In presenting this dissertation/thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations, governing materials of this type. I agree that permission to copy from, or to publish, this thesis/dissertation may be granted by the professor under whose direction it was written, or, in his/her absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this thesis/dissertation which involves potential financial gain will not be allowed without written permission.

_____

Kevin Roland Viel

Refining the roles of genetic, environmental, and endogenous factors in Factor VIII activity levels

By

Kevin Roland Viel
Doctor of Philosophy

Department of Epidemiology

_____
W. Dana Flanders
Adviser

_____
Carolyn D. Drews-Botsch
Committee Member

_____
Michael Epstein
Committee Member

_____
Stephanie Sherman
Committee Member

_____
Laura Almasy
Committee Member

_____
Tom E. Howard
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

_____
Date

Refining the roles of genetic, environmental, and endogenous factors in Factor VIII activity levels

By

Kevin Roland Viel
B.S., B.S. Emory University, 1995

Adviser: W. Dana Flanders, M.D., D.Sc.

An Abstract of
A dissertation submitted to the Faculty of the Graduate School of Emory University in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Department of Epidemiology

2007

ABSTRACT

Background: Factor VIII (FVIII) is a central coagulation protein associated with hemophilia A at one end of the bleeding spectrum and thrombosis at the other end. A wide variety of factors are associated with increases in its activity, such as von Willebrand Factor (VWF), exertion, mental stress, inflammation, diabetes, and overseas air travel, but not mutations within the structural gene, *F8*. Recently, an association between depression and FVIII:C level has been reported.

Goals: The first goal of this dissertation was to characterize the genetic variations in *F8* using 222 X chromosomes. The second was to determine if any of the variants are associated with FVIII:C levels, controlling for appropriate variables. We then attempted to replicate the association in a different population of unrelated white and black women. Finally, we investigated the role of depression as a possible non-genetic determinant of FVIII activity levels in these women.

Results: We identified 47 genetic variants in *F8*, including five missense mutations (W0255C, R0484H, R0776G, D1241E, and M2238V) and five SNPs in the promoter region. Two SNPs (G056010A and C092714G, which encodes D1241E) were strongly associated with FVIII:C levels and, in this study of 21 families, were also in allelic association ($r^2 = 0.84$, unadjusted). We replicated the association of C092714 and FVIII:C level within the white subjects of the second study, but in the black subjects the result was not strong and included the possibility of no association. Due to resource limitations, we were unable to genotype G056010A. Among the black women depression was suggestively associated with higher FVIII:C levels, whereas among white women it was not. Finally, though not the focus of these investigations, we raised the possibility

that VWF antigen level and Factor IX activity level may not be just confounders, but also intermediates in a complex, unknown system of coordinate regulation.

Refining the roles of genetic, environmental, and endogenous factors in Factor VIII activity levels

By

Kevin Roland Viel
B.S., B.S. Emory University, 1995

Adviser: W. Dana Flanders, M.D., D.Sc.

A dissertation submitted to the Faculty of the Graduate School of Emory University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

2007

ACKNOWLEDGEMENTS

First, I am indebted to my advisor, Dana Flanders. Without his selfless guidance, these investigations would have been routine, instead of thorough and superlative, and the quality of the writing roundly mediocre. I would like to thank Carey Drews-Botsch for her inspiring interest in the substantive field and determination to address important methodological issues. Stephanie Sherman and Michael Epstein provided important expertise in genetics, in addition to wisdom and direction. Laura Almasy perfectly balanced my inclination to pursue myriad investigations with steady forward progress while providing a research assistantship in the second half of my studies. Finally, I would like to thank Tom Howard, who enabled the investigations and had unparalleled curiosity in the research. To each of the committee members, I am eternally grateful.

As studies have become enterprises, it would be impossible to acknowledge all of the participants required to be successful. In the interest of brevity, I would like to thank the members of the GAIT project in Barcelona, Spain; the members of CDC, especially Connie Miller; the members of the Howard Lab at Emory University; and the faculty and staff at the Genomics Computing Center of the Department of Genetics at the Southwest Foundation for Biomedical Research in San Antonio. Whether these people recruited patients, collected samples, analyzed samples, instructed me on programming or statistics, they were invaluable to this dissertation. In addition, I would like to thank the members of the studies for their time and, literally, blood, without which we could not elucidate the pathogenesis of diseases nor push the boundaries of science.

I reserve this special place to acknowledge the interactions I have had and hope to continue to have with Mitch Klein. A true scholar, Mitch has played a major role in my

education and in my dissertation. Not only has Mitch always been available, but numerous times he has also initiated contact with me. Mitch, I am truly grateful for your help and hope that for many decades to come students will benefit from your role as a teacher.

Finally, I would like to thank my family and friends for support and patience. Foremost among them is my son, Alex. Not only has he endured with me the sacrifices of scholarship, but his medical condition reminds me daily of the ultimate rewards of this field-innovative and improved patient care.

Table of Contents

Chapter 1. Introduction 1
I. Focus of the dissertation 2
II. Overview of the biological process of coagulation
A. Introduction 3
B. Zymogen precursors 4
C. Enzymatic coagulation cascade 5
D. The role of platelets 6
E. The importance of cofactors 7
F. Attenuation, inactivation, clearance, and fibrinolysis 8
III. Overview of the FVIII protein
A. Introduction 10
B. Structural organization 10
C. Biosynthesis 11
D. Post-translational modification 12
E. Secretion 13
F. Activation/Inactivation 14
G. Clearance 16
IV. Overview of measurement
A. Introduction 17
B. Antigen level 17
C. Activity level 18
D. Measurement variability 19
V. Overview of the genic data
A. Introduction 21
B. Overview of the *F8* structural gene 22
C. Generating DNA sequence data 26
VI. Family-based studies
A. Introduction 28
B. Identity by descent / Degree of relation 29
C. Violations of Mendelian inheritance 31
D. Heritability 33
E. Linkage Analysis 41
VII. Determinants of FVIII:C Levels
A. Introduction 46
B. Environmental 47
C. Endogenous 54
VIII. FVIII:C levels and thrombosis 80
IX. Goals of the dissertation 85

Chapter 2. A sequence variation scan of the coagulation Factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels. 87
• Abstract 88
• Introduction 88

Table of Contents (continued)

- Patients, materials, and methods 89
- Results 92
- Discussion 95
- Acknowledgements 97
- Authorships 97
- References 97

Chapter 3. The C allele of the Factor VIII (*F8*) 092714 single nucleotide polymorphism is additively associated with increased Factor VIII activity levels in white women, but only weakly associated in black women. 100

I. Introduction 101
II. Materials and Methods
A. Study Population 102
B. Genotyping 103
C. Statistical Analysis 104
III. Results 106
A. Main Findings 106
B. Sensitivity Analyses 107
C. Additional Adjustments 108
IV. Discussion 110
V. Appendix
A. Possible Selection bias 113
B. Control of confounding 114
C. Influential observations 115
D. Confounders or potentially co-regulated factors 115

Chapter 4. The association between Depression and plasma Factor VIII activity (FVIII:C) levels in reproductive aged women 118

I. Introduction 119
II. Materials and Methods
A. Study Population 120
B. CES-D instrument 121
C. Statistical Analysis. 122
III. Results 124
IV. Discussion 128
V. Appendix
A. Possible Selection bias 131
B. Impact of Adjustment for Factor IX and for VWF:Ag 133

Chapter 5. Conclusion 135

Table of Contents (continued)

I. Goals, Brief Overview of Background, and Importance 136
A. FVIII is central to proper hemostasis 136
B. Epidemiologic and Public Health Importance 137
II. Review of Major Findings
A. Genetic Variation in *F8* 138
B. Association between *F8* variants and FVIII:C level 139
C. Association between depression and FVIII:C level 139
D. New perspective on established covariates 140
III. Cohesion of results
A. Pre-requisite variant discovery 141
B. Replication is necessary 141
C. Adjustment for VWF:Ag and FIX:C level 142
D. Implication of replication 142
IV. Agreement with current knowledge
A. The association between C092714G and FVIII:C levels was novel 143
B. Depression and FVIII:C levels 143
V. Important Strengths 144
VI. Important Weaknesses 145
VII. Conclusions
A. FVIII has multiple forms 146
B. C092714G is associated with FVIII:C levels 146
C. The association between FVIII:C and depression was weakly supported 147
VIII. Future directions
A. Measurement of FVIII:Ag levels 147
B. Characterizing the proteins with which FVIII interacts 147
C. Explore associations with haplotypes 148
D. The role of Depression 148
E. Refine the definition of diabetes 149

References 150

List of Figures

Chapter 1.

Figure 1-1. A chromatogram displaying a subject doubly heterozygous at adjacent nucleotides 27

Figure 1-2. Hypothetical family of three full siblings demonstrating IBD and IBS status at an autosomal locus 31

Figure 1-3. Odd and even number of cross overs during meiosis 42

Figure 1-4. A hypothetical pedigree in which the disease cosegregates with the marker locus 44

Chapter 2.

Figure 2-1. Genomic structure and common polymorphisms of the human *F8* 93

Figure 2-2. Pattern and degree of LD (Linkage Disequilibrium) across *F8* 94

List of Tables

Chapter 1.

