Population-specific spectrum of factor XI mutations in Koreans: evidence for a founder effect

Juwon Kim

Department of Medicine

The Graduate School, Yonsei University

Population-specific spectrum of factor XI mutations in Koreans: evidence for a founder effect

Directed by Professor Kyung-A Lee

The Doctoral Dissertation
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Doctor of Philosophy

Juwon Kim

June 2012

This certifies that the Doctoral Dissertation of Juwon Kim is approved.

Thesis Supervisor : Kyung-A Lee

Thesis Committee Member#1 : Jong Rak Choi

Thesis Committee Member#2 : Chuhl Joo Lyu

Thesis Committee Member#3: Yu Ri Kim

Thesis Committee Member#4: Myung-Hyun Nam

The Graduate School
Yonsei University

June 2012

ACKNOWLEDGEMENT

Writing a thesis is like running a marathon which requires a great deal of patience, diligence and carefully made plans. Tremendous help from great number of people has placed me at the finish line of my long journey. It is a pleasure to convey my gratitude to them all in my humble acknowledgement. First of all, I would like to express the deepest appreciation to professor Kyung-A Lee, who has the attitude and the substance of genius. She continually and convincingly conveyed a spirit of adventure in regard to research and excitement in regard to teaching. Without her guidance and persistent help this dissertation would not have been possible. Her truly scientist intuition has made her as a constant oasis of ideas and passions in science, which inspired and enriched my growth as a student, a researcher and a medical scientist. I am grateful for the lab member, Myungsook Chang, for technical assistance and maintenance of the tools used in this study. Also, I would like to thank professor Jong Rak Choi for providing me with a good environment and facilities to complete this study, and professors Yu Ri Kim, Chul Joo Lyu, and Myung-Hyun Nam for their crucial and valuable advice and using their precious time to read this thesis and giving out critical comments. An honorable mention goes to our families and friends for their understanding and supports on me in completing the thesis. I am forever indebted to my parents and my husband, Kyuhyung Kim, for their understanding, endless patience and encouragement when it is most required. Words fail me to express my appreciation and love for my sons, Minjae and Minjune. Finally, I would like to thank everyone who was important to the successful realization of thesis, as well as expressing my apology that I could not mention personally one by one.

Juwon Kim

April 23rd, 2012

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Population-specific spectrum of factor XI mutations in Koreans: evidence for a founder effect

Juwon Kim

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Kyung-A Lee)

FXI deficiency (MIM#264900), also called hemophilia C, is a rare autosomal recessive bleeding disorder which is particularly common among Ashkenazi Jews. According to a *F11* mutation database, more than 191 disease causing mutations in *F11* have been reported. While FXI deficient patients in the Ashkenazi Jewish population commonly present the E117X and F283L mutations, non-Jewish patients with the same deficiency are found with different mutations. The aim of this study was to investigate a mutation spectrum of *F11* among Korean patients with FXI deficiency and to determine the haplotypes of mutations frequently found in Koreans. Thirteen unrelated patients from non-consanguineous families with FXI deficiency were included in the study. In the mutation analysis, the most frequently found mutations were Q263X (4 cases; 31%) and Q226X (3 cases; 23%). The five new missense mutations were: Q14R, R234T, H35Q, W383R, and D508E. The frequency of Q263X bearing haplotype was significantly different between normal and patient groups (P=0.001). These two mutations were on the haplotype

background distinct from those of mutations reported in Jewish or Italian populations. The mutation ages for Q226X and Q263X for mean population growth rate of 1.3% were 10.7 and 10.8 generations, respectively. The high frequency of the mutations Q226X and Q263X provides evidence for the founder effect. Testing for the presence of these two mutations should be the first genetic screening in Korean patients with FXI deficiency.

Key words: Factor XI deficiency, Factor XI gene, mutation spectrum

population-specific spectrum of factor XI mutations in Koreans: evidence for a founder effect

Juwon Kim

Department of Medicine
The Graduate School, Yonsei University

(Directed by Professor Kyung-A Lee)

I. INTRODUCTION

Factor XI (FXI) is a zymogen that contributes to hemostasis by activating factor IX[1]. A crystal structure of F11 revealed that it is composed of four apple domains (A1, A2, A3, and A4) which form a saucer-shaped platform for the catalytic domain with the apple 3 domain being at the back of the molecule in the right monomer unit[2, 3]. Apple domains are about 90 amino acids long, and the catalytic domain is 238 amino acids long[4]. A1 contains binding sites for the prothrombin, thrombin, high molecular kininogen, A2 and A3 domains provide binding sites for FIX, platelet, and heparin while A4 is important for dimerization and FXII binding. F11 circulates in the blood as a homodimer, which is unique among coagulation proteases, and it is of note that some of the mutations found in F11 deficient patients perturb dimerization[5]. FXI deficiency (MIM#264900), also called hemophilia C, is a rare autosomal recessive

bleeding disorder which is particularly common among Ashkenazi Jews with a heterozygote frequency of 9% and homozygote frequency of 0.22%[6]. The common modes of presentation of FXI deficiency are bleeding following injury, including dental procedures, tonsillectomy, nose surgery, urologic procedures, and an incidental finding of prolonged activated partial thromboplastin time (APTT) during routine physical checkup or pre-operation laboratory work up[7]. In patients with FXI deficiency, mutations were found in *F11* (GenBank accession number, NM_000128.3) which is located on chromosome 4 (4q35) and is 23 kb in length with 15 exons and 14 introns[8]. Exon1 encodes for the promoter region, exon 2 encodes for the 18 amino acid long-signal peptide that is cleaved during FXI biosynthesis, and exon 3 to 15 encode for the 607 amino acid long mature protein[9]. According to a *F11* mutation database at http://www.factorxi.org, more than 191 disease causing mutations in *F11*[10] have been reported. FXI deficiency has been associated with a wide range of mutations including missense, nonsense, splice site, insertion, and deletion mutations. Missense mutations account for 53% of the mutations. Most patients with severe FXI deficiency (FXI level < 20%) are found to harbor a homozygote mutation or compound heterozygote, while heterozygote individuals show milder presentation or show no symptoms[7].

Three mutations were initially identified in six Ashkenazi Jewish patients and were designated type I, type II, and type III[11]. Type I mutation is a G to A substitution at the donor splice site in the last intron of the F11 gene; type II is a

nonsense mutation creating a premature stop codon (E117X) in exon 5 and type III is missense mutation in exon 9 (F283L) which codes for protein that fails to be properly secreted from the cells[12]. Type II and III are the most frequent among all the mutation types, which account for about 95% of the cases[4]. Notably, while FXI deficient patients in the Ashkenazi Jewish population commonly present the E117X and F283L mutations, non-Jewish patients with the same deficiency are found with different mutations. In Basques, C38R was found to be the predominant mutation[13], while C128X and Q88X were more frequently observed in Caucasians from the United Kingdom and families living in west France, respectively[14, 15]. Haplotype analysis of these findings showed a founder effect. Interestingly, while the E117X and F283L mutations are frequently observed in the Jewish population and some European populations, they have never been found in Asians, suggesting possible existence of founder mutation. Although specific population studies have not been carried out in Asians, two mutations (Q226X and Q263X) have been frequently reported in Japanese and Chinese patients[16-20].

