Thrombosis Research 135 (2015) 1057–1063

Contents lists available at ScienceDirect

Thrombosis Research

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

Regular Article

Overrepresentation of missense mutations in mild hemophilia A patients from Belgium: founder effect or independent occurrence?

N. Lannoy [a,b,*], C. Lambert [c], M. Vikkula [a,d], C. Hermans [b,c]

[a] Center of Human Genetics, Université Catholique de Louvain, Cliniques Universitaires Saint-Luc, Brussels, Belgium
[b] Institut de Recherche Expérimentale et Clinique (IREC), Université catholique de Louvain, Brussels, Belgium
[c] Hemostasis and Thrombosis Unit, Hemophilia Clinic, Division of Hematology, Cliniques Universitaires Saint-Luc, Brussels, Belgium
[d] Laboratory of Human Molecular Genetics of the Duve Institute, Université catholique de Louvain, Brussels, Belgium

ARTICLE INFO

Article history:
Received 15 December 2014
Received in revised form 26 February 2015
Accepted 8 March 2015
Available online 17 March 2015

Keywords:
hemophilia A
Factor VIII
Genetics of Thrombosis and Haemostasis
mutation analysis
microsatellites

ABSTRACT

Roughly 40% of observed mutations responsible for hemophilia A (HA) are novel and present in either a single family or a limited number of unrelated families. During routine diagnostic analysis of 73 unrelated Belgian patients with mild HA, 4 out of 43 different mutations (p.Ser2030Asn, p.Arg2178Cys, p.Arg2178His, and p.Pro2311His) were detected in more than one family, representing 35% of total identified mutations.
To discriminate between an independent recurrence or a founder effect, an analysis of intra- and -extragenic single nucleotide polymorphisms (SNPs) and short tandem repeats (STRs) flanking the *F8* gene was conducted. SNP haplotype and microsatellite analysis revealed strong evidence that p.Ser2030Asn and p.Pro2311His mutations were probably associated with a founder effect. The two other mutations localized in an *F8* cytosine-phosphate-guanine (CpG) site likely resulted from recurrent *de novo* events.
This study suggests that missense mutations producing C-to-T or G-to-A substitutions in CpG dinucleotide can occur *de novo* with more repetition than other causal substitutions that do not affect the CpG site. Analysis of F8 database implied that CpG sites throughout the *F8* gene are not all mutated with the same frequency. Causes are still unknown and remain to be identified.

Introduction

Hemophilia A (HA) (OMIM 306700), an X-linked hereditary disease, is one of the most common coagulation disorders, caused by a spectrum of mutations in the *F8* gene. With the exception of intron 1 and 22 inversions, present in approximately 45% of severe hemophilia A cases [2,3], more than 2015 unique variants spread throughout the *F8* gene's 26 exons have been identified and recorded in the international hemophilia A mutation database [http://www.eahad-db.org/February 2014 update].

Due to the high rate of spontaneous *F8* gene mutations, approximately 40% of families, with phenotypes of variable severity, have affected individuals being either sporadic or confined to one family branch [11]. However, several publications have described *F8* mutations that are highly prevalent in specific populations sharing the same haplotype, indicating the existence of a single, common ancestor. Examples of this can be seen with the c.6104 T > C or p.Val2035Alanine (legacy AA No.2016) mutation, identified in an isolated population in rural Newfoundland [14], the exon 13 duplication located in northern Italy [1], and the novel G-to-A mutation (c.1538-18G > A) in intron 10 of the *F8* gene, which was identified as a putative cause of mild hemophilia A in southern Italy [13]. All these mutations were associated with the mild HA phenotype.

Within the setting of a study comprising 73 apparently unrelated French-speaking Belgian HA families with mild severity screened for *F8* gene mutations, 4 different missense mutations were detected as occurring between 4 and 10 times. By way of haplotype analysis constructed with single nucleotide polymorphisms (SNPs) and microsatellites repeats, this study sought to assess the origins of these mutations. We investigated whether they represent a recurrent molecular defect, occurring repeatedly *de novo,* or whether they may have manifested as a single molecular event that was then disseminated throughout French-speaking Belgium via a founder effect. This study and analysis of the EAHAD database (EAHAD db) will attempt to demonstrate that certain *F8* CpG sites ("CpG" standing for "—C—phosphate—G—) are more susceptible to mutations than others, and that mutations frequently reported outside of a CpG site are possibly the result of a founder effect.