Table 1-1.	Domains and regions in FVIII protein	11
Table 1-2.	FVIII proteolysis sites of relevant proteases	14
Table 1-3.	The variance components estimated for FVIII:C using internal quality control materials	20
Table 1-4.	Genic regions of *F8* with nucleotide numbering referenced to hg17	24
Table 1-5.	Contribution of biallelic genotypes to the population phenotypic mean	34
Table 1-6.	Average effect and mean phenotypic value by gametic allele	35
Table 1-7.	Mean phenotypic value of the progeny by biallelic genotype	38
Table 1-8.	Breeding values by biallelic genotype	39
Table 1-9.	Parental haplotypes and possible gametes	43
Table 1-10.	Environment factors investigated for their potential effects on FVIII levels	54
Table 1-11.	The effect of age on mean FVIII:C levels	57
Table 1-12.	Mean FVIII:C levels by race, sex, and age from ARIC	58
Table 1-13.	The associations of sex with FVIII:C levels	59
Table 1-14.	The associations of race with FVIII:C levels	61
Table 1-15.	The association between BMI and mean FVIII:C levels adjusted for age, sex, and race in the ARIC study	61
Table 1-16.	The effects of BMI on mean FVIII:C levels	63
Table 1-17.	The association between mean FVIII:C levels adjusted for age, sex, and race, and diabetes and serum insulin in the ARIC study	64
Table 1-18.	The association between age-adjusted glucose and insulin levels and mean FVIII:C levels in the CHS study	65

List of Tables (continued)

Table 1-19. The age-adjusted mean FVIII:C levels by cholesterol level in the CHS study 67

Table 1-20. The FVIII:C levels of mothers obtained cross-sectionally across gestational age as found by Sie el al. 70

Table 1-21. The FVIII:C levels of mothers obtained cross-sectionally across gestational age as found by Sanchez-Luceros et al. 71

Table 1-22. The ANOVA results of the log transformed FVIII:C levels of 136 female blood donors 77

Table 1-23. FVIII:C levels by ABO phenotype, sex, and race in the CARDIA study 78

Chapter 2.

Table 2-1. *F8* regions resequenced to identify and genotype common polymorphisms 90

Table 2-2. Characteristics, status, and allelic associations of *F8* polymorphisms with FVIII:C levels 91

Table 2-3. Multicovariate analysis of 92714C>G (D1241E) and FVIII:C levels in the GAIT population 95

Chapter 3.

Table 3-1. The mean FVIII:C levels by race and genotype at the C097214G allele of the *F8* gene, which encodes the D1241E allele in the B domain, among the women of the CDC study of menorrhagia 107

Table 3-2. The characteristics of the 185 women in the CDC case-control study of menorrhagia genotyped at the *F8* 092714 nucleotide, which encodes amino acid 1241 in the B domain of FVIII 108

Table 3-3. Results of the multiple linear regressions of FVIII:C levels in the separate black-only and white-only populations 109

List of Tables (continued)

Table 3-A-1. The effects of adjustments for the potential bias resulting from unequal selection probabilities on the association between FVIII:C levels and the *F8* C(097214)G SNP 114

Table 3-A-2. The effects of excluding FIX:C and VWF:Ag levels on the estimate for the effect of the *F8* 097214 locus 116

Chapter 4.

Table 4-1. The mean natural log transformed FVIII:C level by race and depression status among the women of the CDC study of menorrhagia 124

Table 4-2. The characteristics of the women in the CDC case-control study of menorrhagia 125

Table 4-3. Percent Difference in Geometric Mean FVIII:C levels, from multivariate analyses in the black- and white-only populations 127

Table 4-A-1. The effects of adjustments for the potential bias resulting from unequal selection probabilities on the association between the FVIII:C level and depression 132

Table 4-A-2. The effects of excluding FIX:C and VWF:Ag levels on the association between depression and FVIII:C levels 133

Chapter 1

Introduction

1-I. Focus of the dissertations

Factor VIII (FVIII) activity, an *in vitro* measure of the action of the blood clotting Factor VIII protein, is indispensable to adequate coagulation. Partly because of the fatal nature of Factor VIII deficiency which results in the bleeding disorder known as hemophilia A, many factors affecting FVIII:C levels have been identified. As high FVIII:C levels have emerged as a risk factor for thrombosis, investigations of factors that increase FVIII:C levels have begun.

The associations between many of the preliminarily-identified risk factors and FVIII:C level may yet still be spurious, possibly due to uncontrolled confounding. Few investigations of FVIII:C levels have had extensive covariate data available, and fewer still have attempted adequate control in the analysis. Importantly, no study has identified a DNA sequence variation within the gene that encodes the FVIII protein, *F8*, that is associated with FVIII:C level. Thus, two main gaps in existing literature are present – identification of DNA polymorphisms that explain variation in FVIII activity, and studies that thoroughly control for confounding.

The first focus of this dissertation is to investigate and fill in these shortcomings. Thus, one focus will be to study a previously unidentified genetic component of the variance of FVIII:C level while providing adequate adjustment for confounding. Thus, the DNA sequence variations within the *F8* genes of the populations under study will be identified and subsequently tested for association with FVIII:C levels.

In addition to a genetic contribution to the variance of FVIII:C levels, many environmental and endogenous factors are also thought to affect FVIII:C levels, both

acutely and chronically. It is also important to investigate these non-genetic factors and to refine their relationships with FVIII:C levels. One such factor which has been reported to be associated with FVIII activity is depression. This common psychiatric illness and its possible link with FVIII activity levels is of potentially large public health importance, because of its relatively high frequency. The reported association of depression with procoagulant factors is worrisome, since high FVIII:C levels are both common and a risk factor for thrombosis. The second focus of this dissertation, therefore, is to investigate the association of depression with FVIII:C level while providing adequate control of potential confounding. The results of this dissertation should provide important knowledge that will help explain the large variance of FVIII:C level and thus begin to address the causal pathways involved in the pathogenesis of thrombotic diseases, such as deep vein thrombosis, myocardial infarction, and stroke.

1-II. Overview of the biological process of coagulation

1-II-A. Introduction

Blood is a connective tissue that transports vital materials throughout the body. Disrupting the integrity a component of the circulatory system could result in fatal blood loss. Further, the circulatory system may disseminate autologous or foreign matter, such as cancerous cell or bacteria, respectively, to the detriment of the host. To combat these events and to restore the integrity of the vasculature, a host defense mechanism exists which culminates in stasis through the generation of a platelet plug and, potentially, a

fibrin clot followed by deposition of white cells in the area of injury, and the ultimate restoration of the vessel through repair and, potentially, remodeling(1). Factor VIII (FVIII) is an indispensable component of an enzymatic complex central to the intrinsic pathway of coagulation.

1-II-B. Zymogen precursors

To protect against inappropriate or overly aggressive reactions that could result in ischemia and then infarction, the coagulation system as a whole is quiescent with most of its components in the form of **zymogens (**proenzymes) or procofactors at any given time during homeostatic conditions. Unlike fellow serine proteases of the chymotrypsin family, such as the homologous proteinases of digestion, the coagulation enzymes have high specificity apparently due to recognition of substrates by exosites, i.e. sites distinct from the active site of the enzyme(2). Activation of a given component requires proteolytic cleavage. At least for the vitamin-K dependent coagulation zymogens, such as Factors II, VII, IX, or X, macromolecular complexes of enzymes and cofactors reversibly assembled on cell membranes catalyze proteolysis, with the notable exception of activation of zymogenic Factor IX by activated Factor XI. With the exception of thrombin, which is relatively pan-specific, the coagulation enzymes activate a narrow range of cognate protein substrates.

1-II-C. Enzymatic coagulation cascade

The hallmark of coagulation system is the **coagulation cascade,** the amplifying progression of activations that greatly enhances the velocity of the reaction that results in the formation of the insoluble fibrin polymer that is impermeable to blood. For example, thrombin, in addition to proteolyzing fibrinogen to fibrin, is capable of activating upstream reactants such as FVIII and FX that lead to the formation of more thrombin.

The initiation of cascade occurs in response to two broad stimuli: exposure of blood to components that are not present physiologically and biochemical alterations. The former event, such as a mechanical injury that disrupts the endothelium or that ruptures a vessel, exposes the flowing blood to activators. The latter event, such as an immune response, which may involve Tissue Necrosis Factor or Interluekin-1, induces the synthesis of the activators.

The activator to which these events expose blood is Tissue Factor (TF), a single-chain membrane receptor for Factor VII (FVII). Most cells, with the exception of hepatocytes, constitutively express TF. Only when TF binds activated FVII (FVIIa) is the complex enzymatically active, i.e. capable of acting on a substrate. Under normal conditions, 99% of circulating FVII is in the inactive state(3, 4), illustrating the dormant nature of the system.

The TF:FVIIa complex has the zymogens Factor IX (FIX) and Factor X (FX) as substrates which it converts to activated FIX (FIXa) and activated FX (FXa), respectively. FIXa occupies a platelet binding site in which it complexes with its cofactor activated Factor VIII (FVIIIa) forming the tenase complex, so called because it

activates FX. FXa complexes with activated Factor V (FVa) on the surface of a platelet to form prothrombinase, so called because it cleaves prothrombin (FII) into thrombin (FIIa) and fragment 1.2. Thrombin has fibrinogen as its principal substrate which it cleaves into a fibrin monomer and two molecules each of the acidic fibrinopeptides A and B. Fibrinogen is a dimer consisting of two identical heterotrimers each composed of Aα, Bβ, and γ polypeptides explaining the total of four fibrinopeptides generated. Spontaneous polymerization of the monomers forms protofibrin and then a fibrin array. In the presence of Ca^{2+}, thrombin activates Factor XIII (FXIIIa), a transamidase that cross-links fibrin to form γ-dimers and α-polymers. This process ultimately creates the impermeable fibrin polymer and platelet plug that staunches the flow of blood.

