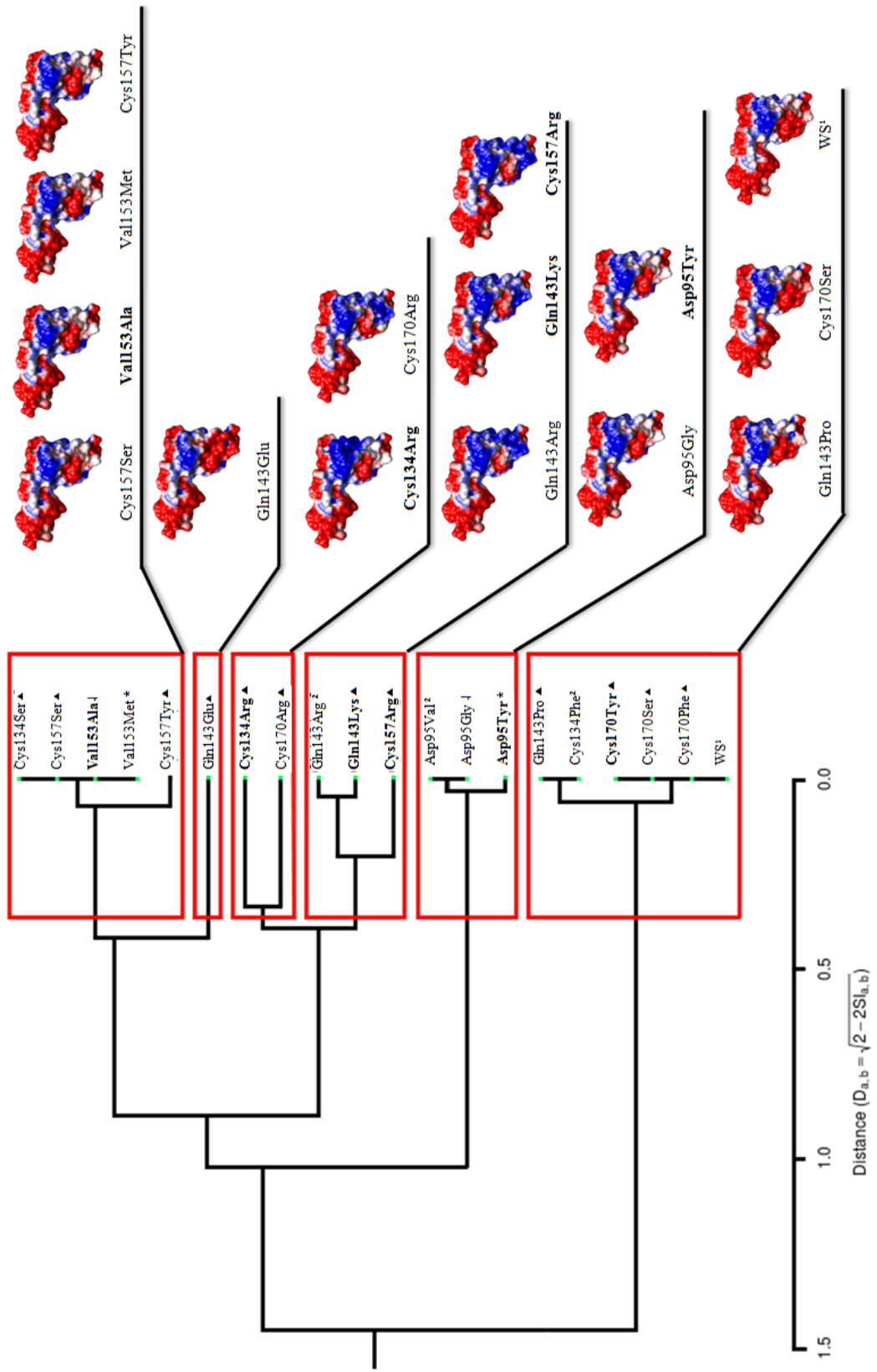


Figure 2

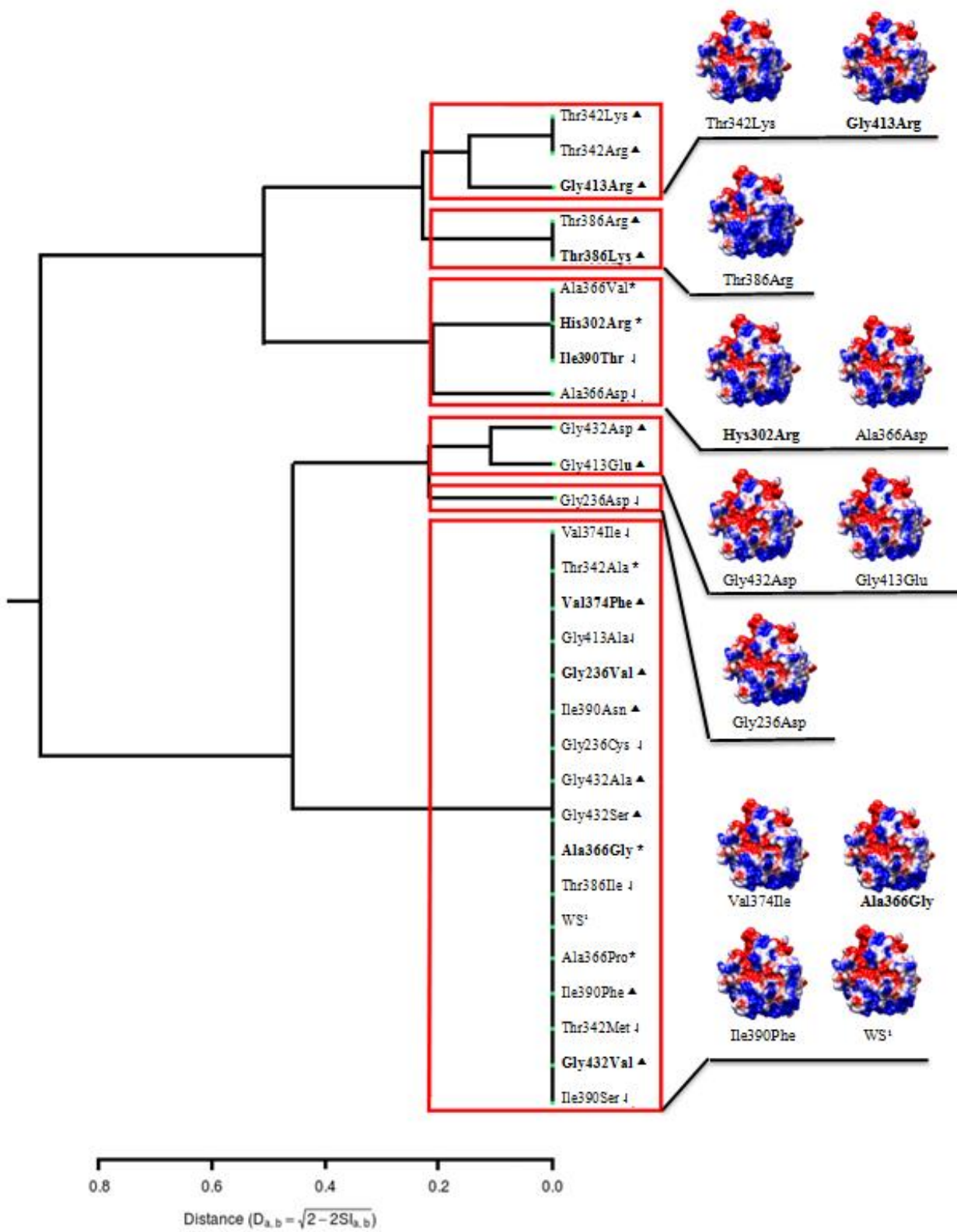


Figure 3

SUPPORTING INFORMATION

Fig. S1. Heatmap illustrating similarities among models generated by a pair-to-pair comparison in the GLA domain. The picture was created with the webPIPSA tool. Color key and density plots are indicated in the upper left side of the figure. In **bold**, mutations previously detected by our group [1] in the southern Brazilian population.