In this study, we studied 13 unrelated Korean patients diagnosed with FXI deficiency to analyze the spectrum of *F11* mutations. In addition, haplotype analysis was carried out to determine a common genetic background of a specific mutation, and to investigate a possible founder effect in the Korean population.

II. MATERIALS AND METHODS

1. Patients

Patients were selected based on the FXI level (< 60% in normal plasma) and prolonged APTT. The clinical histories and laboratory results of the patients are summarized in Table 1. We selected a total of 13 patients and 50 healthy controls among individuals who visited our hospital for physical checkup. The controls were screened for the prevalence of the two frequently found nonsense mutations and all novel mutations found in this study. Before the study began, informed written consent was collected from all participants according to the ethical guidance of the Institutional Review Board and carried out according to the principles of the Declaration of Helsinki. For patients who were under age of 14, informed consent was obtained from their parents.

2. Laboratory assessment

Citrated blood was used for coagulation factor assays which examined the levels of factors II through XII, prothrombin time (PT), and APTT which were measured using an automated coagulation analyzer (STA-R Evolution®, Diagnostica Stago, Asnieres, France). Other coagulation parameters, including lupus anticoagulant and fibrinogen, were also measured to rule out a possible secondary FXI deficiency.

Table 1. Phenotypic and genotypic results in patients with FXI deficiency

Case	Age (y)	Sex	Reason for referral	FXI:C (%)	FXI:Ag (ng/ml)[c]	Zygosity	Mutation Information				
							Codon (HGVS nomenclature)[a]	Amino acid change[b]	Exon	Geographic origin	Reference
1	74	M	Bleeding during urologic procedure	4	-	Homo	c.841C>T	p.Gln263X (Q263X)	8	Japanese, Chinese	16, 17, 19, 20
2	51	F	History of postoperative bleeding	2	-	Homo	c.841C>T	p.Gln263X (Q263X)	8	Japanese, Chinese	16, 17, 19, 20
3	46	M	Pre-operation screening	5	-	Homo	c.841C>T	p.Gln263X (Q263X)	8	Japanese, Chinese	16, 17, 19, 20
4	7	F	Pre-operation screening	53	7.1	Hetero	c.94G>A	p.Gly14Arg (G14R)	3	Novel	Present study
5	25	M	Pre-operation screening	3	-	Comp Hetero	c.730C>T c.755G>C	p.Gln226X (Q226X) p.Arg234Thr (R234T)	7 7	Japanese Novel	18 Present study
6	27	M	Health check up	48	-	Hetero	c.688T>C	p.Cys212Arg (C212R)	7	NA	21
7	33	F	Pre-operation screening	38	2.6	Hetero	c.159C>A	p.His35Gln (H35Q)	3	Novel	Present study
8	11	M	Pre-operation screening	9	-	Comp Hetero	c.1546G>A c.1560dupG	p.Val498Met (V498M) p.Tyr503ValfsX32	13 13	Korean Korean	22 22
9	7	F	Pre-operation screening Epistaxis	39	-	Hetero	c.1201T>C	p.Trp383Arg (W383R)	11	Novel	Present study
10	7	M	Pre-operation screening	12	-	Comp Hetero	c.730C>T c.1578C>G	p.Gln226X (Q226X) p.Asp508Glu (D508E)	7 14	Japanese Novel	18 Present study
11	68	F	Pre-operation screening	2	-	Comp Hetero	c.730C>T c.1556G>A	p.Gln226X (Q226X) p.Trp501X (W501X)	7 13	Japanese Japanese	18 23
12	22	M	Difficulty in bleeding control during nose surgery	20	-	Hetero	c.1253G>T	p.Gly400Val (G400V)	11	Italian(Czech), Chinese, Japanese	5, 18
13	22	F	Difficulty in bleeding control during plastic surgery	<1	-	Comp Hetero	c.841C>T c.1135+1G>A	p.Gln263X (Q263X) NA	8 Splice Site	Japanese, Chinese Japanese	16, 17, 19, 20 18

[a]The approved Human Genome Variation Society (HGVS) nomenclature with 'A' of the translation initiation codon ATG numbered as +1 was used to describe nucleotide numbering for coding level.
[b]Conventional numbering according to Asakai *et al.* [9] at the protein level omitting signal peptide counting start codon ATG as -18.
[c]Reference range 7.3-17.5 ng/ml
Abbreviations: Comp Hetero, compound heterozygote; Hetero, heterozygote; Homo, homozygote; F, female; M, male; NA, not available

FXI antigen was measured by an AssayMax Human Factor XI ELISA kit based on a biotynylated polyclonal antibody against FXI recognized by a streptavidin-peroxidase conjugate (Gentaur, Kampenhout, Belgium). Briefly, 50 ul of standard or samples per well was added and the wells were covered with sealing tape and incubated for 2 hours. The timer was set to start after the last sample addition. The wells were washed five times with wash buffer manually and 50 ul of Biotinylated FXI antibody was added to each well and incubated for 1 hour. The microplates was washed and 50 ul of streptavidin-peroxidase conjugate was added and incubated for 30 min. After washing the microplate, 50 ul of chromogen substrate per well and incubated for about 15 min until the color density developed. 50 ul of stop solution was added to each well the absorbance was read on a microplate reader at a wavelength of 450 nm immediately. The mean value of the triplicate for each standard and sample was calculated. The standard curve was generated using standard concentration on the x-axis and the corresponding mean 450 nm absorbance on the y-axis. The unknown sample concentration was determined from the standard curve and the dilution factor was multiplied to the value. The FXI antigen level was measured in two patients who were heterozygote for novel missense mutations. The reference range was determined by measuring FXI antigen levels in 30 individuals with normal APTT and PT.