1-II-D. The role of platelets

Platelets function as a first line of defense against blood loss during a mechanical disruption of the vasculature. The presence of platelets at a site of injury and, hence, their availability as a component of a plug that mechanically impedes blood loss is not a function of an immense number of circulating platelets. Instead, platelets adhere to sites of disrupted vascular integrity via von Willebrand Factor (VWF), which couples exposed collagen in the extra-cellular matrix (ECM) to glycoprotein (GP) Ib. Further aggregation occurs when platelets use fibrinogen to couple each other via GP IIb-IIIa, resulting in affective hemostasis with reduced platelet numbers.

Importantly, activated platelets, which aggregate at the site of injury, bind certain components of the cascade. The anionic phospholipid surface of platelets, which

becomes enriched with phosphotidylserine upon activation(4), increases the frequency of favorably oriented interactions of enzymes, cofactors, and substrates by restricting them to a two-dimensional surface (5, 6). Thus, the propagation and amplification of coagulation reactions occurs on localized cell surfaces, instead of what would otherwise be the vastness, and potential remoteness, of the three-dimensional lumen of a vessel.

The tenase complex amply illustrates the efficiency gained by focal concentration of the reactants and products: FIXa, FVIIIa, and FX bind their respective receptors on the activated platelet, platelet-bound FIXa has a high affinity site for platelet-bound FVIIIa which has a high affinity site for platelet-bound FX(7), which is the substrate for tenase (the membrane bound FIXa:FVIIIa complex). The product, FXa, is already localized to the activated platelet and able to form, with FVa, the membrane bound prothrombinase complex mentioned earlier. Since these platelets are localized to the site of injury, so is the thrombin produced on their cell surfaces.

1-II-E. The importance of cofactors

The interactions between enzymes, cofactors, and substrates may seem like an improbable series of events, even when assisted by platelet binding and amplifying effects of the cascade. For example, the assembly and function of the tenase complex may involve a series of six binary interactions: FVIIIa/membrane, FIXa/membrane, FX/membrane, FVIIIa/FIXa, FIXa/FX, and FVIIIa/FX (though not necessary in that order)(8). The presence of cofactors, however, not only increases the velocity of the reaction, they are indispensable to proper coagulation.

For example, FVIIIa enhances the relative catalytic efficacy of FIXa by several orders of magnitude(3, 5, 9). A deficiency in FVIII activity, however, leads to the bleeding associated with hemophilia A emphasizing a central role of this cofactor in proper hemostasis.

1-II-F. Attenuation, inactivation, clearance, and fibrinolysis

To be effective either at stemming blood loss or inhibiting dissemination of an inappropriate material, the reactions of coagulation must be rapid. Otherwise, blood flow might disseminate the reactants, now activated and in complex, from the site or might disperse material that may escape the reaction meant to trap it. The speed of the reaction must be balanced by **attenuation**, which protects against an inappropriately aggressive response through inhibition, inactivation, or clearance.

Inhibiting the (further) activation of products, such as the actions of Tissue Factor Pathway Inhibitor (TFPI) or antithrombin, limits the intensity of the localized reaction. TFPI ligates FXa to the TF:FVIIa complex, thereby suppressing continued activation of FIX and FX. Antithrombin, however, inhibits the activated products thrombin, FXa, and, FIXa which dampens the reaction.

Inactivation of the activated products, in addition to terminating the reaction, may also attenuate it. Activated Protein C (APC) which has as a cofactor Protein S (PS), inactivates FVa and FVIIIa. Interestingly, a widely known genetic risk factor for thrombosis, the FV-Leiden mutation, is an amino acid change at residue 506 resulting

from a single nucleotide polymorphism (SNP) that destroys the active site for PC cleavage of FVa(10).

Clearance of the activated products may also attenuate the reaction(5). The *Clearance section (II-F)* of the overview of the FVIII protein presents a more thorough discussion of the clearance of FVIII. For example, Ananyeva et al.(11) suggested that the low-density lipoprotein receptor-related protein (LRP), which not only binds FVIIIa, FIXa, and FXa but is also abundant on the types of cells, such as monocytes, smooth muscle cells (SMC), and fibroblast, that injury exposes to blood may clear these ligands by endocytosis. Further, Ananyeva et al.(11) characterized the potential to prolong the half-life of infused recombinant FVIII through an engineered disruption in the binding sites that interact with the receptors of a clearance pathway as "attractive", perhaps indicating the importance of clearance in the attenuation of the coagulation reaction.

To allow the return of normal blood flow, **fibrinolysis** must proteolyze arginyl-valinyl bonds of the fibrin polymer resulting in the soluble fibrin fragments D and E. The serine protease responsible for this action is the zymogen plasminogen. For activation of plasminogen to plasmin to occur, plasminogen activator inhibitor-1 (PAI-1) must first relinquish it. PC may function in initiation fibrinolysis by freeing plasminogen from PAI-1(9). The activators of plasminogen, tissue plasminogen activator (t-PA) and urokinase plasminogen activator (u-PA), are synthesized by the endothelial cells after stimulation by thrombin and perturbation, respectively. Both plasminogen and t-PA bind fibrin, providing yet again efficient localization the reactants. When stasis occurs and the reaction terminates, repair and remodeling of the vessel architecture may occur.

1-III. Overview of the FVIII protein

1-III-A. Introduction

FVIII, a procofactor for FIXa, is a labile glycoprotein with extensive post-translational modifications that exists as a heterogeneous population of heterodimers. The estimates of concentration in "normal" individuals range from 100-200 ng/mL (9, 12). It circulates in a noncovalent complex with VWF, which protects it from proteolytic attack(4, 13, 14), prematurely binding to components of FX-activating complex(6, 15), and clearance(11, 16, 17). FVIII has a short half-life of 0.3-0.5 days(3) which is greatly shortened in the absence of VWF(13, 14), as occurs in the autosomal recessive von Willebrand disease.

1-III-B. Structural organization

FVIII is composed of 2351 amino acids, the first 19 of which comprise a signal peptide yielding a mature protein of 2332 amino acids. It shares structural homology with FV and ceruloplasmin (a copper-binding protein), among other proteins. We classifying the linear sequence of amino acids into ordered domains: A1-A2-B-A3-C1-C2. This unconventional use of the term "domain", however, does not connote an independent folding unit. Rather, it reflects function and internal homology, such as the 30% homology between the A domains. In addition, FVIII contains three noteworthy acid regions a1, a2, and a3. a1 follows A1 while a2 and a3 flank the B domain. **Table 1-1**

presents the delineation of Saenko et al.(5) Alterative assignments of the amino acids to the domains exist(5, 8) and novel functional regions or homology may further refine classification the domains.

Table 1-1. Domains and regions in FVIII protein

Domain/Region	Amino Acids
A1	0001-0336
a1	0337-0372
A2	0373-0719
a2	0720-0740
B	0741-1648
a3	1649-1689
A3	1690-2019
C1	2020-2172
C2	2173-2332

The light (A3-C1-C2) and heavy (A1-A2-B) chains circulate as a heterodimer. The association of these chains depends upon a copper cation, potentially Cu(I), the monovalent form(18). For details of the chains and the processes yielding them, see on the *Post-translational modification section* (1-III-D) and *Activation section* (1-III-F).

1-III-C. Biosynthesis

The hepatocytes **synthesize** FVIII, though several other tissues may express FVIII mRNA. Studies in which liver transplantation corrected the bleeding tendency in hemophilia A patients(19, 20) or a case study of a patient acquiring hemophilia A after liver transplantation(21) suggest that the liver is an important site of FVIII synthesis.

Based on an extension rate of 10-15 nucleotides per second(15, 22), a hepatocyte will require approximately 3-5 hours to transcribe the 180 Kb base gene. Though the translation of FVIII and FV mRNA may proceed with similar efficiencies(23), the Factor VIII mRNA is less stable than that of FV(24).

1-III-D. Post-translational modification

After translation, the FVIII protein translocates to the lumen of the endoplasmic reticulum (ER). This process cleaves the signal peptide and adds high-mannose-containing core oligosaccharides to multiple asparagine residues once FVIII is inside the lumen(4). Several chaperon molecules, including immunoglobulin-binding protein (BiP), calnexin (CNX), and calreticulin (CRT), bind FVIII. The B domain may be important with respect to this process(4). BiP immediately binds FVIII, releasing it after ATP hydrolysis which requires high levels of intracellular ATP(24). Misfolded FVIII then interacts with CNX and CRT, which may directed it to the ER-associated degradation pathway(15).

A large portion (4) of the FVIII molecules fail to escape the chaperon-mediated retention and, therefore, degradation. Fully folded FVIII, however, transits to the Golgi apparatus via the endoplasmic reticulum-Golgi intermediate compartment (ERGIC) in anterograde trafficking as the cargo proteins of vesicles. An ERGIC localized mannose-binding type 1 transmembrane protein, LMAN1 (formerly ERGIC-53), assists in this trafficking. Recently, Nichols et al.(25) identified LMAN1 as the defect responsible for combined FV/FVIII deficiency (CF5F8D). Patients with this autosomal recessive disease

have greatly reduced, but significant residual FV and FVIII activity, suggesting that LMAN1 affects the efficiency of secretion(24).