Fig. S2. Heatmap illustrating similarities among models generated by a pair-to-pair comparison in the EGF1 and EGF2 domains. The picture was created with the webPIPSA tool. Color key and density plots are indicated in the upper left side of the figure. WS¹ indicates the wild structure. In **bold**, mutations previously detected by our group [1] in the southern Brazilian population.

Fig. S3. Heatmap illustrating similarities among models generated by a pair-to-pair comparison in the Serine protease domain. The picture was created with the webPIPSA tool. Color key and density plots are indicated in the upper left side of the figure. WS¹ indicates the wild structure. In **bold**, mutations previously detected by our group [1] in the southern Brazilian population.

Fig. S4. Identification of a pattern in electrostatic changes at the EGF1 and EGF2 domains. In A, a representative mutated model; in B, the pattern derived (the specific surface region that have EP alterations independent of the residues position are shown in pink); in C, the wild structure; in D, mutated model with a semi-transparent surface showing amino acids residues in the backbone; and in E, the pattern isolated in the surface with the model backbone.

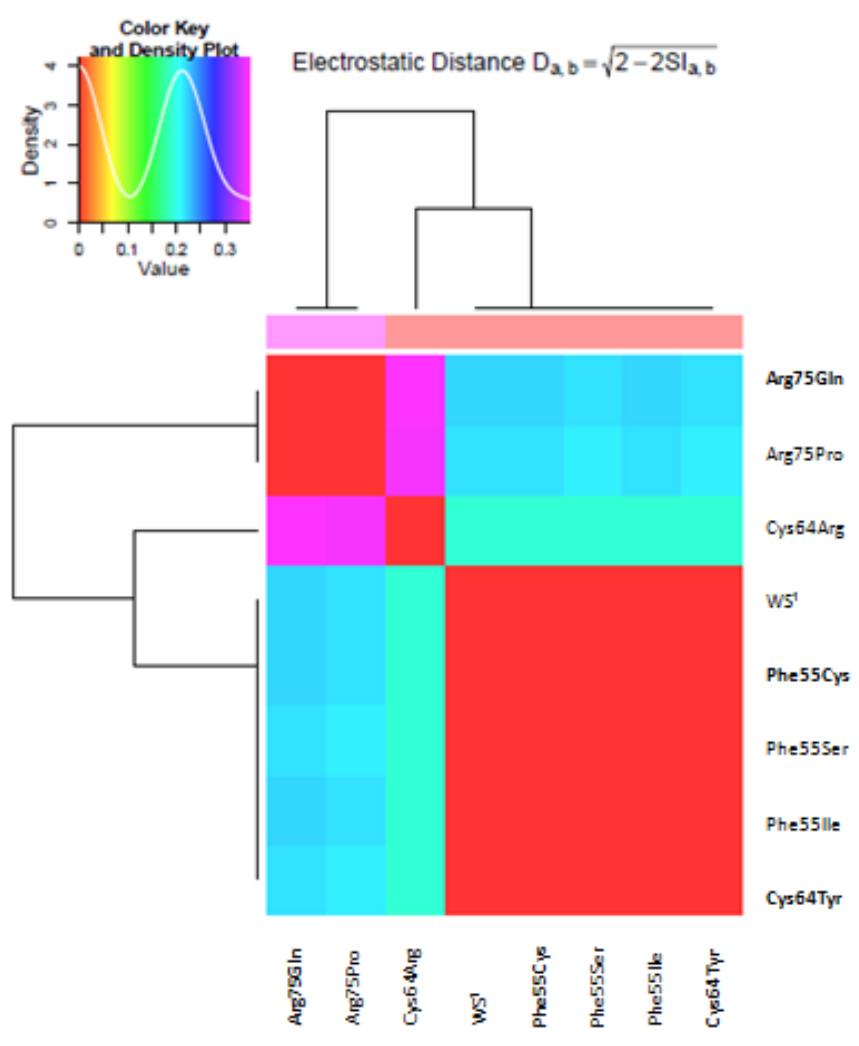


Figure S1

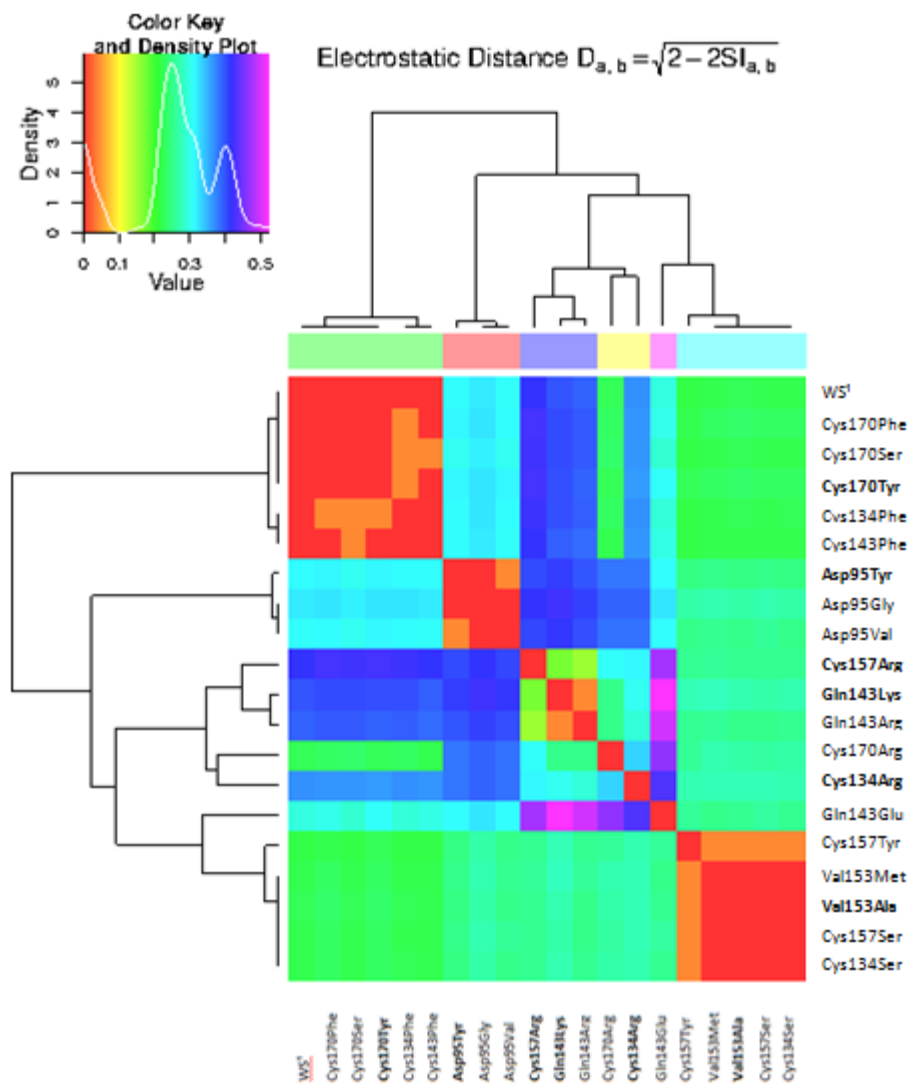


Figure S2

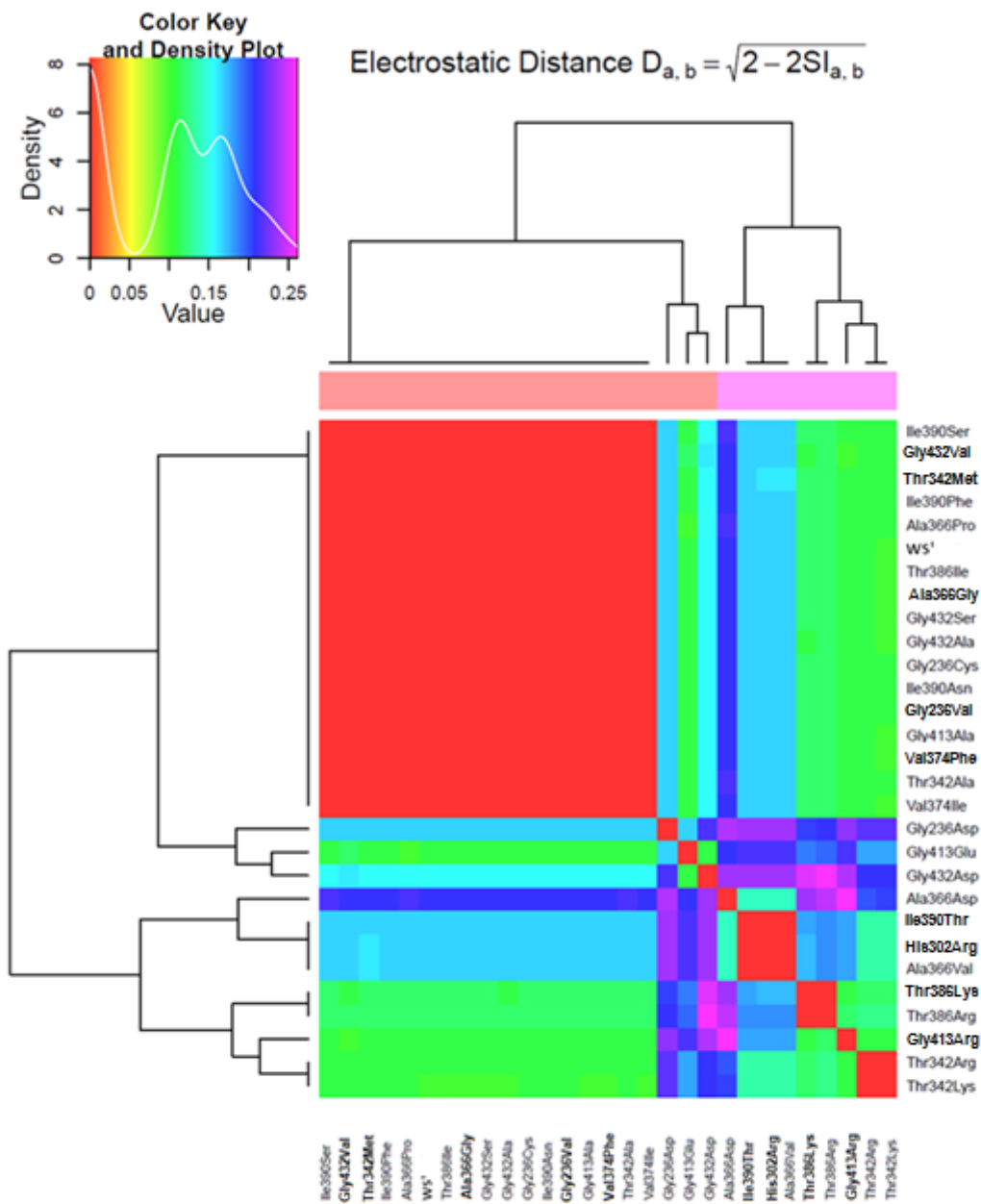


Figure S3

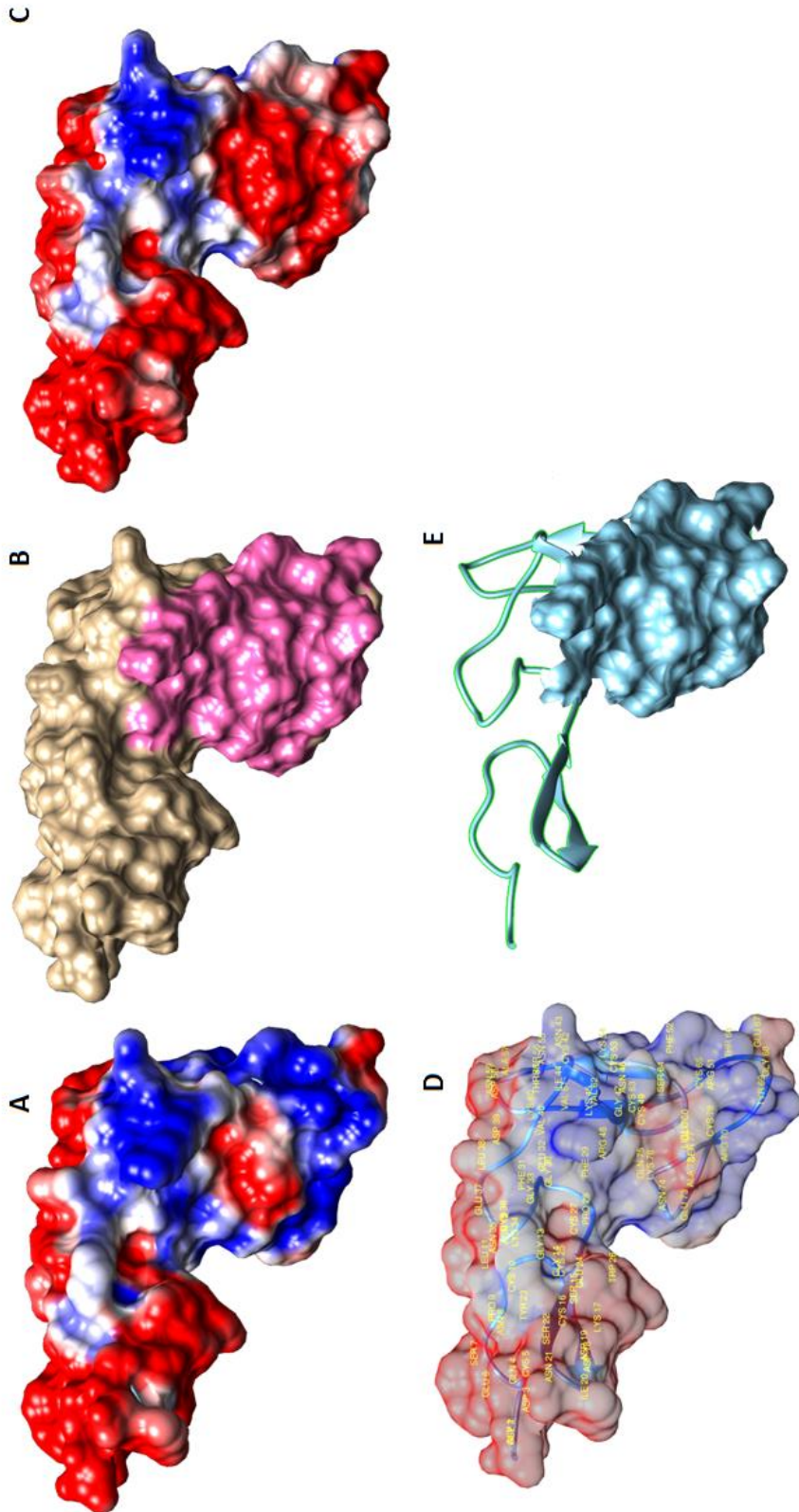


Figure S4

CAPÍTULO 4 – DISCUSSÃO GERAL

A hemofilia tem o seu primeiro registro no livro de condutas do povo judeu (Talmude), do século II d.C. Este livro fornece os princípios religiosos, baseados na lei oral, na doutrina, na moral e nas tradições. O mesmo isenta da circuncisão crianças provenientes de famílias que tenham sofrido duas mortes decorrentes de hemorragias após o procedimento (DiMichele e Neufeld, 1998). A hemofilia possui, portanto, mais de 1800 anos de história, e com o advento de novas tecnologias tem sido possível firmar uma nova luz às investigações dessa condição.

Desde a descoberta do gene do fator IX, muitos estudos foram publicados com a temática da hemofilia B. Recentemente, foi feito um artigo de revisão para definir o *status quo* de pesquisas publicadas e experimentos de tratamentos registrados (Berger et al., 2015). Para tal, foi feita uma busca sistemática nos bancos de dados EMBASE e MEDLINE usando os termos “*Hemophilia B*”, “*Haemophilia B*” e “*Factor IX*”, tendo sido obtidos um total de 14.908 estudos. Em seguida houve limitação para pesquisas que ocorreram após 2009 e em humanos, e seu total passou para 1.639 artigos e resumos. Foram encontrados estudos de revisão, relatos de casos, comentários, editoriais, cartas aos editores e notas; inclusive sobre a segurança, eficácia e farmacocinética de produtos de tratamento e sobre terapia gênica.