3. DNA extraction, PCR amplification, and direct sequencing

DNA was extracted from EDTA whole blood samples with an Easy-DNA^{TM} Kit (Invitrogen Corporation, Carlsbad, CA, USA). Briefly, 100 ul of fresh blood was mixed to form a homogeneous solution. 50 ul of solution A was added and vortexed in 1 second intervals until evenly dispersed. After incubation at 65°C for 10 min, 20 ul of solution B was added, and vortexed vigorously until the precipitate moves freely in the tube and the sample is uniformly viscous. 70 ul of chloroform was added until the viscosity decreased and mixture was homogeneous. The solution was centrifuged at maximum speed in a microcentrifuge for 10 min at 4°C to separate the phases and form the interface. The aqueous phase was transferred to a fresh microcentrifuge tube. In order to precipitate DNA, 445 ul of TE buffer and 1 ml of 100% ethanol (-20°C) was added and mixed by inverting the tube. Then the mixture was incubated on ice for 30 min. After centrifuge at maximum speed in a microcentrifuge for 10 min at 4°C, the ethanol was removed. 500 ul of 80% ethanol was added to the tube and mixed by inversion. After centrifugation, the ethanol was removed and let dried for 5 min in a heat chamber. The DNA pellet was resuspended in distilled water for further experiment. The concentration and quality of genomic DNA was evaluated by Nanodrop (ND-1000, Thermo Scientific,Wilmington, DE, USA). We designed primers to amplify 15 exons and flanking introns of *F11* based on the sequences provided by Zucker et al[24]. The modified primer sequences are listed in Table 2.

One hundred haploid genomes from unrelated control individuals were also analyzed by direct sequencing of the exons bearing the novel missense mutations to confirm that control individuals did not have the mutations. PCR was performed on 100 ng of genomic DNA using an AccuPower^{TM} Premix (Bioneer, Daejeon, Korea) under the following amplification conditions: 94°C for 3 min followed by 50 cycles of 94°C for 1 min, 62°C for 10 sec and 72°C for 15 sec, and final extension at 72°C for 15 min. The PCR products were then purified using a QIAquick Gel Extraction Kit (Qiagen, Düsseldorf, Germany). For gel extraction, the PCR product was excised from the agarose gel and put in a clear tube to be weighted. 3 gel volume of QG buffer was added to the tube and incubated at 504°C for 10 min. After the agarose gel solubilize completely, 1 gel volume of isopropanol was added to the sample and transferred to a QIAquick spin column in a 2 ml collection tube. To bind DNA, the tube was centrifuged for 1 min at 8000 rpm. The flow-through was discarded and 500 ml of QG buffer was added and centrifuged for 1 min. In order to wash, 750 ml of PE buffer was added to the column and incubated for 5 min and centrifuged for 1 min. Flow through was again discarded and column was centrifuged for an additional 1 min at maximum speed to be dried. The column was transferred to a clean 1.5 ml microtube and 50 ul of distilled water was added and incubated for 10 min. The flow through was collected and labeled. The gel purified DNA was directly sequenced using a cycle method with the same primers for PCR and a Big Dye Terminator Cycle Sequencing Ready Reaction Kit (Applied

Biosystems, Foster City, CA, USA) with following conditions: 96°C for 5 min followed by 24 cycles of 96°C for 10 sec, 50°C for 5 sec and 60°C for 4 min, and final extension at 72°C for 5 min, in conjunction with an ABI Prism 310 automated genetic analyzer (Applied Biosystems). For the purification of product from sequencing PCR, 80 ul of distilled water was added to 20 ul of PCR product and centrifuged for 30 min at 12000 rpm. The supernatant was discarded and 250 ul of 70% ethanol was added and centrifuged for 15 min at maximum speed. The supernatant was discarded and dried in an oven for about 10 min. The product should not be overdried. 15 ul of HiDi solution was added to the sample and loaded on to the sequencing analyzer. To detect any sequence variation, the sequences were compared to the reference sequences using Sequencher software (Gene Codes Corporation, Ann Arbor, MI, USA).

4. Bioinformatics study

In order to predict functional effects of novel missense variations, we compared inter-species amino acid conservation using ClustalW[25] and assessed *in silico* prediction using protein function predicting software such as SIFT[26], Polyphen[27],and PMut[28] based on the information obtained from the Uniprot (http://www.ebi.uniprot.org/; Swiss-prot code/Uniprot number: P03951).

Table 2. Primer sequences for *F11* gene and haplotype analysis

	5' Forward 3'	5' Reverse 3'
F11 promoter	GACGCAGGACAATCACTTGAAC	GCTTAGAGGAATGTTAAATAGTAAGTTCTG
Exon 2	AGTCACACTAAGGAATGCTCCAGG	CAGTGTGATTTCCCTCTCCCAGC
Exon 3	ACATAACGCATGCCATGTAC	AAAAATCTGTCTCCTCGATG
Exon 4	GCTTTCTGTGTGCTGACTTT	CAGCTGGTATTTGTTGATTC
Exon 5	TGCCCCTAGAATCTGGAAGGT	TTTCATCGAACCACTCGAATGTCCTG
Exon 6	CTTAGCAACACTGCTGGG	GTGAGCATAAGCTGGTATC
Exon 7	TCCTGATAGCTGGTGAATTG	GAAGATAACAAATTATCCTTACTTG
Exon 8-10	CTGACTTTACTTTCTCTAGGTG	GTTCTCCCTTCTGTGGCTAT
Exon 11	AATGCTTCTGTTGCAGAGT	TTATAAATGTGTGAAGAAGATGAAC
Exon 12	GCCACACACTTCACAATGTC	TCAGGCCGTAAGTCTAGT
Exon 13	AAAATACACGACAACAAGGC	TCTAGGATGGAGCACATATAAC
Exon 14	ATGGTTATTCTACAAACGAACC	AACAGAGCGAGACTCTGTC
Exon 15	TCTGAGTTGATCTGTGCACC	CCAGAATCCAGTCCACGTACTC
-231C>T	GGCCATGAGAAGGATCTGAA	TTTCTTTCCCAATCCTGGAG
(CA)n	CACTCCCATCCTCCCATCTA	CGACAGAGCAAGACCCAGTT
-431G>A/-361C>T	GTAAGAAGGACTTAGCCA	CACATGAAGAAGGAGAGTG

5. Haplotype analysis

Haplotypes were determined by analyzing 10 polymorphic markers of *F11* (Table 3) in DNA samples obtained from 24 normal individuals (48 chromosomes) and 13 patients (26 chromosomes). The oligonucleotide primers used in *F11* haplotype analysis are listed in Table 2. In addition to polymorphisms -231C>T, -138A>C, $(CA)_n$, -431G>A, -361C>T, c.801A>G, c.1191T>C and c.1812G>T which were previously used in Jewish and Italian haplotyping studies[24, 25], -196T>G and c.1481-34G>T were also used to reconstruct haplotypes in Koreans in our study. Haplotype reconstruction and frequency estimation were performed using PHASE version 2.1 (http://www.stat.washington.edu/stephens/).