Once in the Golgi apparatus, an unknown endoprotease cleaves within the B-domain and at the A2-B domainal junction, yielding the constant-sized light and variable-sized heavy chains of the heterodimeric FVIII(4, 15). Further processing includes modification of the asparagine-linked high-mannose containing oligosaccharides to complex types, addition of carbohydrate to multiple serine and threonine residues within the B domain, addition of sulfate to six tyrosine residues, disulfide bonding, and protein folding(15, 26).

1-III-E. Secretion

Post-translationally modified FVIII proceeds via the trans-Golgi compartment to the secretory compartment(27). After secretion into the blood, FVIII immediately forms noncovalent complexes with VWF (5, 15), which may displace the light chain from the cell surface and then promote its association with the heavy chain(18). From this point on, approximately 95% of FVIII exists in a bound state under equilibrium conditions(26). Apparently, FVIII is labile or otherwise subject to proteolysis or clearance. VWF, while acting as a carrier protein, protects FVIII from proteolysis (4, 5, 15) and may prevent premature (inappropriate) formation of the tenase complex(5, 6, 15).

1-III-F. Activation/Inactivation

Four proteases function in relevant proteolysis of FVIII. Thrombin and FXa procoagulantly cleave FVIII, while FIXa, FXa, and APC cleave FVIII which inactivates it. **Table 1-2** presents the amino acid preceding the cleavage site and the proteases involved.

Table 1-2. FVIII proteolysis sites of relevant proteases

Amino Acid	Protease(s)	(P)rocoagulant/(I)nactivation
Lys36	FXa	I
Arg336	FIXa, FXa, APC	I
Arg372	Thrombin, FXa	P
Arg562	APC	I
Arg740	Thrombin, FXa	P
Arg1689	Thrombin, FXa	P
Arg1721	FXa	likely benign(6)

Thrombin proteolysis of FVIII occurs in the apparently ordered cleavage of Arg740 followed by Arg336. Not only does single cleavage at site Arg740 yield a more catalytic FVIIIa than single cleavage at site Arg336, but the FVIIIa produced by thrombin proteolysis is also more catalytically active than that produced by FXa(6, 8).

As mentioned earlier, VWF binding protects FVIII from proteolytic cleavage. However, cleavage of FVIII at the A2-B domainal junction induces a conformational change in FVIII which disrupts VWF binding. The C2 domain and perhaps the NH_2-terminal region seem important for binding wWf(8), the affect of which may be to inhibit

the interaction of FVIII with phospholipid surfaces(6). Collectively, these cleavages release the B domain, yielding the heterotrimeric activated FVIII molecule (see **Table 1-1** for the domainal delineations).

Though both FIXa and FXa may proteolytically inactivate FVIIIa and are clearly in proximity to do the job, they may not provide much physiologic relevance to the loss of tenase function. In fact, association with FIXa more likely prolongs functionality of FVIIIa by stabilizing the A2 domain(6). It is this dissociation of the A2 domain that is a more important mode of the loss of tenase function, likely to the point of being regulatory(6, 8).

While FVIIIa is in complex with FIXa, APC is not able to proteolyze it at Arg562(6). However, in the presence of its cofactor, PS, APC may proteolyze this bond. Regardless, it appears that this pathway is secondary in the dampening of tenase activity(6).

The B domain is intriguing, if for nothing else than its apparent lack of similarity to other proteins or the absence of its role in FVIII activity(28). The heterogeneous nature of the size of the heavy chain (A1-A2-B) of FVIII is due to variable cleavage within the B. Indeed, naturally occurring and engineered recombinant heterodimers may lack this region and have “normal” activity(6, 26). Regardless of the content, activation entails cleavages which free what domain may be present since no interactions between the products exist. This domain, which accounts for approximately half of the mass of FVIII(8), has no apparent homology to date(18), and has wide inter-species variability(26). The lack of homology extends to FV, a molecule of relevance to FVIII due to similarity in homology and cofactor function, but the B domains of both molecules

contain a number of asparagine-like oligosaccharides(18). Considering the expenditure of cellular resources required to make this domain, this seems quite perplexing.

1-III-G. Clearance

FVIII, non-covalently bound to VWF, binds the heparan sulfate proteoglycans (HSPGs) which are a constituent of the extracellular matrix of cells(29), such as monocytes, SMC, fibroblast, and prominently hepatocytes(11, 30). The HSPG presents the VWF/FVIII complex to the low-density lipoprotein receptor-related protein (LRP), but may also function as an independent catabolic receptor(29). LRP endocytoses FVIII after without VWF resulting in ultimate lysomal degradation of FVIII and VWF released to the plasma(29).

Results from transgenic mice and in vitro experiments suggest that the clearance of FVIII occurs more efficiency in the absence of VWF(16, 17). Saenko et al. reported that a region in the A2 domain binds LRP, with a secondary region in the C2 domain(17). The spontaneous dissociation of the A2 domain, a major process in the dampening of the tenase complex, makes this notable.

Infusion of receptor-associated protein, which binds LRP with high affinity, suppresses but does not completely inhibit clearance, suggesting the presence of at least one other pathway(16, 17). Bovenshen et al. used LRP-deficient mice to demonstrate an elevation in FVIII activity levels (31). Two other members of this receptor family, low density lipoprotein receptor (LDLR) and very low density lipoprotein receptor (VLDLR),

may act in the clearance of FVIII, but their effects may be in concert with LRP and may not be important for *in vivo* regulation(32, 33).

1-IV. Overview of measurement

1-IV-A. Introduction

Broadly, there are two related measurements for most components of the cascade. The first measures the amount of protein present and the second quantifies the proteolytic action of the protein or its complex. These measurements may be closely correlated, especially if the components of the coagulations systems in question are marked similar. As an example to the contrary, consider a proteinase in which a mutation disrupted the active site. As long as the mutation does not affect synthesis, secretion, or clearance, then the amount of this protein would be "normal", however this mutation would diminish or obliterate the functionality of the protein.

1-IV-B. Antigen level

Typically, measurement of the amount (concentration in moles or mass per unit volume) of a protein employs monoclonal antibodies against the protein of interest. A stationary surface immobilizes the first antibody which binds to an epitope on the protein of interest. A separate epitope of the protein binds a second anti-body to which a marker

enzyme is conjugated. The immobilized marker enzyme allows for the determination of the **antigen level**. The symbol ":Ag " suffixed to the protein name indicates the results of this test, i.e. FVIII:Ag. The enzyme-linked immunoabsorbent (ELISA) test is one such anti-body based measurement.

1-IV-C. Activity level

Apart from the amount of protein present, we are also interested in the **activity level** of the protein, that is to quantify the ability of the protein to perform its task. Usually the endpoint in question is the formation of a clot, which might require characterizing the quantification as relative for reasons which will be apparent upon understanding the principal of the bioassay. To measure the activity of a protein, the protocol must standardize every other component of the reaction. Upon achieving this standardization, one may infer that differences in the time to endpoint result from different activities of the protein of interest. Typically, one obtains plasma deficient in the protein of interest. In the case of FVIII, patients with hemophilia A were once a source of this plasma. However, the use of an monoclonal anti-body against the protein of interest will artificially deplete normal plasma(34). The reference for this measurement is usually plasma pooled from a sufficient number of normal donors which has been calibrated against the standard. The investigator then serially dilutes the reference plasma and the test plasma with an excess of depleted plasma. This excess should obscure any differences in the activities of other factors that might exist between the reference plasma and the test plasma. Over a suitable range, decreasing dilutions should

linearly decrease the time to clot formation in a given plasma. If the results suggest that the regression lines are parallel, then one may interpolate the activity of the test plasma using as the clotting time of the reference plasma with 100 U/dL(4). In the case of FVIII, one unit (U) is the amount of activity in 1 mL of pooled normal plasma(28, 35). The symbol ":C" suffixed to the protein name indicates the results of this test, i.e. FVIII:C. The availability of commercial instruments and kits has provided automation and increased standardization to facilities performing these measurements.

1-IV-D. Measurement variability

Levels of FVIII have broad variability with a clinical "normal" range of 50%-150%. For instance, from the baseline measurements of subjects in the Atherosclerosis Risk in Community (ARIC) study, which consisted of more than 4,000 each of white men and women and more than 1,000 each of black men and women, the FVIII:C had the following mean (SD): 121(31), 126(33), 138(39), and 145(43), respectively, which did not account for the weighted sampling scheme(36). The precision of the measurements may have had a non- negligible impact on these results, but the design and analyses of the study did not quantify it. It is necessary, therefore, to identify the potential sources of error (deviations from the true level) in an investigation of the broad biologic variability of FVIII levels.