Patients and Methods

Patients

This study involved 73 apparently unrelated subjects exhibiting mild HA with Factor VIII clotting activity from >5 to <40 IU/dL (>5% to <40%

* Corresponding author at: Centre de Génétique Humaine Tour Rosalind Franklin Avenue Mounier entrée F 1200 Brussels Belgium. Tel.: +32 2 764 6723; fax: +32 2 764 6939.
E-mail address: nlannoy@dmsnet.be (N. Lannoy).

http://dx.doi.org/10.1016/j.thromres.2015.03.017
0049-3848/

of normal levels) recruited for genetic testing from hospitals in Brussels and the French-speaking part of Belgium. The study was approved by the medical ethics committee of the medical faculty of the Université catholique de Louvain, Brussels (Belgium). Blood samples were obtained during medical care where patients were informed and gave their consent for DNA analysis.

Methods

DNA extraction and genotyping of coding sequences

Genomic DNA was extracted from white blood cells using standard procedures, and mutation screening was performed as previously described [9]. Causal mutations were defined in line with recommendations of the Human Genome Variation Society (HGVS; http://www.HGVS.org; June 2010 update) with codon 1 representing the first residue (Met) of the signal peptide (this is -19 in Legacy numbering). In Legacy numbering, codon +1 refers to that coding for the first amino-acid of the mature FVIII protein (in HGVS numbering, this is codon +20). As Legacy numbering is extensively used in FVIII publications, particularly before the year 2000, this numbering is maintained within parenthesis.

In addition, 50 DNA samples from anonymous French-speaking Belgian males, referred for pre-conception genetic counselling, were selected from the DNA bank and genotyped as below. This was in order to facilitate the interpretation of results obtained by haplotype analysis.

SNP and microsatellite genotyping

Flanking markers from 1.4 Mb centromeric to 0.6 Mb telomeric of the *F8* locus were genotyped, including the four previously reported [6,8,12,16] extragenic and intronic microsatellites (3' DXS15–DXSint22–DXSint13–DXS1108-5') and ten SNPs, as displayed in Table 1. The PCR was conducted in a 25 μL containing 1x PCR Faststart buffer, 1U of Faststart Taq polymerase (Hoffmann-La Roche, Basel, Switzerland), 0.2 mM of each dNTP, and 10pmol of each primer. Cycling conditions were set to initial denaturation at 94 °C for 5 minutes, followed by 30 cycles of amplification. These comprised 30 second denaturation at 94 °C; 1 minute annealing at 55 °C; 1 minute elongation at 72 °C; and 5 minutes final extension at 72 °C. A 1.5 μL aliquot of STR PCR product was added to 10 μL of Hi-Di Formamide (Applied Biosystems, Foster city, CA, USA). Amplification products were identified using an ABI3130xl capillary sequencer by employing the Genescan-ROX-500 (Applied Biosystem, Foster City, CA, USA) size standard, and were interpreted with GeneMarker software V1.5 (SoftGenetics LLC, State College, PA, USA). Amplified SNP DNA fragments were purified and subjected to direct cycle sequence analysis using the Taq di-deoxy terminator method and an ABI PRISM 3130xl Genetic Analyzer sequencer, in accordance with the manufacturer's instructions.

Results

In this cohort, routine genetic analysis of the *F8* gene was undertaken, and 43 different missense mutations were identified in a group of 73 unrelated subjects from separate families with mild HA [Table S1]. Recurrent missense causal mutations were noted 10 times for the p.Ser2030Asn (c.6089G > A) in exon 19, 7 times for the p.Arg2178Cys (c.6532C > T) in exon 23, 4 times for the p.Arg2178His (c.6533G > A) in exon 23, and 4 times for the p.Pro2311His (c.6932C > A) in exon 26. All mutations were associated with the mild phenotype. These 4 mutations accounted for 35% of all mutations identified in the routine genetic *F8* analysis of mild phenotype patients. Other recurrent missense mutations were also observed (p.Leu431Phe and p.Gly1769Arg). However, these families were not included in this study due to a lack of DNA.

In the first part of this study, haplotypes at the *F8* locus were generated with extra- and intragenic SNPs from Belgian male control individuals among whom 6 different haplotypes were identified, as displayed in Table 2a. The most common haplotype (haplotype 4) manifested at 58% prevalence, while 3 others (haplotypes 2, 5, and 6) were found at ≤10% prevalence in the normal population. It was of particular interest to note that 2 haplotypes in the Belgian control population were not present in the general CEU population, which consists of Utah residents with ancestry from northern and western Europe. In contrast, one haplotype in the general CEU population was not present in our Belgian male controls (Table 2b). In order to increase the discriminative value of the Belgian SNP haplotype, we conducted further analyses of the microsatellite data at the *F8* locus. A total of 35 different haplotypes were identified in the 50 controls (Table 2c).