6. Estimating the Age of Mutation

DMLE +version 2.3 (http://dmle.org/) which uses a Bayesian inference with the Markov chain Monte Carlo method was utilized to estimate the age of the *F11* Q263X and Q226X from linkage disequilibrium data [29]. To estimate the rate of population growth in Korea, the data from the Korean Census and Statistics Department (http://www.kosis.kr/nsportal/common/meta_onedepth.jsp?vwcd=MT_CTITLE&listid=MT_CTITLE_A) (accessed 20 August 2010) was used which dates back to 1949 when the South and North Koreas separately began to conduct census. Since population has steadily decreased from 1949 to 2010, we used a

mean growth rate (1.3%) across the whole time period as well as the lowest (0.5%) and highest (2.8%) growth rates in 1960 and 2005, respectively. Since population has steadily decreased from 1949 to 2010, we used a mean growth rate (1.3%) across the whole time period as well as the lowest (0.5%) and highest (2.8%) growth rates in 1960 and 2005, respectively. In order to obtain a good convergence of the posterior probability frequency, 20 million iterations were performed. The genotypic data, genetic model, number of chromosomes in the disease sample and number of marker loci as well as map distances, and mutation location were included in the analysis. Other parameter ranges were set up as recommended by the program. Haplotypes used in the analysis consisted of 11 markers, including 10 polymorphic markers and 2 markers for the common mutations, in 12 chromosomes of the patients bearing two common mutations and 48 chromosomes of the control group. The genotypic data, genetic model, number of chromosomes in the disease sample and number of marker loci as well as map distances, and mutation location were included in the analysis. Other parameter ranges were set up as recommended by the program.

7. Statistical analysis

To show difference in haplotypes between the patient group and normal controls, chi-square analysis was performed chromosomes bearing Q263X or Q226X mutation and mutation negative chromosomes using SPSS (Statistical

Package for the Social Sciences, Chicago, IL) 14.0 for Windows. P value of <0.05 was considered as significant.

Table 3. *F11* haplotype of patients with factor XI deficiency and two common mutations observed in Koreans

Marker name[a]	Location	RS numbers	Position (bp)	Jewish or Italian haplotype	Cases 1/2/3		Case 5		Case 10		Case 11		Case 13	
-231C>T*	intron A	rs4253398	187425055	C	C	C	T	T	T	T	T	T	C	T
-196T>G*	intron A	rs4253399	187425088	-	T	T	T	G	T	T	T	T	T	T
-138A>C*	intron A	rs3822057	187425146	A	A	A	C	C	C	C	C	C	A	C
(CA)n	intron B		187428799	11	10	10	9	9	9	9	9	9	10	9
-431G>A*	intron E	rs2055916	187433504	G	G	G	A	A	A	A	A	A	G	A
-361C>T*	intron E	rs4253413	187433574	C	C	C	C	T	C	C	C	T	C	C
c.730C>T†	**exon 7**	**Q226X§**	**187434513**	**-**	**-**	**-**	**+**	**-**	**+**	**-**	**+**	**-**	**-**	**-**
c.755G>C†	exon 7	R234T§	187434538	**-**	**-**	**-**	**-**	**+**	**-**	**-**	**-**	**-**	**-**	**-**
c.801A>G†	exon 8	rs5974	187438205	G	A	A	A	A	A	A	A	A	A	A
c.841C>T†	**exon 8**	**Q263X§**	**187438245**	**-**	**+**	**+**	**-**	**-**	**-**	**-**	**-**	**-**	**+**	**-**
c.1135+1G>A†	intron J		187442240	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**+**
c.1191T>C†	exon 11	rs5970	187442295	C	T	T	T	T	T	T	T	T	T	T
c.1481-34G>T†	intron M	rs2289253	187444529	-	T	T	G	T	G	G	G	G	T	G
c.1556G>A†	exon 15	W501X§	187444638	-	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**+**	**-**	**-**
c.1578C>G†	exon 15	D508E§	187444660	-	**-**	**-**	**-**	**-**	**-**	**+**	**-**	**-**	**-**	**-**
c.1812G>T†	exon 15	rs5971	187446696	G	T	T	G	T	G	G	G	G	T	G

*Nucleotide number of the intronic polymorphisms have been assigned starting from the nearest splicing junction.

†The approved Human Genome Variation Society (HGVS) nomenclature with 'A' of the translation initiation codon ATG numbered as +1 was used to describe nucleotide numbering for coding level.

§Conventional numbering according to Asakai *et al* [9] at the protein level omitting signal peptide counting start codon ATG as -18.

[a]Markers used in Jewish or Italian haplotyping in previous studies are shaded in grey [12, 30].

III. RESULTS

Direct sequencing of all exons and their flanking introns from all patients revealed 13 distinct mutations (Table 1), including five novel missense mutations (Fig. 1A). In most patients (9 cases; 69%), mutations were identified during pre-operative screening or routine physical checkup. Others were found when examining for minor bleeding during surgical procedures or bleeding that is difficult to control (Table 1). Antigen levels were measured in patients with novel heterozygote mutations, and the levels corresponded to the FXI levels (Table 1). In the mutation analysis, the most frequently found mutations were Q263X (4 cases; 31%) and Q226X (3 cases; 23%) (Table 1). These two nonsense mutations were not found in 100 normal alleles. In this study, we identified five novel missense mutations: Q14R, R234T, H35Q, W383R, and D508E (Fig. 1A), which were not found in 100 normal alleles. All five were found at residues that are highly conserved across different species by using the ClustalW program (Fig. 1B). The Q14R and H35Q is located in A1 domain while R234T and W383R are located in A3 and SP domain respectively (Fig. 2A and 2B). The D508E was found in the region that is not fully determined structurally. Furthermore, the novel missense mutations were predicted to be pathogenic by two or more *in-silico* software program except for D508E (Table 4). Two missense mutations, R234L and R234I, were located at residue R234, and changing from arginine to lysine caused a disruption in normal mRNA splicing[14, 31]. At residue W383, a nonsense mutation was found with W228X in

a Chinese family with FXI deficiency which was caused by reduced secretion of F11[32]. The D508E mutation was found to be in a conserved region by ClustalW, suggesting that the mutated residue is crucial in the protein function. Considering that the patient harboring this mutation with Q226X showed a F11 activity level of 12% which seems to be low for a heterozygous patient, the mutation could possibly be pathogenic. Thus, further analysis of this sequence change is essential to confirm its significance in the function of the F11 protein. In this study, we also observed previously reported mutations[5, 18, 21-23] of C212R, V498M, c.1560dupG, W501X, G400V, and a splice site mutation c.1135+1G>A (Table 1). One splice site mutation c.1135+1G>A , previously reported in a Japanese patient, was also found, which causes a change in invariant nucleotide of the splicing site recognition sequence resulting in aberrant splicing of F11 mRNA[18]. The mutations occurred in exons 3, 7, 8, 11, 13, and 14, and most of the mutations were found in the apple 3 and catalytic domains, and others were located in the apple 1domain. One novel sequence variation, c.1-150_1-149InsAT, presumably a polymorphism, was also found in case 5 but *in silico* analysis using Human Splicing Finder 2.4 (http://www.umd.be/HSF/) revealed that the variation does not impose an effect on splicing. Interestingly, two mutations, V498M and c.1560dupG, which were together identified in a Korean patient[22], were also found in one of our patients (Case 8). To establish whether the Q226X and Q263X mutations are on different haplotypes, 10 variants found between intron A and exon 15 were used