Chambless et al.(37) performed a quality control evaluation for the ARIC study which included FVIII:C levels. In each of the 2-3 daily runs they included 3-4 internal

quality control samples, which included the pooled plasma of 20 healthy subjects (see the above section for details). They used a nested two-factor random-effects model:

$$\mathbf{Y_{ijk} = \mu_i + \alpha_{j(i)} + \varepsilon_{k(ij)}}$$

$$\mathbf{\sigma^2_T = \sigma^2_{BD} + \sigma^2_{BR} + \sigma^2_{WR}}$$

where Y_{ijk} is the measurement of the k^{th} sample on the j^{th} run of the i^{th} day, μ_i is the mean on the i^{th} day, $\alpha_{j(i)}$ is the deviation of the j^{th} run nested in the i^{th} day, $\varepsilon_{k(ij)}$ is the (measurement) error for the k^{th} sample nested in the j^{th} run of the i^{th} day, σ^2_T is the total variance, σ^2_{BD} is the between-day variance, σ^2_{BR} is the between-run variance, and σ^2_{WR} is the within-run variance. **Table 1-3** presents the variance components as percentages for the periods October 1987 – September 1988 and October 1988 – October 1989. Notably, the estimated between day variance accounted for 40.0% and 25.7%, respectively, of the estimated total variance(37). The coefficient of variations were 6.0% and 7.6%, respectively.

Table 1-3. The variance components estimated for FVIII:C using internal quality control materials(37)

Period	n	Mean (%)	SD	% of total variance*		
				σ^2_{BD}	σ^2_{BR}	σ^2_{WR}
October 1987 – September 1988	1,671	89.7	5.4	40.0	0	60.0
October 1988 – October 1989	2,178	90.4	6.9	25.7	0	74.3

*See the model description above for an explanation of the symbols.

Results from 816 replicated measurement made blindly a week apart in this same study indicated that the measurement of FVIII:C levels were highly repeatable(37). Thompson et al.(35), using a volunteer panel of 14 subjects, determined that FVIII:C had a variance components profile similar to that of triglycerides, but not as favorable as that of cholesterol. It appears that the contribution of method variance, i.e. variance due to phlebotomy, handling, storage, shipping, or assay procedures, constitutes an acceptable proportion of the total variance of FVIII levels for epidemiologic studies.

1-V. Overview of the genic data

1-V-A. Introduction

The absence of FVIII functionality may follow two prominent inheritance patterns. The first is autosomal recessive and the second is a recessive, sex-linked inheritance pattern. These patterns, when available, discriminate between hemophilia A from the functional absence of FVIII associated with certain forms of von Willebrand disease, which is located on autosomal chromosome 12. The structural FVIII gene, in contrast, is located at chromosome position Xq28.1. In an epic race, two groups cloned FVIII in 1984(22, 38), marking the dawn of recombinant treatment of hemophilia A and a more thorough characterization of the FVIII gene.

1-V-B. Overview of the *F8* structural gene

FVIII refers to the Factor VIII protein, whereas *F8* refers to the structural gene. The 187 kilobase (Kb) *F8* gene accounts for approximately 0.1% of the 155 Mb X chromosome. It is oriented with its 5' untranslated region towards the telomere (end) of the long arm of the X chromosome. Thus, transcription proceeds towards the centromere.

I obtained the *F8* sequence plus 50 Kb of both 5' and 3' genomic sequence from the UCSC Genome Browser on October 12, 2004(39), which corresponds to NCBI Build 35. This is the first reference sequence to contain all of the exons; specifically, it included exon 21 and exon 22. Mansvelt et al.(40) provided sequence of transcription start site. For purposes of referencing the gene and variants, I complimented this sequence and assigned +1 to the transcription start site and –1 to the base immediately 5' to it. For instance, the promoter contains nucleotides with negative numbers. To indicate references to this nucleotide numbering, we use hg17.

A segment containing most of the (exonic) 5' untranslated region (UTR) comprising 1.2 Kb sufficiently promoted transcription in several liver-derived cell lines(12, 41). Furthermore, a segment containing approximately 300 bases 5' to the transcription start site showed maximal promoter activity(41). Using the convention that enumerates the first base of the start of transcription as +1 and the base immediately 5' to this site as –1, then beginning at base –30 there exists a GATAAA sequence, the putative TATA box that is typically relevant to transcription.

This region of 300 bases contains several cis-acting elements that bind transcription factors such as Hepatocyte Nuclear Factor (HNF), CCAATT/enhancer

binding proteins (CEBP), D-site binding protein (DBP), and Nuclear Factor kappa-B (NF-κB) that are relevant to initiation of transcription. Importantly, disruption of the alternative TATA sequence does not suppress transcription in vitro(12). Additionally, a repressor element may exist approximately 1 Kb 5' to the gene(12).

This suggests that promotion of transcription may be robust to genetic variation in the promoter, particularly in the case single nucleotide polymorphisms (SNPs) which do not cause length changes, which could radically affect promotion since the DNA-binding proteins overlap the transcription start site, potentially acting to stabilize the preinitiation complex(12). Therefore, the presence of variants in this region of the promoter, instead of causing pathology such as transcription failure, might rather contribute to "normal" variation of FVIII levels through variable transcription rates.

The length of the 5' UTR is 171 base pairs(22). FVIII contains 26 exons, with all of the splice donor/acceptors sites conforming to the canonical GT…AG pattern(22). On October 11, 2004, the National Center for Biotechnology Information (NCBI) updated its record for FVIII which filled a gap the spanned part of Intron 20 to Intron 23. **Table 1-4** presents the size and order of the genic regions of FVIII included in the NCBI human genome build 35 release.

The vast majority of the 187 kilobase (Kb) gene is bulk and repetitive content, typical of the genome in general. The exons, which account for approximately 5% of the gene, code for 2351 amino acids, inclusive of a 19 amino acid propeptide. Exon 14 is the largest, with 3,106 nucleotide bases which code for the entire B domain and parts of the acidic regions a2 and a3 (see **Table 1-1**, section 1-III-B).

Table 1-4. Genic regions of *F8* with nucleotide numbering referenced to hg17

Region	Base pairs	Nucleotides	Amino Acids	Region	Base pairs	Nucleotides
5' UTR	171	1–171				
Exon 01	143	172–314	0001-0029	**Intron 01**	22809	315-23123
Exon 02	122	23124–23245	0029-0070	**Intron 02**	2383	23246-25628
Exon 03	123	25629–25751	0070-0111	**Intron 03**	3824	25752-29575
Exon 04	213	29576–29788	0111-0182	**Intron 04**	5630	29789-35418
Exon 05	69	35419–35487	0182-0205	**Intron 05**	2433	35488-37920
Exon 06	117	37921–38037	0205-0244	**Intron 06**	15134	38038-53171
Exon 07	222	53172–53393	0244-0318	**Intron 07**	2643	53394-56036
Exon 08	262	56037–56298	0318-0405	**Intron 08**	284	56299-56582
Exon 09	172	56583–56754	0405-0462	**Intron 09**	4801	56755-61555
Exon 10	94	61556–61649	0463-0494	**Intron 10**	3903	61650-65552
Exon 11	215	65553–65767	0494-0565	**Intron 11**	2914	65768-68681
Exon 12	151	68682–68832	0566-0616	**Intron 12**	5984	68833-74816
Exon 13	210	74817–75026	0616-0686	**Intron 13**	16021	75027-91047
Exon 14	3106	91048–94153	0686-1721	**Intron 14**	21997	94154-116150
Exon 15	154	116151–116304	1721-1772	**Intron 15**	1396	116305-117700
Exon 16	213	117701–117913	1773-1843	**Intron 16**	286	117914-118199
Exon 17	229	118200–118428	1844-1920	**Intron 17**	207	118429-118635
Exon 18	183	118636–118818	1920-1981	**Intron 18**	1738	118819-120556
Exon 19	117	120557–120673	1981-2020	**Intron 19**	608	120674-121281
Exon 20	72	121282–121353	2020-2044	**Intron 20**	1419	121354-122772
Exon 21	86	122773–122858	2044-2072	**Intron 21**	3633	122859-126491
Exon 22	156	126492–126647	2073-2124	**Intron 22**	32849	126648-159496
Exon 23	145	159497–159641	2125-2173	**Intron 23**	1216	159642-160857
Exon 24	149	160858–161006	2173-2222	**Intron 24**	1109	161007-162115
Exon 25	177	162116–162292	2223-2281	**Intron 25**	22679	162293-184971
Exon 26	153	184972–185124	2282-2332	**3' UTR**	1806	185125–186930

Interestingly, but potentially irrelevant to the investigation of the broad variability of FVIII levels observed in non-hemophilic populations, there exists two known non-FVIII transcripts within intron 22 flanking a bi-directional promoter(42, 43). For the first, *F8A*, transcription of its single exon occurs centromeric to telomeric, opposite to that of FVIII(42). The mRNA for the second, *F8B*, includes exons 23 to 26.

Two copies of *F8A* exists telomeric to FVIII, which is of grave importance since crossover (i.e. homologous recombination, see the *Degrees of relation/Identity by decent*

section [1-VI-B] for details) between *F8A* and one of these copies may occur(44). A second intrachromosomal cross-over may occur between a 1 Kb sequence within Intron 1 and a copy of this sequence that is telomeric to FVIII. Both of these events lead to severe forms of hemophilia in males, resulting in pathologic (absent) FVIII:C levels. In female carriers, translation products from *F8A*, *F8B*, or any chromosome suffering such recombination, may be of little direction importance to FVIII function since primary structure, i.e. amino acid, similarity does not necessary lead to higher order similarity, i.e. domainal topology. In any case, these events are sufficiently rare that an investigation of "normal" FVIII levels would be unlikely to sample more than one female carrier. These events likely occur during spermatogenesis exclusively, since the presence of the second X chromosome of females may inhibit intrachromosomal recombination.