Um grande número de pesquisas tem seu foco na caracterização de mutações em populações específicas. Se observarmos os dados da *World Federation of Hemophilia*, pode-se verificar que a frequência da doença varia nos diferentes países. Além disso, a hemofilia B possui grande heterogeneidade alélica, possibilitando que existam mutações específicas a determinadas populações. Tal fenômeno torna importante que haja estudos de ampla caracterização populacional, para que se disponha de informações que possibilitem uma melhor compreensão global da genética molecular da doença.

A identificação de variantes no F9 fornece informações para que sejam aprimoradas análises bioquímicas e moleculares de predições de alterações na estrutura tridimensional da molécula da proteína. A partir dessas informações, podem-se realizar melhores inferências para correlacionar o genótipo com o fenótipo e sugerir mecanismos de causas

da doença. Levando em conta o padrão de herança da hemofilia B, o diagnóstico molecular também se torna uma importante ferramenta para o aconselhamento genético em famílias, incluindo testes pré-natais em membros de genealogias com histórico familiar da doença, identificação de mulheres heterozigotas para as mutações, e particularidades da doença com risco de recorrência.

O conhecimento sobre as mutações também é importante para o tratamento da HB, pois algumas mutações podem levar a uma resposta inesperada, sendo necessário outro tipo de terapia. Alguns pacientes com hemofilia B podem desenvolver inibidores ao tratamento. Conforme Rallapalli et al. (2013), de 3741 pacientes analisados, 1.6% (n=59) desenvolveram inibidores ao FIX recombinante. Na maioria dos casos, o desenvolvimento de inibidor está associado com grandes deleções de um ou mais éxons, ou mutações sem sentido. Considerando as grandes deleções, 25% daquelas descritas no banco (27/107) levaram à formação de inibidores. A presença de inibidor foi registrada apenas em uma mutação de sentido trocado (p.Gln237Lys) e em duas mudanças de nucleotídeo único que modificaram o processamento. Em nosso estudo, a presença de inibidores foi observada em apenas um paciente. Ele possui uma mutação sem sentido, já descrita no banco de dados (p.Arg75*) em 73 pacientes, dos quais oito apresentaram inibidor. Outros dois pacientes independentes também apresentaram essa mutação em nosso estudo, porém não desenvolveram inibidor. Não foram encontradas grandes deleções em nossa análise, apenas uma pequena deleção p.Arg358Serfs* que já havia sido descrita no banco de dados em um paciente com hemofilia B grave (Costa et al., 2000), o mesmo fenótipo que encontramos.