to reconstruct *F11* haplotypes (Table 3). As a result, haplotypes bearing Q263X and Q226X in our patients were CTA10GCATTT and TTC9ACATGG, respectively (Table 5). The frequency of Q263X bearing haplotype was significantly different between controls (2% of the normal alleles) and patients (31% of mutant alleles) (P=0.001) (Table 5). The frequency of Q226X bearing haplotype was 19% in controls and 27% in patients, but the difference was not significant (P=0.596). The haplotypes bearing Q263X (CA10GCATT) or Q226X (TC9ACATG) mutation differed from the haplotypes bearing the type II mutation (CA11GCGCG) in the Jewish and European populations (Table 5**).** The SNPs flanking Q226X and Q263X were also different, indicating that these two mutations have arisen through a different route. For age estimation of the mutations Q226X and Q263X, DMLE+2.3 software was used using haplotype data from six affected patients with two common mutations and 24 unaffected controls (Table 3). Figure 3 shows the histogram data for the posterior distribution of the mutation age for Q226X and Q263X for the mean population growth rate of 1.3% with the highest peak at 10.7 and 10.8 generations, respectively. This indicates that the mutations Q226X and Q263X have arisen almost at the same time (around 214-321 and 216-324 years ago, respectively, assuming 20-30 years for a generation). By rerunning the analysis with the lowest (0.5%) and highest (2.8%) growth rates, the age of the mutations for Q226X were 24.4 and 5.3 generations and, for Q263X, were 24.3 and 5.3 generations, respectively (Fig. 3).

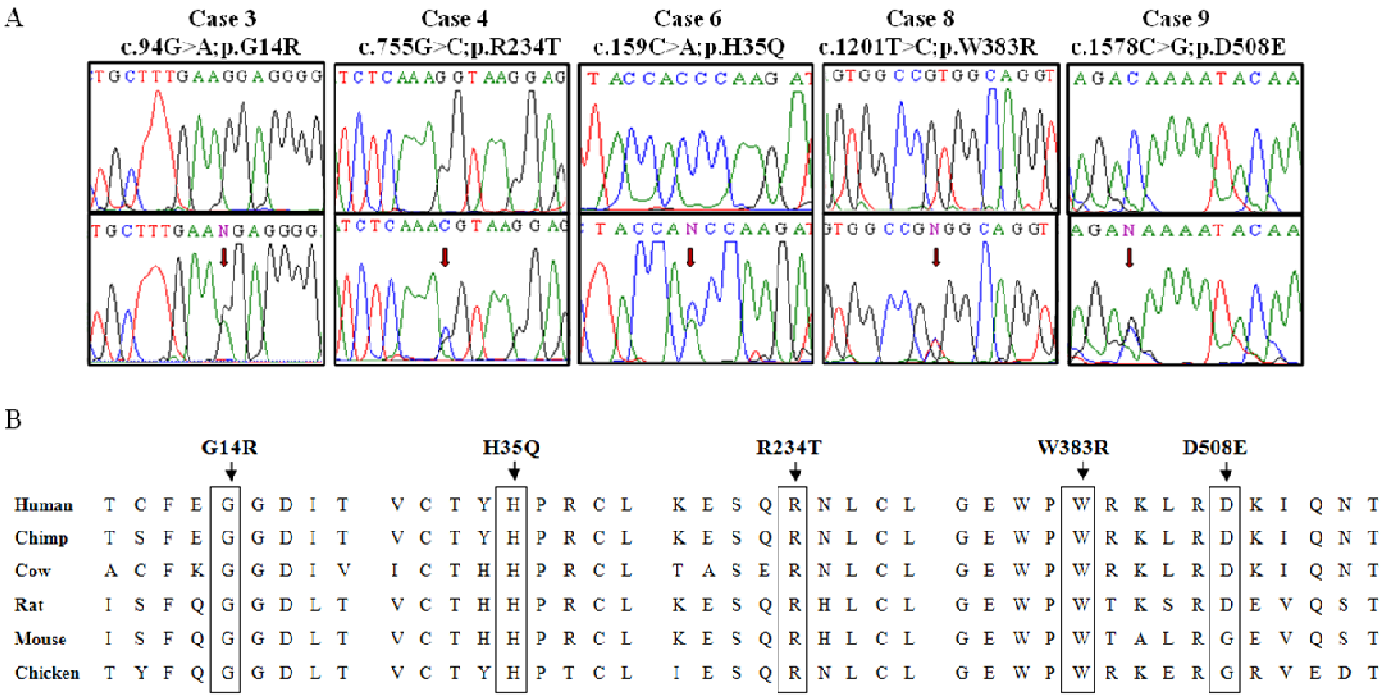

Figure 1. Sequence analysis and evolutionary assessment of *F11*. (A) Sequence chromatogram of five novel mutations found in this study. The mutation sites are indicated by red arrows. (B) Alignments of FXI amino acid sequences of different species using ClustalW. The sequences where the novel missense mutations occurred are conserved