As of build 123, the single nucleotide polymorphism (SNP) database of the National Center for Biotechnology Information (NCBI) website, which was formerly dbSNP (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi), contained 391 records. Appendix 1 completely lists these SNPs and other variants with hg17 numbering. It is important to note that this repository is public and that some, if not many, of these SNPs likely are pathological, i.e. borne by hemophilia patients, or deep within introns, i.e. of dubious functional importance. The University of Washington, together with the Fred Hutchinson Cancer Research Center, also performed a scan (see the *Generation of DNA sequence data* section [1-V-C] for details) of FVIII and FIX(42, 45). Appendix 1 also lists the variants found in this scan.

Several investigations have included scans of certain regions or codons in FVIII; however, there have been no nucleotide sequence variations in these regions. The region

of interest were the promoter(40, 46), the 3' UTR(40), and select exonic regions(47-50). Apart from hemophilic mutations, current reports suggest a dearth of nucleotide sequence variations in *F8*. This absence might be due to restrictive scope of those scans, however.

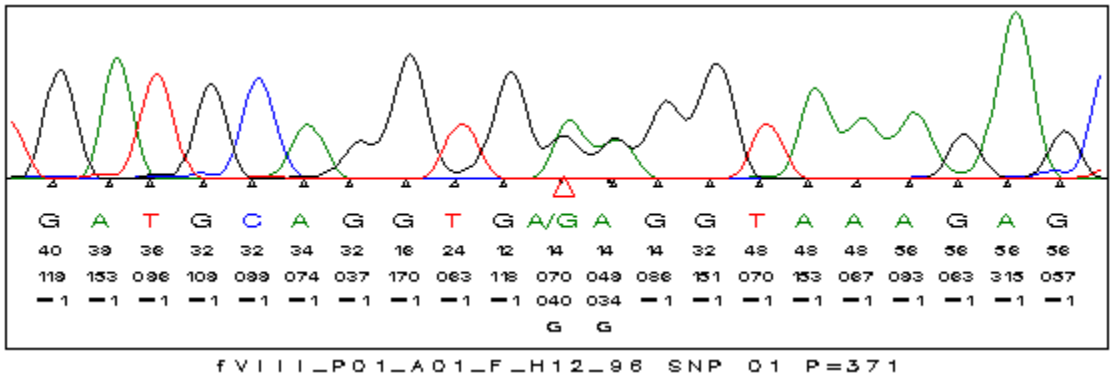
1-V-C. Generating DNA sequence data

DNA polymerase, traveling 3' to 5', synthesizes a 5' to 3' sequence of nucleotides complimentary to its template strand of DNA (A→C, C→G, G→C, and T→A) by adding the new base to the 3' of the elongation product. The two strands of double stranded DNA bind each other through hydrogen bonding of complimentary bases. Heating the double stranded DNA causes separation and thus adequate templates. The polymerase, however, requires a small stretch of double stranded DNA to prime the reaction. Oligonucleotide sequences that anneal to their complimentary sequence within the template satisfy this requirement. The specificity of the interactions means that it is possible to amplify any pinpointed location with genome. The range of the reaction, i.e. length of the product, is limited, however. By pairing two oligonucleotide primers targeted to the different strands within a reasonable number of base pairs, the products also serve as templates. Due to the exponential nature of the amplification of the reaction, these products vastly outnumber the original templates. Mullis et al.(51), recognizing the cyclic potential of the reaction, developed the polymerase chain reaction (PCR). This diverse and robust reaction requires mere micrograms of total DNA and is very amenable to automation on the desktops of the average laboratory(52).

It is the sequence of these products, or amplicons, in which we are interested.

By modifying the bases, we are able to both inhibit the addition of a new base onto the 3' end of the elongating product and to fluorescently tag that base. A reaction involving these modified bases and an excess of unmodified bases, generates products of varying lengths since the incorporation of a modified base terminates the reaction at random stages of elongation. The elution of the products separates them by size and since only the last base contains a tag, we are able to determine the sequence of an approximately 1.2 Kb target region. This reaction results in a linear amplification of the target, therefore a strong signal requires sufficiently amplified PCR product. The elution currently occurs through the 36 cm or 50 cm capillaries of automated sequencer in about two hours. An increase in either the time or the length of the capillary allows for the sequence determination of longer amplicons since the proportion difference of retention factors decreases with increasing size. **Figure 1-1** displays a subregion plot of a typical sequencing run.

Figure 1-1. A chromatogram displaying a subject doubly heterozygous at adjacent nucleotides.



Line 1: called bases; Line 2: quality values; Line 3: relative area of the called base peak; Line 4: relative area of the uncalled base peak, Line 5: Uncalled base, if any

The sequencers process their raw data and usually include programs that call the bases. A widely used stand alone program is phred(53, 54). The authors of this program used sequence data from clones as a training data set for their algorithms. Phred not only calls the bases, it assigns a quality value (QV) that reflects the probability of an error in basecalling. It is important to not, that clones do not contain heterogeneous loci and therefore, a decrease in the QV occurs. Specifically, $QV = -10 \log_{10}(P_e)$, where P_e is the probability of a basecalling error. **Figure 1-1** displays some of the data phred produces and uses to call bases and assign QVs. I have used phred data to determine the genotype of SNP, since phred will not call heterozygous bases, which base 11 (following the A/G) in Figure 1-1 illustrates.

1-VI. Family-based studies

1-VI-A. Introduction

It is clear that unknown genes, environmental factors, and endogenous factors have a causal role in FVIII levels. Until we identify those genes by first mapping them to genetic loci within chromosomes and then discover variations between them, any proposed mechanism, including the elucidation of the effects of environmental or endogenous factors, will be incomplete. Family-based studies, the broad class of observational studies that incorporate relatives, are particularly efficient designs for determining whether familial clustering of a trait is due to a shared environment alone or

a combination of the environments and a genetic effect. One endpoint of these studies is the potential determination of the approximate location of functional genes. Given the vastness of the genome and the need to standardize and automate laboratory procedures, it is unlikely that the variants of interest would be among those under investigation. Thus, this approach would fail if not for the existence of associations between distinct syntenic alleles, i.e. alleles on the same chromosome, that occurs as a result of recombination. Using family-based studies, we examine loci of known location that show an association with the phenotype of interest, knowing that these are not likely in the functional gene, but rather cosegregating with it due to allelic association.

1-VI-B. Identity by descent / Degree of relation

Parents pass a haploid set of autosomes and one sex chromosome to their offspring. It might appear that this would result in a lack of diversity of genotypes, the combinations of alleles at a given locus. Recombination of homologous chromosomes during meiosis, however, adds to the diversity since parents transmit only recombined chromosomes under normal conditions.

If two individuals possess exact copies of the same ancestral locus, then we say that the alleles are **identical by descent (IBD)**. This requires that their respective intervening antecedents, if any, between each of them and a common ancestor that possessed the original locus all replicated the locus with unfailing fidelity and transmitted it as an unrecombined region in each mating of the lineage leading to the individuals.

Two alleles from two individuals may be the same, but not derived from a common ancestor, in which case we say that the alleles are identical by state. Identity by descent implies identity by state, but the converse is not necessary true.

For a given automsomal locus, two individuals can share 0, 1, or 2 alleles IBD. Obviously, the more distantly two individuals are related, i.e. the more collective intervening antecedents between them and the common ancestor, the less likely they are to share alleles IBD. For two individuals to share 2 alleles IBD, then they must have two common ancestors, one in the maternal lineage and one in the paternal lineage.

The limit to the number of common antecedents of two individuals is, in the extreme case of full siblings, the number of individuals in the lineage to the common ancestor of interest. Other relationships, such as the double first cousins that Almasy et al.(55) present in Figure 2 of their paper, emphasize the potential complexity, but the probabilities that alleles at given loci are IBD (IBD probability) sufficient characterize the relationship between two individuals from a genetics stand point(56).

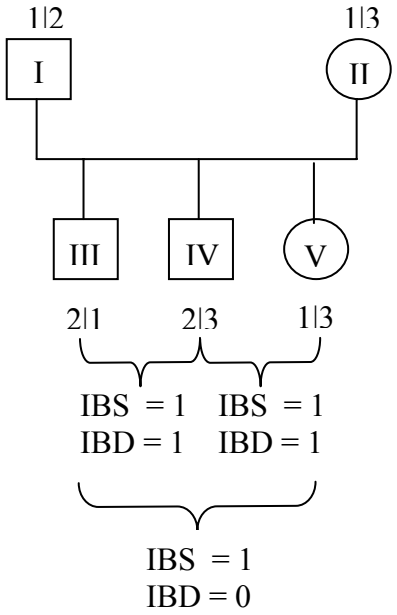
Figure 1-2 presents a possible family with three full siblings that are offspring of parents that are each heterozygous at a given locus. In this case, it is possible to deduce the parental origin of the alleles. In the case of individual V, either the mother or the father could have transmitted allele 1. Her second allele is 3, which only the mother could have transmitted. This means that she obtained her father's allele 1 and, thus, is not IBD for this allele with her brother III, who obtained his allele 1 from his mother, which we know since his allele 2 must have come from his father. Thus individuals III and V share 1 allele IBS, but 0 alleles IBD. The other two pairs, III and IV and IV and V, each share one allele IBD. In the former pair, they each obtained the 2 allele from

their father whereas in the latter pair, they each obtained the 3 allele from their mother. In both cases, the second alleles are not IBS, and, thus, cannot be IBD.