As mutações no F9 podem ser recorrentes dentro de uma população. Este fenômeno pode ser verificado especialmente devido ao efeito do fundador, que leva a uma reprodução dessa mutação em gerações sucessivas. Considerando a população no sul do Brasil, oito mutações diferentes foram encontradas em mais de um paciente independente, são elas: Arg43Gln, Arg75*, Arg75Gln, Cys170Tyr, Arg226Gln, Arg298*, Thr342Met e Cys157Arg, sendo que a última não está descrita em www.factorix.org. As sete primeiras foram encontradas em outras pesquisas, respectivamente em 85, 73, 18, 6, 52, 63 e 128 pacientes, distribuídos em diferentes países. Algumas destas mutações foram encontradas em pacientes da América Latina, em países como a Argentina e o México.

O presente trabalho é o primeiro que aborda a hemofilia B no sul do Brasil, e forneceu o diagnóstico molecular para 43 pacientes. Foram encontradas 30 mutações distintas que foram consideradas como causadoras da desordem. Cabe ressaltar que seis delas não haviam sido descritas no banco de dados do fator IX. Todas as mutações de sentido trocado eram em resíduos altamente conservados entre diferentes espécies de mamíferos. A partir dessa informação pode-se sugerir que tais resíduos estão envolvidos na função da proteína, seja por contribuírem a detalhes de sua estrutura tridimensional, seja por conferirem a sua estabilidade ou por serem locais de interação com outras moléculas.

Quando se comparam os 54 modelos de mutações de sentido trocado com as estruturas normais dos domínios GLA, EGF1, EGF2 e Serino-protease em termos de diferença eletrostática, pode-se verificar alguns padrões. Aminoácidos mutantes com carga positiva (pH=7) sempre mostraram diferenças com relação ao fenótipo normal. Todas as substituições envolvendo arginina, histidina e lisina, seja como resíduo normal ou como resíduo mutante, levaram os modelos a afastarem-se em termos eletrostáticos da estrutura sem mutações. A carga positiva é determinante para as interações eletrostáticas em uma proteína e, portanto, a perda ou ganho dela possivelmente interfere com a função protéica. Aminoácidos carregados negativamente (possuem COO^- em pH= 7), como o aspartato (*Asp*) e o glutamato (*Glu*) funcionam como sítios de ligação a metais; e todos os modelos analisados que envolviam mutações com esses resíduos também divergiram eletrostaticamente do modelo estrutural normal.