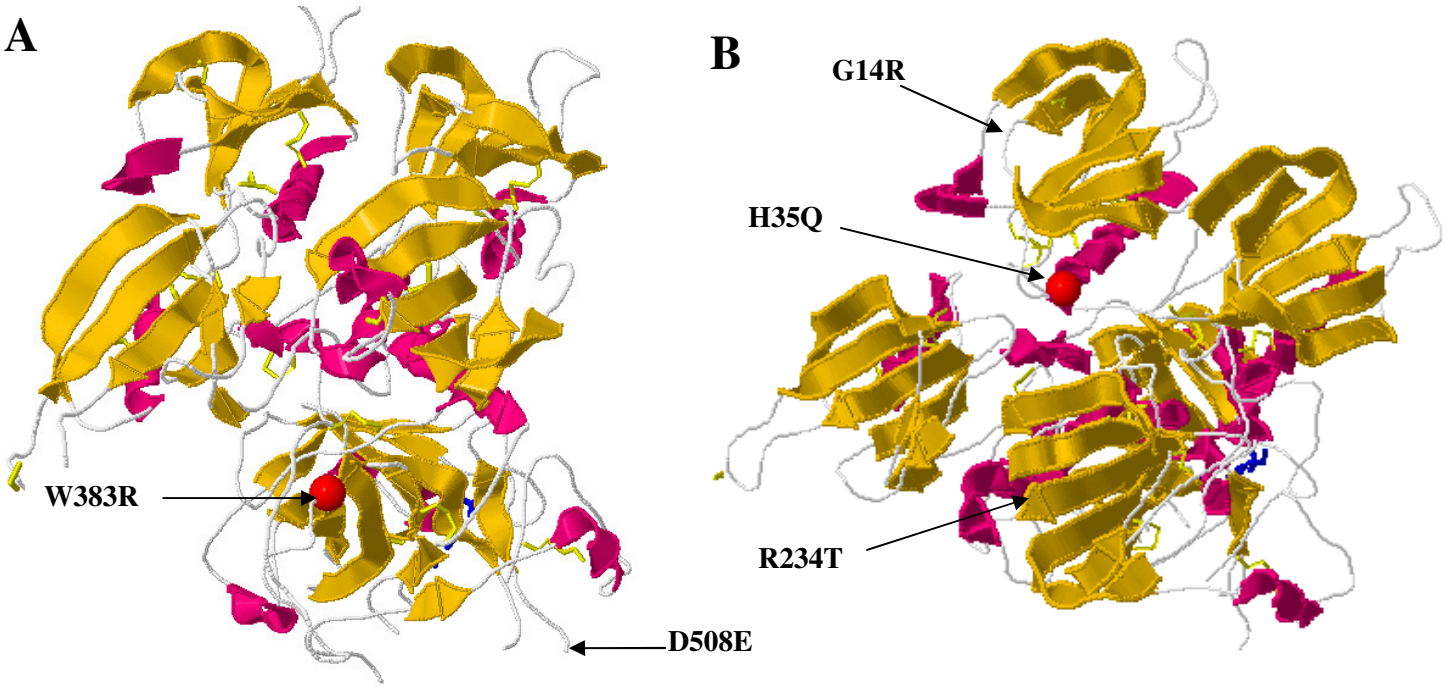

Figure 2. Ribbon view of the SP domain and four apple domains showing (A) mutations W383R and D508 in SP domain, and (B) G14R and H35Q in A1 and R234T in A4. The crystal structure was brought from the FXI database at www.FactorXI.org.

Table 4. *In silico* analysis results of 5 novel mutations

Case number	Mutation	Amino Acid Change according to HGVS (ATG at -18)	Polyphen	PMut	SIFT
3	c.94G>A	G32R (G14R)	Probable damage	Pathological	Not tolerated
4	c.755G>C	R252T (R234T)	Probable damage	Pathological	Not tolerated
6	c.159C>A	H53Q (H35Q)	Probable damage	Neutral	Not tolerated
8	c.1201T>C	W401R (W383R)	Probable damage	Pathological	Not tolerated
9	c.1578C>G	D526E (D508E)	Benign	Neutral	Not tolerated

Table 5. Haplotypes of *F11* polymorphisms in control and patient groups

Hap no	**Haplotype**	**Control**	**Patient**
1	TGC9ATATGG	13	3
2	TGC9GCATGG	1	0
3	TGC10GCATGG	1	0
4	TTC9ATATGG	2	0
5 *	TTC9ACATGG	9	**7**
6	TTC9ACATTT	1	0
7	TTA9ACATGG	1	1
8	TTA10GCATGG	2	1
9	TTA11GCATGG	1	1
10	TTA11GCGCGG	4	0
11	CTC9ACATGG	4	0
12	CTA9ACATGG	3	0
13	CTA9GCATTT	2	0
14	CTA10ACATGG	1	0
15	CTA10GCATGG	2	1
16 †	CTA10GCATTT	1	8
17	TTA11GCGTGG	0	1
18	TGC9ATATTT	0	1
19	TGC10ATACGG	0	2
	Total number of chromosomes:	48	26

*Q226X bearing haplotype
†Q263X bearing haplotype

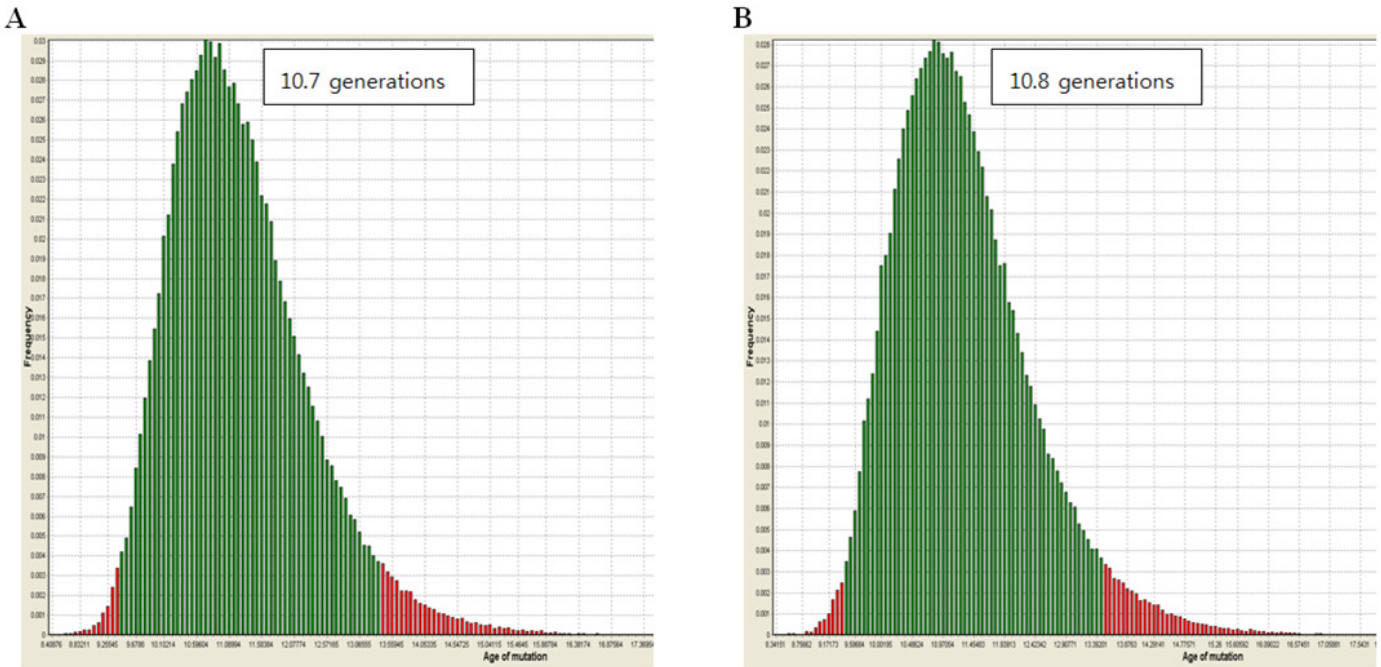

Figure 3. Age estimates for the *F11* Q263X (A) and Q226X (B) mutations with the mean population growth rate of 1.3. The histogram data for the posterior distribution of the mutation age, estimated by the software DMLE+2.3, are shown. In each graph, the generation at which the highest frequency reached is indicated in the box. The iterations were distributed in 200 bins shown as the outline of the columns in the figure.