1-VI-C. Violations of Mendelian inheritance

Typically, genome scans utilize microsatellite markers. As the genome has an abundance of bulk repetitive material, these markers are usually di- or tri-nucleotide

Figure 1-2. Hypothetical family of three full siblings demonstrating IBD and IBS status at an autosomal locus.



repeats resulting in polymorphic lengths. The marker DXS1073 is an example of a dinucleotide repeat microsatellite in which the repeating unit is $(TG)_n$, where the n indicates the number of times this unit appears. The X indicates that this marker is on the X chromosome.

These markers have high heterozygosity, i.e. the probability that the genotype is not homozygous. The heterozygosity affects the information content of the marker, since it is easier to infer IBD status without extended pedigrees, which may not be available. Despite this, more genome scans are taking advantage of the advances in the mass genotyping of known SNPs, which can have a maximum of four alleles, but three allele SNPs are rare. The ease with which a laboratory can determine the genotypes of SNP compared to the length polymorphic microsatellites is one of the leading reason for this change.

In either case, genotyping errors occur. The easiest to detect involve violations of Mendelian inheritance. For example, a 5|6 offspring of 5|5 and 4|5 parents most likely resulted from a genotyping error since the probability of a de novo mutation is extremely low. The program INFER of the PEDSYS detects such straight forward transmission errors in three successive phases(57). In the first, it examines each offspring-parent triplet inferring unknown genotypes and recording errors according to the Mendelian laws of transmission. In the second phases it repeats the process including the potentially inferred data. The final phase extends the triplet to offspring-grand parental triplets.

1-VI-D. Heritability

An important first step in identifying a gene that may affect a phenotype is to determine whether phenotype segregates within families. Further, segregation analysis may also elucidate the mode of inheritance with the shortcoming that most diseases have complex, i.e. undeciphered, modes. To formalize this concept, geneticists use the term heritability to express the proportion of total phenotypic variance due to additive genetic effects. This section explains **heritability in the narrow sense** within the framework of a quantitative phenotype and relies heavily on Falconer's Introduction to Quantitative Genetics(58).

In general, we consider an outcome of interest, the phenotype, which is determined by a genotype. The environment in which the individual exists or develops may cause a deviation from the value that the genotype confers. Symbolically,

$$\mathbf{P = G + E}$$

In the population as a whole, E[E]=0, so E[P]=E[G + E]=E[G] + E[E]=E[G].

Within a biallelic framework, we assign the resulting phenotypic values to the genotypes: $A_1A_1 = a$, $A_1A_2 = d$, $A_2A_2 = -a$. This allows for allelic interaction, or dominance. In the absence of dominance $A_1A_2 = 0$, i.e. the midpoint between the two phenotypic values resulting from the two homozygous genotypes. If we assume that $f(A_1)=p$ and $f(A_2)=(1-p)=q$ then we have **Table 1-5**.

Table 1-5. Contribution of biallelic genotypes to the population phenotypic mean

Genotype	Phenotypic Value	f(Genotype)	Contribution to population phenotypic mean
A_1A_1	a	p^2	ap^2
A_1A2	d	2pq	2dpq
A_2A_2	-a	q^2	$-aq^2$

Thus, the population phenotypic mean:

$$M = ap^2 + 2dpq - aq^2 = a(p^2 - q^2) + 2dpq = a(p-q)(p+q) + 2dpq = a(p-q) + 2dpq \quad \textit{Eq 1}$$

The ***average effect*** of an allele is the mean difference between the population phenotypic mean (M) and the mean of phenotypic values resulting when that allele is united at random with one of the two alleles in that population. For a given population under random mating, an allele passed by the parent under consideration will be united with the A_1 allele with probability p and with the A_2 allele with probability q, i.e. $f(A_1) = p$ and $f(A_2) = q$. Clearly, the resulting genotype is a function of the allele frequencies, which govern the identities of both the uniting allele and the allele donated by the parent under consideration. The population phenotypic mean is also a property of the population, i.e. described completely by the phenotypic values and the allele frequencies. **Table 1-6** presents the mean phenotypic value of offspring considering each allele as the focus of the average effect measure.

Table 1-6. Average effect and mean phenotypic value by gametic allele

Allele in the gamete produced by the parent under consideration	**Genotypes that results under random mating phenotypic value**			**Resulting mean phenotypic value**
	A_1A_1 a	A_1A_2 d	A_2A_2 -a	
A_1	p	q	0	ap + qd
A_2	0	p	q	dp – aq

To explain, when the allele in the donated gamete is A_1, the gamete with which it unites carries the A_1 allele with probability p and A_2 with probability q, i.e. the frequency at which these alleles are present in the population. This is a conditional probability that this specific genotype results from a random union given the identity of the donated allele, i.e. P(offspring = A_1A_1 | allele = A_1) = p, the probability of randomly encountering the A_1 allele in a random gamete. Similarly, P(A_1A_2 | allele = A_1) = q, the probability of randomly encountering the A_2 allele in a random gamete. These are the only two nonzero probability genotypes given that the gamete under consideration carries the A_1 allele.

The average affect is then the deviation of the resulting mean phenotypic value (column 4, **Table 1-6**) from the population M (*Eq 1*):

$$\alpha_1 = ap + qd - a(p - q) - 2dpq$$

$$= q[a + d(1 - 2p)]$$

$$= q[a + d(1 - p - p)]$$

$$= q[a + d(q - p)] \qquad \textit{Eq 2}$$

$$\alpha_2 = dp - aq - a(p - q) - 2dpq$$

$$= -p[a + d(1 - 2q)]$$

$$= -p[a + d(1 - q - q)]$$

$$= -p[a + d(p - q)] \qquad \textit{Eq 3}$$

The difference between the average effects has an interesting interpretation. The ***average effect of the gene substitution, α***, is the change in the mean phenotypic value that would occur if an allele selected at random were replaced with the other allele. For instance, consider replacing A_2 with A_1. A_1A_2 would result in A_1A_1 and the phenotypic value would change from d to a, thus $\Delta = a - d$. A_2A_2 would result in A_1A_2 (or A_2A_1, which has the same result) and the phenotypic value would change from –a to d, thus $\Delta = d - (-a) = d + a$.

To determine the frequency with which these changes occur, and thus the change in the mean phenotypic value, consider selecting A_2 at random from the population. You could select from A_2A_2 individuals, which offer two A_2 alleles, or you could select A_1A_2 individuals, which offer one A_2 allele:

$$\underline{A_2}A_2 \quad A_2\underline{A_2} \quad A_1\underline{A_2}$$

$A_1\underline{A_2} \rightarrow A_1\underline{A_1}$: $2pq / [q^2 + q^2 + 2pq] = 2pq / [2q^2 + 2pq] = pq / q[q + p] = p$

$\underline{A_2}A_2 \rightarrow \underline{A_1}A_2$: $(q^2 + q^2) / [q^2 + q^2 + 2pq] = 2q^2 / [2q^2 + 2pq] = q^2 / q[q + p] = q$

Thus,

$$\alpha = p(a-d) + q(d+a) = a(p+q) + d(q-p) = a + d(q-p) \qquad \textit{Eq 4}$$

Noting *Eq 2* and *Eq 3*, we have

$$\alpha = \alpha_1 - \alpha_2$$

$$\alpha_1 = q\alpha$$

$$\alpha_2 = -p\alpha$$

The concept of average effects embodies the haplotypic nature of sexual reproduction: meiosis deconstructs diploid genotypes and the separated genes (alleles) unite to form genotypes anew under the constraints of the allele frequencies. Thus, the mean value of the offspring of an individual which mates at random is a property of a population, as are the allele frequencies and the phenotypic effect of a genotype. **Table 1-7** presents the mean value of the progeny for each of the three genotypes possible at the biallelic site.

The deviation of the mean value of the progeny from the population mean is half of the ***breeding value*** of the individual. The individual only provides half of the genes to its progeny, thus the need to double the deviation. **Table 1-8** presents the breeding values for this individual in the population.

Table 1-7. Mean phenotypic value of the progeny by biallelic genotype

Genotype of interest	Proportion of Gamete	Uniting Gamete	Probability	Resultant Genotype	Phenotypic Value	Contribution
A_1A_1						
Gamete						
A_1	0.5	A_1	p	A_1A_1	a	0.5ap
		A_2	q	A_1A_2	d	0.5dq
A_1	0.5	A_1	p	A_1A_1	a	0.5ap
		A_2	q	A_1A_2	d	0.5dq
					Progeny Mean	ap + dq
A_1A_2						
Gamete						
A_1	0.5	A_1	p	A_1A_1	a	0.5ap
		A_2	q	A_1A_2	d	0.5dq
A_2	0.5	A_1	p	A_1A_2	d	0.5dp
		A_2	q	A_2A_2	-a	-0.5aq
					Progeny Mean	0.5[a(p – q) + d]
A_2A_2						
Gamete						
A_2	0.5	A_1	p	A_1A_2	d	0.5dp
		A_2	q	A_2A_2	-a	-0.5aq
A_2	0.5	A_1	p	A_1A_2	d	0.5dp
		A_2	q	A_2A_2	-a	-0.5aq
					Progeny Mean	dp – aq