As características dos aminoácidos são cruciais para o entendimento das mutações, especialmente aquelas de sentido trocado (*missense*). Isso porque, tanto o aminoácido normal quanto o mutante têm características próprias, e diferentes substituições em uma mesma posição podem originar fenótipos divergentes. A estrutura da proteína depende da sequência primária de aminoácidos e de interações com outras moléculas. O número de padrões de dobramento das proteínas é muito grande, mas limitado; conforme o dobramento as proteínas terão uma melhor ou pior capacidade de movimentação, e o movimento está vinculado com a função.

Do ponto de vista estrutural, as proteínas possuem interações covalentes e não covalentes, que estabilizam a sua conformação. Nas interações covalentes, temos as pontes dissulfeto entre resíduos de cisteína; nas interações não-covalentes temos as pontes de

hidrogênio, que estão presentes em número elevado e fazem ligações intramoleculares para deixar a estrutura o mais estável possível. Exemplos são as interações hidrofóbicas, que são um importante “impulso” no dobramento da proteína, induzindo a formação de uma estrutura compacta; interações iônicas; e as forças de Van der Waals, que embora sejam fracas, contribuem para a estabilidade da molécula. A estrutura secundária das proteínas também relaciona-se aos seus resíduos. Alguns aminoácidos tendem a formar alfa-hélice, por exemplo, enquanto que a prolina tende a quebrá-las. Regiões de alças geralmente são mais desorganizadas e móveis, e a glicina é um aminoácido conveniente nesse caso. Já a prolina força a formação de alças beta. As estruturas secundárias associam-se e são influenciadas por interações, que fazem o dobramento, reduzindo a superfície exposta ao solvente e estabilizando a proteína.

Muitas das mutações que foram observadas no nosso estudo não apresentam diferença eletrostática com relação à estrutura normal. Porém, devem-se levar em conta todos os aspectos dos aminoácidos (carga, polaridade, influência na estrutura secundária), as interações intramoleculares existentes no FIX (pontes de hidrogênio e dissulfeto), as regiões de ligações com outras moléculas (ligação a outros fatores da coagulação ou a outras enzimas), a acessibilidade ao solvente, e regiões de modificações pós-traducionais (glicosilação, fosforilação), para entender melhor os fenótipos apresentados.

CAPÍTULO 5 – CONCLUSÕES E PERSPECTIVAS

O nosso estudo indicou a causa da hemofilia B em 43 pacientes sem parentesco próximo, e detectou seis mutações que não estavam descritas no banco de dados do fator IX, quatro mutações de sentido trocado, uma em sítio de processamento e outra de mudança na fase de leitura. A identificação dessas novas mutações reforça indicações anteriores sobre a alta mutabilidade do gene F9, e contribui para compreensão da doença. Até o momento não existem muitos estudos que abordem a caracterização molecular da hemofilia B no Brasil, que são muito importantes para o tratamento e aconselhamento genético da doença.

Análises *in silico* envolvendo ferramentas de bioinformática têm trazido relevantes informações para estudos em genética e biologia molecular, seja como análises prévias ou complementares. Em nosso estudo, foi possível o reconhecimento de alterações eletrostáticas geradas por 29 de 54 mutações (54%) de sentido trocado. Nesses casos, a substituição de aminoácidos afastou os modelos estruturais da proteína normal. Verificamos também que muitos dos modelos apresentaram perda de pontes dissulfeto ou de hidrogênio, e modificações em sua polaridade e/ou hidrofobicidade. Apesar da grande validade de tais informações, o quanto cada um desses fatores está influenciando no fenótipo permanece uma incógnita. O projeto realizou investigações relevantes no campo da bioinformática, que envolveram o uso de preditores, modelagem por homologia, cálculos de diferenças eletrostáticas. Porém, todas as avaliações foram realizadas em modelos estáticos, e não foi possível uma real exploração do quanto tais propriedades afetam a proteína. Para isso, outras análises *in silico* envolvendo dinâmica molecular e ancoragens moleculares seriam valiosas, possibilitando também esclarecer os sítios de interação do FIX com outros fatores da coagulação sanguínea e com outras moléculas. Essas são as metas de nossos próximos estudos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Anson DS, Choo KH, Rees DJ, Giannelli F, Gould K, Huddleston JA, Brownlee GG (1984). The gene structure of human anti-haemophilic factor IX. *EMBO J* **3**, 1053-60.
- Arruda VR, Hagstrom JN, Deitch J, Heiman-Patterson T, Camire RM, Chu K, Fields PA, Herzog RW, Couto LB, Larson PJ, High KA (2001). Posttranslational modifications of recombinant myotube-synthesized human factor IX. *Blood* **97**, 130-8.
- Autin L, Miteva MA, Lee WH, Mertens K, Radtke KP, Villoutreix BO (2005). Molecular models of the procoagulant factor VIIIa-factor IXa complex. *J Thromb Haemost* **3**, 2044-56.
- Becker RC (2005). Cell-based models of coagulation: a paradigm in evolution. *J Thromb Thrombolysis* **20**, 65-8.
- Bentley AK, Rees DJ, Rizza C, Brownlee GG (1986). Defective propeptide processing of blood clotting factor IX caused by mutation of arginine to glutamine at position -4. *Cell* **45**, 343-8.
- Berger K, Schopohl D, Hilger A, Behr Gross ME, Giangrande P, Peyvandi F, Seitz R, Schramm W (2015). Research in haemophilia B--approaching the request for high evidence levels in a rare disease. *Haemophilia* **21**, 4-20.
- Berndt MC, Metharom P, Andrews RK (2014). Primary haemostasis: newer insights. *Haemophilia* **20 Suppl 4**, 15-22.
- Bertina RM, Veltkamp JJ (1978). The abnormal factor IX of hemophilia B+ variants. *Thromb Haemost* **40**, 335-49.
- Bloem E, Meems H, van den Biggelaar M, Mertens K, Meijer AB (2013). A3 domain region 1803-1818 contributes to the stability of activated factor VIII and includes a binding site for activated factor IX. *J Biol Chem* **288**, 26105-11.
- Bolton-Maggs PH, Pasi KJ (2003). Haemophilias A and B. *Lancet* **361**, 1801-9.
- Bowen DJ (2002). Haemophilia A and haemophilia B: molecular insights. *Mol Pathol* **55**, 127-44.