IV. DISCUSSION

Factor XI deficiency is a rare congenital bleeding disorder. This study demonstrates a high prevalence of the Q263X and Q226X mutations in the Korean population and that the mutation spectrum of *F11* causing FXI deficiency in the Korean population is distinct from that of other populations including the Jewish or European populations.

The Q263X mutation has been reported repeatedly in Japanese and Chinese patients[16, 17, 19, 20]. Structurally, the Q263X mutation codes for a stop codon after a sequence coding for the third apple domain of the FXI protein, causing truncation of the fourth apple domain which is important for dimerization and secretion of FXI and the serine protease domain[17].

The high frequency of the Q263X mutation in Korean patients with FXI deficiency and the presence of significant difference in the frequency of the mutation bearing haplotype between the control and patient groups revealed a founder effect of the mutation. In addition, the results of the mutation age analysis indicate that the ancestral Q226X and Q263X mutations are most likely over 10 generations old, which indicates that these mutations were introduced to the Korean population between the late 1600s and late 1700s. Under the premise that E117X and F283L have arisen between 586 B.C. and 70 A.D. and between 19th and 20th century, respectively, the mutations frequently found in our patients may have arisen before F283L and after E117X appeared in the Jewish population[12]. In addition, based on the observation that the Q263X has

been reported in Chinese, Japanese, and Korean whereas the Q226X has never been reported in Chinese despite screening of a large number of patients may indicate that the Q263X has arisen in China and spread to Korea and Japan while Q226X has arisen later than Q263X in Korea and spread down to Japan. This presumption has to be validated through an interpopulation study among Chinese, Japanese and Korean.

According to our experience, it can be predicted that there may exist more F11 deficient patients in the Korean population than what is predicted for a non-Jewish population, which has been assumed to be one in one million [33]. Indeed, several studies have demonstrated that it may be more common in Caucasians in the U.K. and the Basque population, which was about one in 10,000[21, 34], representing about 1% of the population. In a study involving mostly Asian subjects with APTT prolongation[35], eight patients (4.5%) were diagnosed with F11 deficiency: six patients (3.4%) were considered to be partial (F11>20%) and two patients (1.1%) severe (F11<20%). Presumably, a partial deficiency can represent heterozygote individuals, whilst the severe phenotype indicates probably the homozygote or compound heterozygotes. Considering that APTT prolongation is found in about 5% of the Asian population[36], the incidence of F11 deficiency in the Asian population can be estimated to be about 17 in 10,000 for heterozygotes and 6 in 10,000 for homozygotes, providing an evidence that F11 deficiency may be more prevalent in the Asian population.

The Q226X mutation has been reported in Japanese patients with FXI deficiency[18], but never been reported in Chinese. The Q226X mutation disrupts the third apple domain, which is crucial for ligand binding to platelets, heparin, and other factor molecules in addition to losing apple 4 and serine protease domains[5, 18, 37]. Although the difference in the frequency of the haplotype bearing Q226X mutation between control and patient groups was not significant, the Q226X bearing haplotype was relatively more frequent in the patient group, suggesting another possible founder mutation in the Asian population. Apparently, the haplotype surrounding the Q263X mutation was distinct from the haplotype of Q226, showing that the two mutations originated from two different founders. It will be of great interest to study Japanese and Chinese patients with the same mutations to see if they share the mutation- bearing haplotypes and use this information in tracing the racial and historical relationship in different ethnicities.

V. CONCLUSION

The high frequency of the Q226X and Q263X mutations in our study suggests that testing for the presence of these two mutations should be the first genetic screening in Korean FXI deficient patients. By analyzing exons 7 and 8 of *F11* alone can detect up to 57.1% of the mutations, and by adding exon 13 to the analysis, the detection rate can be increased to 71.4%.

In conclusion, our study provides an opportunity to understand the mutation

spectrum of *F11* in the Korean population. The identification of a founder effect improves the genetic screening strategy to be followed and facilitates the clinical diagnosis of the disease.

REFERENCE

1. Seligsohn U. Factor XI in haemostasis and thrombosis: past, present and future. Thromb Haemost 2007;98:84-9.

2. Gomez K, Bolton-Maggs P. Factor XI deficiency. Haemophilia 2008;14:1183-9.

3. Papagrigoriou E, McEwan PA, Walsh PN, Emsley J. Crystal structure of the factor XI zymogen reveals a pathway for transactivation. Nat Struct Mol Biol 2006;13:557-8.

4. Berber E. Molecular characterization of FXI deficiency. Clinical and applied thrombosis/hemostasis : official journal of the International Academy of Clinical and Applied Thrombosis/Hemostasis 2011;17:27-32.

5. Kravtsov DV, Wu W, Meijers JC, Sun MF, Blinder MA, Dang TP, et al. Dominant factor XI deficiency caused by mutations in the factor XI catalytic domain. Blood 2004;104:128-34.

6. Shpilberg O, Peretz H, Zivelin A, Yatuv R, Chetrit A, Kulka T, et al.

One of the two common mutations causing factor XI deficiency in Ashkenazi Jews (type II) is also prevalent in Iraqi Jews, who represent the ancient gene pool of Jews. Blood 1995;85:429-32.

7. Seligsohn U. Factor XI deficiency in humans. J Thromb Haemost 2009;7 Suppl 1:84-7.

8. Kato A, Asakai R, Davie EW, Aoki N. Factor XI gene (F11) is located on the distal end of the long arm of human chromosome 4. Cytogenet Cell Genet 1989;52:77-8.

9. Asakai R, Davie EW, Chung DW. Organization of the gene for human factor XI. Biochemistry 1987;26:7221-8.

10. Saunders RE, O'Connell NM, Lee CA, Perry DJ, Perkins SJ. Factor XI deficiency database: an interactive web database of mutations, phenotypes, and structural analysis tools. Hum Mutat 2005;26:192-8.

11. Asakai R, Chung DW, Davie EW, Seligsohn U. Factor XI deficiency in Ashkenazi Jews in Israel. N Engl J Med 1991;325:153-8.

12. Peretz H, Mulai A, Usher S, Zivelin A, Segal A, Weisman Z, et al. The

two common mutations causing factor XI deficiency in Jews stem from distinct founders: one of ancient Middle Eastern origin and another of more recent European origin. Blood 1997;90:2654-9.

13. Zivelin A, Bauduer F, Ducout L, Peretz H, Rosenberg N, Yatuv R, et al. Factor XI deficiency in French Basques is caused predominantly by an ancestral Cys38Arg mutation in the factor XI gene. Blood 2002;99:2448-54.