For the p.Ser2030Asn mutation in *F8* exon 19, all 10 unrelated patients exhibited the same SNP haplotype, corresponding to the most commonly found (haplotype 4) in the Belgian control population (Tables 2 and 3). Analysis of short tandem repeat (STR) microsatellites identified the same microsatellite haplotype in eight patients, while the two other haplotypes differed in only one (family HA55 for DXS15 marker) or two alleles (family HA39 for DXS15 and STR13 markers). The HA39 STR haplotype was absent in the Belgian male controls while the HA55 microsatellites allele was present in 2% of the same control population.

For the p.Pro2311His causal mutation in *F8* exon 26, haplotyping by way of SNP markers revealed that the 4 unrelated index cases shared the same haplotype (Table 3). Yet while this SNP haplotype was absent in the Belgian control population, it was present in 4.5% of cases in the CEU population, as reported in the international HapMap project (Haplotype Map of the human genome: http://www.ncbi.nlm.nih.gov/projects/SNP/snp_gf.cgi, consulted on 12 December 2009). These results were confirmed by STR haplotype, except for the 139 bp/STR allele in family HA56.

For the two other exon 23 mutations at the same codon, namely p.Arg2178Cys and p.Arg2178His, genotyping of unrelated patients by means of SNP and STR markers clearly revealed 3 different haplotypes

Table 1
Intronic and extragenic SNPs markers.

SNP	Physical position at Xq28	RefSNP Alleles	Minor allele	MAF*	Forward sequence	Reverse sequence
rs6643620	~169 kb 3' of the *F8* gene	A/G	G	0.452/747	5'-TTTCTGGTTGCTAGAATTATGGA3'	5'-AAAGAGTGGGGCATTTGCTA-3'
rs5945250	~162 kb 3' of the *F8* gene	G/T	T	0.349/577	5'-TCTCACCATCACGAACATAAGG-3'	5'-TGTTGCACAACAATGTGAATG-3'
rs1050705	3'-UTR variant	A/G	C	0.374/619	5'-CTGAAGAAACCAGCAGGAAAA-3'	5'-GCAAATGGTTTATAGCCCTGT-3'
rs4898352	intron 18 of the *F8* gene	A/T	T	0.3761/621	5'-TGTGACATTAATTATCCATTTATTCTG-3'	5'-TGGCACTGTACAATCTCTATCCA-3'
rs6643714	intron 14 of the *F8* gene	C/T	T	0.4907/811	5'-TGAGAAAATCTTTATCACCCCTTC-3'	5'-TTCTCTGATGGCCAGTGATG-3'
rs1800291	exon 14 of the *F8* gene	C/G	C	0.2461/406	5'-TCGTCAACAGAGAGTGGTAGG-3'	5'-TTTCTCAGACAAGAAATTCTGACC-3'
rs5987079	intron 2 of the *F8* gene	A/G	A	0.2866/474	5'-GAACCCACTAATGATCAAATTCC-3'	5'-CATTTGCTGTTTTGGTTTTTGT-3'
rs6649625	intron 2 of the *F8* gene	C/T	A	0.2424/201	5'-GCACTCATAAATGGAAATAGGAGAA-3'	5'-ATCTGGTGGGTGAAAGCAAT-3'
rs2313058	~13 kb 3' of the *F8* gene	C/T	A	0.2956/489	5'- CAACAGTTCATTCCTTTTTATTGC-3'	5'-GCACTTTAAAAACTGGCAGGTA-3'
rs5945279	~22 kb 3' of the *F8* gene	C/T	T	0.2183/361	5'-TAGAGCAGGGGTCAGCAAAG-3'	5'-TCTGCTGGGTCTATACCATGA-3'

MAF*: Minor allele frequency (MAF) is the ratio of chromosomes in the population carrying the less common variant to those with the more common variant. This means that for rs6643620, the minor allele is 'G' and has a frequency of 45.2% in the 1000 Genomes Phase 1 population, and that 'G' is observed 747 times in the sample population of 629 people (or 1,258 chromosomes) [http://www.ncbi.nlm.nih.gov/snp/ and http://www.ncbi.nlm.nih.gov/projects/SNP/docs/rs_attributes.html#gmaf]