- Bristol JA, Ratcliffe JV, Roth DA, Jacobs MA, Furie BC, Furie B (1996). Biosynthesis of prothrombin: intracellular localization of the vitamin K-dependent carboxylase and the sites of gamma-carboxylation. *Blood* **88**, 2585-93.
- Broze GJ (1995). Tissue factor pathway inhibitor and the revised theory of coagulation. *Annu Rev Med* **46**, 103-12.
- Brunetti-Pierri N, Grove NC, Zuo Y, Edwards R, Palmer D, Cerullo V, Teruya J, Ng P (2009). Bioengineered factor IX molecules with increased catalytic activity improve the therapeutic index of gene therapy vectors for hemophilia B. *Hum Gene Ther* **20**, 479-85.
- Butenas S, Mann KG (2002). Blood coagulation. *Biochemistry (Mosc)* **67**, 3-12.
- Camerino G, Grzeschik KH, Jaye M, De La Salle H, Tolstoshev P, Lecocq JP, Heilig R, Mandel JL (1984) Regional localization on the human X chromosome and polymorphism of the coagulation factor IX gene (hemophilia B locus). *Proc Natl Acad Sci U S A* **81**, 498-502.
- Casaña P, Haya S, Cid AR, Oltra S, Martínez F, Cabrera N, Aznar JA (2009). Identification of deletion carriers in hemophilia B: quantitative real-time polymerase chain reaction or multiple ligation probe amplification. *Transl Res* **153**, 114-7.
- Chance PF, Dyer KA, Kurachi K, Yoshitake S, Ropers HH, Wieacker P, Gartler SM (1983). Regional localization of the human factor IX gene by molecular hybridization. *Hum Genet* **65**, 207-8.
- Chavali S, Mahajan A, Ghosh S, Mondal B, Bharadwaj D (2011). Protein molecular function influences mutation rates in human genetic diseases with allelic heterogeneity. *Biochem Biophys Res Commun* **412**, 716-22.
- Costa JM, Ernault P, Vidaud D, Vidaud M, Meyer D, Lavergne JM (2000). Fast and efficient mutation detection method using multiplex PCR and cycle sequencing--application to haemophilia B. *Thromb Haemost* **83**, 244-7.
- Davie EW, Ratnoff OD (1964). Waterfall sequence for intrinsic blood clotting. *Science* **145**, 1310-2.
- Di Scipio RG, Kurachi K, Davie EW (1978). Activation of human factor IX (Christmas factor). *J Clin Invest* **61**, 1528-38.
- DiMichele D, Neufeld EJ (1998). Hemophilia. A new approach to an old disease. *Hematol Oncol Clin North Am* **12**, 1315-44.

- Dunn AL (2010). Malignancy in patients with haemophilia: a review of the literature. *Haemophilia* **16**, 427-36.
- Federici AB (2003). The factor VIII/von Willebrand factor complex: basic and clinical issues. *Haematologica* **88**, EREP02.
- Figueiredo MS (1993). Molecular analysis of hemophilia B in Brazilian patients. *Braz J Med Biol Res* **26**, 919-31.
- Gailani D (2000). Activation of factor IX by factor XIa. *Trends Cardiovasc Med* **10**, 198-204.
- Gorziza RP, Vieira IA, Kappel DB, Rosset C, Sinigaglia M, Leiria LB, Salzano FM, Bandinelli E (2013). Genetic changes in severe haemophilia A: new contribution to the aetiology of a complex disease. *Blood Coagul Fibrinolysis* **24**, 164-9.
- Green PM, Bentley DR, Mibashan RS, Nilsson IM, Giannelli F (1989). Molecular pathology of haemophilia B. *EMBO J* **8**, 1067-72.
- Handford PA, Baron M, Mayhew M, Willis A, Beesley T, Brownlee GG, Campbell ID (1990). The first EGF-like domain from human factor IX contains a high-affinity calcium binding site. *EMBO J* **9**, 475-80.
- Harlos K, Holland SK, Boys CW, Burgess AI, Esnouf MP, Blake CC (1987). Vitamin K-dependent blood coagulation proteins form hetero-dimers. *Nature* **330**, 82-4.
- Hoffman, M (2003): Remodeling the blood coagulation cascade. *J Thromb Thrombolysis* **16**, 17-20.
- Hoffman M, Monroe DM (2001). A cell-based model of hemostasis. *Thromb Haemost* **85**, 958-65.
- Huntington JA (2005). Molecular recognition mechanisms of thrombin. *J Thromb Haemost* **3**, 1861-72.
- Lillicrap D (1998). The molecular basis of haemophilia B. *Haemophilia* **4**, 350-7.
- Macfarlane RG (1964). An enzyme cascade in the blood clotting mechanism, and its function as a biochemical amplifier. *Nature* **202**, 498-9.
- Mahajan A, Chavali S, Kabra M, Chowdhury MR, Bharadwaj D (2004). Molecular characterization of hemophilia B in North Indian families: identification of novel and recurrent molecular events in the factor IX gene. *Haematologica* **89**, 1498-503.

- Makino Y, Omichi K, Kuraya N, Ogawa H, Nishimura H, Iwanaga S, Hase S (2000). Structural analysis of N-linked sugar chains of human blood clotting factor IX. *J Biochem* **128**, 175-80.
- Marcus AJ, Safier LB (1993). Thromboregulation: multicellular modulation of platelet reactivity in hemostasis and thrombosis. *FASEB J* **7**, 516-22.
- McMichael, M (2012): New models of hemostasis. *Top Companion Anim Med* **27**, 40-5.
- Mertens K, Celie PH, Kolkman JA, Lenting PJ (1999). Factor VIII-factor IX interactions: molecular sites involved in enzyme-cofactor complex assembly. *Thromb Haemost* **82**, 209-17.
- Nakamura Y, Murata M, Takagi Y, Kozuka T, Nakata Y, Hasebe R, Takagi A, Kitazawa J, Shima M, Kojima T (2015). SVA retrotransposition in exon 6 of the coagulation factor IX gene causing severe hemophilia B. *Int J Hematol* **102**, 134-9.
- Nerich V, Tissot E, Faradji A, Demesmay K, Bertrand MA, Lorenzini JL, Briquel ME, Pouzol P, Woronoff-Lemsi MC (2008). Cost-of-illness study of severe haemophilia A and B in five French haemophilia treatment centres. *Pharm World Sci* **30**, 287-92.
- Orlova NA, Kovnir SV, Vorobiev II, Gabibov AG (2012). Coagulation Factor IX for Hemophilia B Therapy. *Acta Naturae* **4**, 62-73.
- Peyvandi F, Jayandharan G, Chandy M, Srivastava A, Nakaya SM, Johnson MJ, Thompson AR, Goodeve A, Garagiola I, Lavoretano S, Menegatti M, Palla R, Spreafico M, Tagliabue L, Asselta R, Duga S, Mannucci PM (2006). Genetic diagnosis of haemophilia and other inherited bleeding disorders. *Haemophilia* **12 Suppl 3**, 82-9.
- Radic CP, Rossetti LC, Abelleyro MM, Candela M, Pérez Bianco R, de Tezanos Pinto M, Larripa IB, Goodeve A, De Brasi C (2013). Assessment of the F9 genotype-specific FIX inhibitor risks and characterisation of 10 novel severe F9 defects in the first molecular series of Argentinian patients with haemophilia B. *Thromb Haemost* **109**, 24-33.
- Rallapalli PM, Kemball-Cook G, Tuddenham EG, Gomez K, Perkins SJ (2013). An interactive mutation database for human coagulation factor IX provides novel insights into the phenotypes and genetics of hemophilia B. *J Thromb Haemost* **11**, 1329-40.