14. Bolton-Maggs PH, Peretz H, Butler R, Mountford R, Keeney S, Zacharski L, et al. A common ancestral mutation (C128X) occurring in 11 non-Jewish families from the UK with factor XI deficiency. J Thromb Haemost 2004;2:918-24.

15. Quelin F, Trossaert M, Sigaud M, Mazancourt PD, Fressinaud E. Molecular basis of severe factor XI deficiency in seven families from the west of France. Seven novel mutations, including an ancient Q88X mutation. J Thromb Haemost 2004;2:71-6.

16. Au WY, Cheung JW, Lam CC, Kwong YL. Two factor XI mutations in a Chinese family with factor XI deficiency. Am J Hematol 2003;74:136-8.

17. Kawaguchi T, Koga S, Hongo H, Komiyama Y, Li K, Ishihara S, et al. A novel type of factor XI deficiency showing compound genetic abnormalities: a nonsense mutation and an impaired transcription. Int J Hematol 2000;71:84-9.

18. Okumura K, Kyotani M, Kawai R, Takagi A, Murate T, Yamamoto K, et al. Recurrent mutations of factor XI gene in Japanese. Int J Hematol 2006;83:462-3.

19. Sato E, Kawamata N, Kato A, Oshimi K. A novel mutation that leads to a congenital factor XI deficiency in a Japanese family. Am J Hematol 2000;63:165-9.

20. Wang J, Wang X, Dai J, Ding Q, Fu Q, Wang H, et al. A case of factor XI deficiency caused by compound heterozygous F11 gene mutation. Haemophilia 2009;15:603-6.

21. Mitchell M, Mountford R, Butler R, Alhaq A, Dai L, Savidge G, et al. Spectrum of factor XI (F11) mutations in the UK population--116 index cases and 140 mutations. Hum Mutat 2006;27:829.

22. Kwon MJ, Kim HJ, Bang SH, Kim SH. Severe factor XI deficiency in a Korean woman with a novel missense mutation (Val498Met) and duplication G mutation in exon 13 of the F11 gene. Blood Coagul Fibrinolysis 2008;19:679-83.

23. Iijima K, Udagawa A, Kawasaki H, Murakami F, Shimomura T, Ikawa S. A factor XI deficiency associated with a nonsense mutation (Trp501stop) in the catalytic domain. Br J Haematol 2000;111:556-8.

24. Zucker M, Zivelin A, Landau M, Salomon O, Kenet G, Bauduer F, et al. Characterization of seven novel mutations causing factor XI deficiency. Haematologica 2007;92:1375-80.

25. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, et al. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 2003;31:3497-500.

26. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. Genome Res 2001;11:863-74.

27. Ramensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. Nucleic Acids Res 2002;30:3894-900.

28. Ferrer-Costa C, Orozco M, de la Cruz X. Sequence-based prediction of pathological mutations. Proteins 2004;57:811-9.

29. Reeve JP, Rannala B. DMLE+: Bayesian linkage disequilibrium gene mapping. Bioinformatics 2002;18:894-5.

30. Zadra G, Asselta R, Malcovati M, Santagostino E, Peyvandi F, Mannucci PM, et al. Molecular genetic analysis of severe coagulation factor XI deficiency in six Italian patients. Haematologica 2004;89:1332-40.

31. Duncan EM, Casey GJ, Fenech MP, Lerda NV, Casey CR, Rodgers SE, et al. Partial and severe factor XI deficiency in South Australia and the usefulness of factor XI mutation analysis for diagnosis. Pathology 2008;40:401-6.

32. Wu WM, Wang HL, Wang XF, Chu HY, Fu QH, Ding QL, et al. [Identification of two novel factor XI non-sense mutation Trp228stop and Trp383stop in a Chinese pedigree of congenital factor XI deficiency]. Zhonghua Xue Ye Xue Za Zhi 2003;24:126-8.

33. O'Connell N M. Factor XI deficiency. Semin Hematol 2004;41:76-81.

34. Bauduer F, Dupreuilh F, Ducout L, Marti B. Factor XI deficiency in the French Basque Country. Haemophilia 1999;5:187-90.

35. Chng WJ, Sum C, Kuperan P. Causes of isolated prolonged activated partial thromboplastin time in an acute care general hospital. Singapore Med J 2005;46:450-6.

36. Li J, Lai X, Yan C, Xu A, Nie L, Zhou Y, et al. Age-associated developmental changes in the activated partial thromboplastin time (APTT) and causes of prolonged APTT values in healthy Chinese children. Clin Chem Lab Med 2009;47:1531-7.

37. Emsley J, McEwan PA, Gailani D. Structure and function of factor XI. Blood 2010;115:2569-77.

ABSTRACT (IN KOREAN)

한국인 XI응고인자결핍증 (C형 혈우병) 환자의 유전자돌연변이
스펙트럼 분석 및 분자적 특성 연구

<지도교수 이경아>

연세대학교 대학원 의학과

김 주 원

C형 혈우병으로도 불리우는 XI응고인자결핍증 (MIM#2649000)은 상염색체 열성으로 유전되는 유전질환으로 특히 Ashkenazi 유대인에서 가장 흔하게 보고되는 질환이다. 현재까지 *F11* 유전자 돌연변이 데이터베이스에 약 191개의 질병 관련 돌연변이가 보고됨 바 있다. 유대인에서 흔하게 보고된 E117X와 F283L 돌연변이가 C형 혈우병 환자에서 발견되는 돌연변이의 90% 이상을 차지하고 있으며 인종별로 일부 다른 돌연변이 스펙트럼이 보고되고 있다. 본 연구에서는 한국인 XI응고인자결핍증 환자를 대상으로 하여 XI응고인자 유전자의 돌연변이 스펙트럼을 분석하고 병인론적 기전을 규명하고자 하였다. 비 혈연관계에 있는 13명의 C형 혈우병 환자를 대상으로 연구를 진행하였다. 가장 흔하게 발견된 돌연변이는 Q263X (4명; 31%)와 Q226X (3명; 23%)이었으며 Q263X를 구성하는 일배체형의 빈도는 정상인과 환자군에서 유의한 차이를 보였고 이 결과는 Q263X가 한국인의 선조 돌연변이임

을 나타낸다. 5개의 새로운 돌연변이도 발견되었다. 한국인 C형 혈우병 환자에서 *F11* 유전자의 돌연변이를 검사할 경우 위 두개의 돌연변이에 대한 탐색이 가장 효율적인 돌연변이 선별 방법일 것으로 생각된다.

핵심되는 말 : XI응고인자결핍증, F11 유전자, 돌연변이 스펙트럼

PUBLICATION LIST

1. Kim J, Song J, Lyu CJ, Kim YR, Oh SH, Choi YC et al. Population-specific spectrum of the F11 mutations in Koreans: evidence for a founder effect. Clin Genet 2011 Jun, epub ahead of print