- Rezende SM, Pinheiro K, Caram C, Genovez G, Barca D (2009).Registry of inherited coagulopathies in Brazil: first report. *Haemophilia* **15**, 142-9.
- Riddel JP, Aouizerat BE, Miaskowski C, Lillicrap DP (2007). Theories of blood coagulation. *J Pediatr Oncol Nurs* **24**, 123-31.
- Roberts HR, Monroe DM, Escobar MA (2004).Current concepts of hemostasis: implications for therapy. *Anesthesiology* **100**, 722-30.
- Rosset C, Gorziza RP, Botton MR, Salzano FM, Bandinelli E (2014). Factor VIII mutations and inhibitor formation in a southern Brazilian population. *Blood Coagul Fibrinolysis* **25**, 125-7.
- Rosset C, Vieira IA, Salzano FM, and Bandinelli E (2016).A germline variant affects putative miRNA-binding sites at the F8 3'UTR and acts as a potential haemophilia A phenotype modifier in Southern Brazilian patients. *Haemophilia* **22**, e327-9.
- Rosset C, Vieira IA, Sinigaglia M, Gorziza RP, Salzano FM, Bandinelli E (2013).Detection of new mutations and molecular pathology of mild and moderate haemophilia A patients from southern Brazil. *Haemophilia* **19**, 773-81.
- Ruggeri ZM, Mendolicchio GL (2007). Adhesion mechanisms in platelet function. *Circ Res* **100**, 1673-85.
- Salzano FM (1990). *Genética e Farmácia*. Editora Manole LTDA.
- Schenone M, Furie BC, Furie B (2004). The blood coagulation cascade. *Curr Opin Hematol* **11**, 272-7.
- Schmidt AE, Bajaj SP (2003). Structure-function relationships in factor IX and factor IXa. *Trends Cardiovasc Med* **13**, 39-45.
- Stassen JM, Arnout J, Deckmyn H (2004).The hemostatic system. *Curr Med Chem* **11**, 2245-60.
- Stenflo J, Lundwall A, Dahlbäck B (1987). beta-Hydroxyasparagine in domains homologous to the epidermal growth factor precursor in vitamin K-dependent protein S. *Proc Natl Acad Sci U S A* **84**, 368-72.
- Venkateswarlu D (2014). Structural insights into the interaction of blood coagulation co-factor VIIIa with factor IXa: a computational protein-protein docking and molecular dynamics refinement study. *Biochem Biophys Res Commun* **452**, 408-14.

- Versteeg HH, Heemskerk JW, Levi M, Reitsma PH (2013). New fundamentals in hemostasis. *Physiol Rev* **93**, 327-58.
- White GC, Rosendaal F, Aledort LM, Lusher JM, Rothschild C, Ingerslev J, Subcommittee FVaFI (2001). Definitions in hemophilia. Recommendation of the scientific subcommittee on factor VIII and factor IX of the scientific and standardization committee of the International Society on Thrombosis and Haemostasis. *Thromb Haemost* **85**, 560.
- Wilcox JN, Smith KM, Schwartz SM, Gordon D (1989). Localization of tissue factor in the normal vessel wall and in the atherosclerotic plaque. *Proc Natl Acad Sci U S A* **86**, 2839-43.
- World Federation of Hemophilia (2016). Report on the Annual Global Survey 2015. Disponível em: <http://www1.wfh.org/publication/files/pdf-1669.pdf>. Acesso em 20 de janeiro de 2017
- Yoshitake S, Schach BG, Foster DC, Davie EW, Kurachi K (1985). Nucleotide sequence of the gene for human factor IX (antihemophilic factor B). *Biochemistry* **24**, 3736-50.
- Zajicek j (1954). Studies on the histogenesis of blood platelets. *Acta Haematol* **12**, 238-44.
- Zögg, T, and Brandstetter, H (2009): Activation mechanisms of coagulation factor IX. *Biol Chem* **390**, 391-400.

ANEXOS

Anexo 1: Termo de Consentimento Livre e Esclarecido

Anexo 2: Parecer Consubstanciado do Comitê de Ética em Pesquisa da Pró-Reitoria de Pesquisa da UFRGS

Anexo 1. Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.

Projeto de pesquisa: **“NOVOS ESTUDOS DE CARACTERIZAÇÃO MOLECULAR EM PACIENTES COM HEMOFILIA B DO RIO GRANDE DO SUL”**

Pesquisador Responsável: Francisco Mauro Salzano

Pesquisadores Participantes: Mariana Rost Meireles, Eliane Bandinelli, Francisco Mauro Salzano, Gustavo Fioravanti Vieira.

Hemocentro de Porto Alegre

Laboratório de Hemostasia – Departamento de Genética - Universidade Federal do Rio Grande do Sul

Telefone para contato: 3308-6728

O departamento de genética da UFRGS em colaboração com o hemocentro de Porto Alegre está promovendo o projeto de pesquisa **“NOVOS ESTUDOS DE CARACTERIZAÇÃO MOLECULAR EM PACIENTES COM HEMOFILIA B DO RIO GRANDE DO SUL”**. O objetivo deste projeto é a caracterização molecular das mutações em pacientes com Hemofilia B do estado do Rio Grande do Sul e análise do efeito dessas mutações na determinação da doença. A Hemofilia B (HB) está entre os três distúrbios hemorrágicos mais frequentes, atingindo cerca de 1 a cada 25.000-30.000 nascimentos do sexo masculino. A HB é resultante de alterações genéticas no gene F9 que levam a uma deficiência qualitativa ou quantitativa na proteína fator IX (FIX) da coagulação sanguínea.

O quadro clínico dos pacientes com hemofilia B pode ser associado ao nível de atividade do fator IX. Quanto mais graves, maior a propensão a sangramentos espontâneos, que geralmente não ocorrem em pacientes leves, os quais costumam ter apenas sangramentos pós-traumáticos. Cerca de 60-70% dos indivíduos portadores de hemofilia B desenvolvem formas graves ou moderadas da coagulopatia, embora essas proporções não estejam ainda bem estabelecidas. O entendimento das diferentes mutações que ocorrem no

DNA de pacientes com HB pode ajudar a esclarecer os mecanismos da patogênese desta coagulopatia.

A participação no estudo envolve o preenchimento de uma ficha clínica e a doação de 5ml de sangue, que será utilizado para a realização dos testes de coagulação sanguínea e extração de DNA.

A participação neste estudo é sigilosa (bem como os resultados da avaliação clínica e dos exames genéticos) e voluntária. É permitida a desistência em qualquer etapa da avaliação, sem qualquer tipo de prejuízo ao participante. Caso você opte pela não participação saiba que não haverá alterações no atendimento e no tratamento normal a que você tem direito.

O presente projeto foi avaliado e aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal do Rio Grande do Sul. O estudo será financiado por verba de agências de fomento, sendo que nenhum custo incidirá sobre o paciente ou seus familiares.

O material genético excedente poderá ser armazenado ou não, conforme a decisão de cada paciente. O material que ficar armazenado poderá ser utilizado em novos exames ou para uso em novas pesquisas. No caso de serem propostas novas pesquisas com este material, elas serão avaliadas pelo Comitê de Ética.

Em caso de dúvidas, entre em contato com o Comitê de Ética da UFRGS (fone: (051)3308-3629) ou com os pesquisadores Eliane Bandinelli (fone: (051)3308-6736) ou Francisco M. Salzano (fone: (051)3308-6747)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO - PROJETO:

“NOVOS ESTUDOS DE CARACTERIZAÇÃO MOLECULAR EM PACIENTES COM HEMOFILIA B DO RIO GRANDE DO SUL”

Eu, _____, declaro que fui convidado a participar do projeto de pesquisa “NOVOS ESTUDOS DE CARACTERIZAÇÃO MOLECULAR EM PACIENTES COM HEMOFILIA B DO RIO GRANDE DO SUL” acima citado. Fui informado de que minha decisão em participar não comprometerá meu tratamento no hemocentro, sendo meus dados e resultados de meus testes absolutamente sigilosos. Além disso, fui informado de que a qualquer momento posso solicitar minha desistência do estudo, sem qualquer prejuízo. Declaro que consinto

em participar do estudo e que meus dados sejam incluídos na análise coletiva dos resultados.

Com relação ao armazenamento de material genético eu declaro que:

()SIM: autorizo manter meu material genético excedente (DNA) armazenado, sabendo que poderá ser usado, no futuro, para novas pesquisas, as quais serão aprovadas pelo Comitê de Ética.

()NÃO: não autorizo armazenar meu material genético após este exame.

Eu, _____ (ou meu responsável legal), abaixo assinado estou ciente dos termos acima descritos e concordo com a minha participação no estudo.

.....

Assinatura do paciente ou responsável legal

Eu expliquei/expliquei a _____ os objetivos e procedimentos necessários para esta pesquisa, na minha melhor capacidade.

Nome do entrevistador: _____

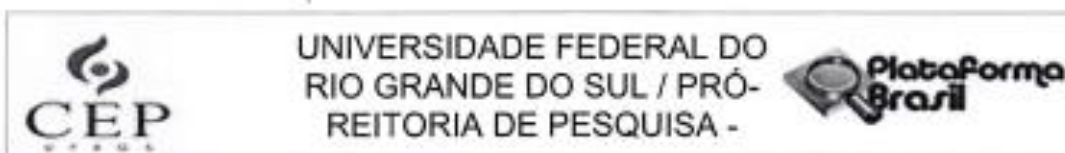
Assinatura do entrevistador: _____

Nome da testemunha: _____

Assinatura da testemunha: _____

Porto Alegre, _____ de _____ de 2015.

Anexo 2. Parecer consubstanciado ao CEP



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Problemas de hemostasia no Rio Grande do Sul: Prevalência e caracterização molecular da Hemofilia B e da Doença de Von Willebrand

Pesquisador: FRANCISCO MAURO SALZANO

Área Temática: Genética Humana;

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP.);

Versão: 4

CAAE: 45164015.3.0000.5347

Instituição Proponente: Universidade Federal do Rio Grande do Sul

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.359.825

Apresentação do Projeto:

Os pesquisadores atenderam as solicitações. Assim, recomenda-se aprovação do projeto de pesquisa.

Sugere-se aos pesquisadores substituir no TCLE o número da Resolução, de 196/12 para 466/12, e a palavra "cópia" por "via".

Objetivo da Pesquisa:

A pesquisa tem como objetivo principal dar continuidade às pesquisas iniciadas há mais de meio século sobre os distúrbios de hemostasia que ocorrem em nosso meio; e mais especificamente seria uma continuação do Projeto de Pesquisa 15667 – "Coagulopatias hereditárias frequentes: melhorias no diagnóstico da doença de von Willebrand e suscetibilidade ao desenvolvimento de inibidores em hemofilia A grave", registrado no sistema de pesquisa da UFRGS e cujo prazo se encerrou em 30/12/2014.

O Projeto de pesquisa atual está dividido em dois Subprojetos:

1. Subprojeto 1: Caracterização molecular das mutações em pacientes com Hemofilia B; e
2. Subprojeto 2: Caracterização molecular das mutações em pacientes com a Doença de von

Endereço: Av. Paulo Gama, 110 - Sala 317 do Prédio Anexo 1 da Reitoria - Campus Centro
Bairro: Farroupilha **CEP:** 91.040-000
UF: RS **Município:** PORTO ALEGRE
Telefone: (51)3308-3738 **Fax:** (51)3308-4085 **E-mail:** etica@propeq.ufrgs.br



Continuação do Parecer: 1.359.823

Willebrand Tipo 3.

Em ambos os casos procurar-se-á: 1. Correlacionar os tipos de alterações encontradas com a prevalência de mutações já descritas em bancos de dados relacionados a essas duas condições; 2. Avaliar as relações genótipo-fenótipo levando em conta os quadros clínicos dos pacientes, bem como análises de bioinformática para a interpretação da ação gênica em uma abordagem molecular; e 3. Contribuir para um melhor conhecimento dos pacientes que vivem em nosso meio, visando estratégias de manejo e aconselhamento genético diferenciados.

Avaliação dos Riscos e Benefícios:

O texto dos riscos e dos benefícios da pesquisa está adequado.

Comentários e Considerações sobre a Pesquisa:

Os pesquisadores atenderam a solicitação de adequação dos nomes dos pesquisadores da equipe de pesquisa. Essa informação é a mesma em toda a documentação do projeto de pesquisa.

Em atendimento à solicitação "Informar como foi obtido o número 60 para a amostra de pacientes com Doença de Von Willebrand, pois foi informado no projeto de pesquisa ser 30 o número de pacientes atualmente diagnosticados com a doença", os pesquisadores informaram: "Atualmente, existem 60 pacientes diagnosticados como apresentando a Doença de von Willebrand, e eles serão abordados e convidados a participar da pesquisa". Assim, foi esclarecido o número amostral 60 para pacientes com doença de von Willebrand.

Ainda, os pesquisadores informaram: "A data do primeiro recrutamento foi alterada como solicitado. A nova data será 04 de janeiro de 2016".

Considerações sobre os Termos de apresentação obrigatória:

Todos os documentos obrigatórios foram anexados. Sugere-se aos pesquisadores substituir no TCLE o número da Resolução, de 196/12 para 466/12, e a palavra "cópia" por "via".

Recomendações:

Recomenda-se aprovação do projeto de pesquisa quanto aos aspectos éticos.

Conclusões ou Pendências e Lista de Inadequações:

Os pesquisadores atenderam às solicitações. A versão atual do projeto de pesquisa está adequada quanto aos aspectos éticos.

Sugere-se aos pesquisadores substituir no TCLE o número da Resolução, de 196/12 para 466/12, e a palavra "cópia" por "via".

Endereço: Av. Paulo Gama, 110 - Sala 317 do Prédio Anexo 1 da Reitoria - Campus Centro
Bairro: Fátima CEP: 90.040-060
UF: RS Município: PORTO ALEGRE
Telefone: (51)3308-3738 Fax: (51)3308-4055 E-mail: etica@propesq.ufrgs.br



UNIVERSIDADE FEDERAL DO
RIO GRANDE DO SUL / PRÓ-
REITORIA DE PESQUISA -



Continuação do Parecer: 1.369.825

Considerações Finais a critério do CEP:

Aprovado.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_513415.pdf	03/12/2015 10:02:56		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	02/09/2015 11:12:55	FRANCISCO MAURO SALZANO	Aceito
Projeto Detalhado / Brochura Investigador	Problemas de Hemostasia, RS.pdf	13/07/2015 12:57:50		Aceito
Folha de Rosto	FOLHA DE ROSTO (Assinada).pdf	14/05/2015 11:18:23		Aceito
Outros	Carta Aceite (HEMOCENTRO).pdf	13/05/2015 13:06:29		Aceito
Outros	ParecerCOMPESQ.pdf	13/05/2015 13:05:23		Aceito

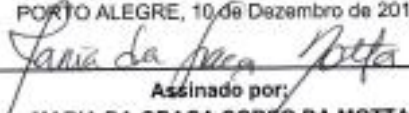
Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

PORTO ALEGRE, 10 de Dezembro de 2015


Assinado por:
MARIA DA GRAÇA CORSO DA MOTTA
(Coordenador)

Endereço: Av. Paulo Gama, 110 - Sala 317 do Prédio Anexo 1 da Reitoria - Campus Centro
Bairro: Fátima CEP: 91.040-000
UF: RS Município: PORTO ALEGRE
Telefone: (51)3308-3738 Fax: (51)3308-4085 E-mail: wica@propesq.ufrgs.br

Página 03 de 03