The breeding value of **A_1A_1** is:

$$2[ap + dq - M] = 2[ap + dq - a(p - q) - 2dpq] = 2q[a + d - 2dp] = 2q[a + d(1 - 2p)] =$$

$$2q[a + d(q + p - 2p)] = 2q[a + d(q - p)] = 2q\alpha = 2\alpha_1$$

The breeding value of A1A2 is:

$$2(0.5[a(p-q)+d]-M) = a(p-q)+d-2a(p-q)-4dpq = -a(p-q)+d(1-4pq)$$

$$= -a(p-q)+d[1-4p(1-p)] = -a(p-q)+d[1-4p-4p^2]$$

$$= -a(p-q)+d(1-2p)^2 = -a(p-q)+d(p+q-2p)^2$$

$$= -a(p-q)+d(q-p)^2 = (q-p)[a+d(q-p)] = (q-p)\alpha$$

Note: $(q-p)\alpha = (q-p)[a+d(q-p)] = q[a+d(q-p)] - p[a+d(q-p)] = \alpha_1 - \alpha_2$

The breeding value of $\mathbf{A_2A_2}$ is:

$$2(dp-aq-M) = 2[dp-aq-a(p-q)-2dpq] = 2p[-a+d-2dq]$$

$$= 2p[-a+d(1-2q)] = 2p[-a+d(p+q-2q)] = 2p[-a+d(p-q)]$$

$$= -2p[a+d(q-p)] = -2p\alpha = 2\alpha_2$$

Table 1-8. Breeding values by biallelic genotype

Genotype	Breeding Value
A_1A_1	$2q\alpha = 2\alpha_1$
A_1A_2	$(q-p)\alpha = \alpha_1 - \alpha_2$
A_2A_2	$-2p\alpha = 2\alpha_2$

For a population in HWE, the mean breeding value is zero:

$$p^2(2q\alpha) + 2pq(q-p)\alpha + q^2(-2p\alpha) = 2p^2q\alpha + 2pq^2\alpha - 2p^2q\alpha - 2pq^2\alpha = 0 \quad Eq\ 5$$

The breeding value, represented as A, is a component of the genetic effects. The other component is the dominance effect, D. If the phenotype of interest derives from a multilocus effect, then a third component may be the interlocus interaction, I. Therefore, we have:

$$P = A + D + I + E, \text{ where } G = A + D + I$$

If we assume 1) that the genotypic values and the environmental deviation are uncorrelated, and 2) an absence interaction between the environment and the genotype, then the variance of the phenotype is straightfoward:

$$V_T = V_A + V_D + V_I + V_E$$

The relative importance of any component is signified by the proportion of total variance for which it accounts:

$$V_X / V_P, \text{ where } X = G, A, D, I, E$$

Overall, V_G / V_P is best described as the *clonal repeatibility* of the trait. Of notable importance is the ratio V_A / V_P, which is the heritability (in the narrow sense) of the trait. V_A is also known as the *additive variance*. Observing the degree of resemblance between relatives enables the estimation of the additive variance.

To calculate V_A one needs the breeding values, their frequency, and two basic statistical results. **Table 1-8** provides the breeding values for each genotype in terms of α. Basic statistical theory provides the following results:

$$Var[X] = E[X^2] - (E[X])^2$$

Noting that E[X]=0 (see Eq 5), we have

$$V_A = p^2(2q\alpha)^2 + 2pq[(p - q)\,\alpha]^2 + q^2(-2p\alpha)^2$$

$$= 4p^2q^2\alpha^2 + 2pq(p^2 - 2pq + q^2)\,\alpha^2 + 4\,p^2q^2\alpha^2$$

$$= 2pq\alpha^2\,(2pq + p^2 - 2pq + q^2 + 2\,pq)$$

$$= 2pq\alpha^2\,(p + q)^2$$

$$= 2pq[a + d(q-p)]^2$$

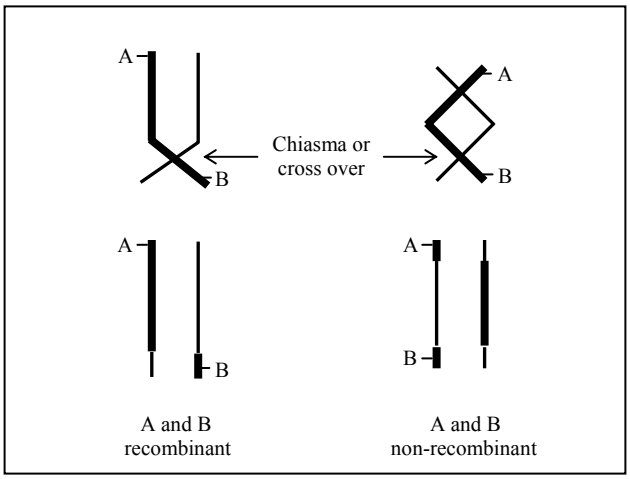
Of the several items to note, observation of the degree of resemblance between relatives allows ready estimation of V_A, unlike the other components of variance. It is not uncommon to group the components, one of which is the observational components. Partitioning the phenotypic variance, namely into components corresponding to the grouping of individuals into families (and classes), allows for the measurement of the degree of resemblance.

1-VI-E. Linkage Analysis

As mentioned in section **V-A** (Identity by descent / Degree of relation**),** recombination disrupts the cosegregation of some syntenic alleles. If there are an odd number of cross overs (chiasmata) between two syntenic loci, then they are recombinant, otherwise they are non-recombinant (see **Figure 1-3**). The proportion of the recombinant

chromosomes (genotypes) out of the total number of chromosomes is the recombinant fraction, θ. The upper limit to E[θ] is 0.5, since the probability of an even number of

Figure 1-3. Odd and even number of cross overs during meiosis.



cross overs is equal to the probability of an odd number of events over a large syntenic region. For a two allele haplotype, i.e. the series of alleles on the same chromosome, there are four possible gametes. **Table 1-9** illustrates the parental haplotypes (represented with a symbolic chromosome as a line or a double line), the resulting set of gametes and the probability of the occurrence.

If the genetic distance is sufficiently large, then recombination fraction is 0.5 and each gamete has probability of 1/4 of occurring. This probability is equivalent to

Table 1-9. Parental haplotypes and possible gametes

Parental haplotype	Resulting Gamete	Type	Probability
M || m D || d	MD	Non-recombinant	(1- θ) / 2
	md	Non-recombinant	(1- θ) / 2
	Md	Recombinant	θ / 2
	mD	Recombinant	θ / 2

assuming that the two loci are on separated chromosome and, therefore, segregation independently.

The number of recombination events *r* occurring among *n* such gametes follows a binomial distribution with probability θ. The corresponding likelihood of observing *r* recombination events takes the form:

$$L(\theta \mid n , r) = \binom{n}{r} \theta^{r} (1 - \theta)^{n-r}$$

under the null hypothesis of no linkage (θ = 0.5), we define the likelihood ratio as:

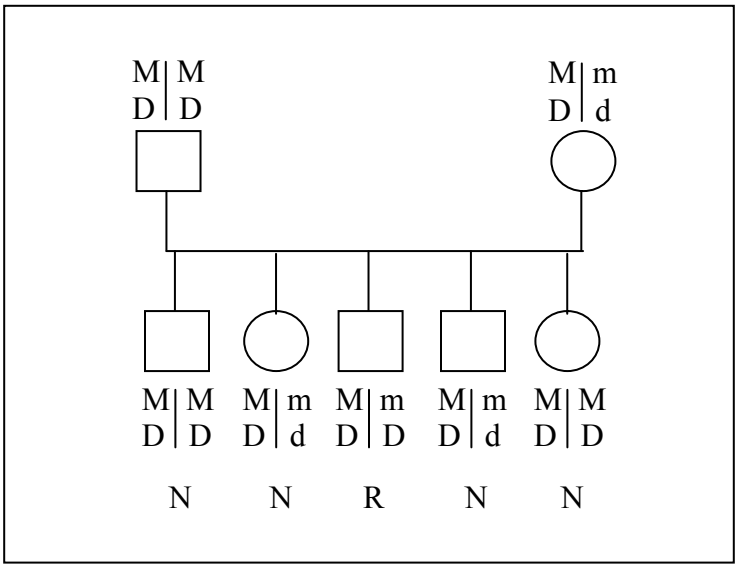
$$LRT(\theta) = L(\theta \mid n , r) / L(\theta = 0.5 \mid n , r)$$

Historically, the base 10 logarithm of the LRT(θ), represented by **Z**, is known as the ***lod score*** (lods). A value of Z = 3.0 corresponds to p = 0.0001, however, we typically require greater stringency for the level of significance due to multiple comparisons across the genome. Alternatively, the strength of linkage evidence can be assessed by calculating a likelihood-ratio statistic, defined as twice the natural log of LR (θ), which is distributed

as a 1:1 mixture of χ^2_1 distribution and a point mass at 0, under the null hypothesis $\theta = 0.5$(59).

To illustrate two point linkage analysis consider a marker locus of known location with alleles M|m, disease or non-disease status as D or d, respectively, and the pedigree of known phase in **Figure 1-4**.

Figure 1-4. A hypothetical pedigree in which the disease cosegregates with the marker locus.



The value that maximizes the likelihood is $\theta_{MLE} = 0.2$ with the resulting maximum lod score:

$$Z = \log_{10}(LR (\theta = 0.2)) = (0.2)^1(0.8)^4 / (0.5)^1(0.5)^4 = 2.62$$

With the value of Z = 3.0 as the typical minimum value to offer enough support to reject the null hypothesis of no linkage, we might, at best, say that the results from this family *suggest* linkage. Assuming, however, that the recombination fraction is the same across separate families, then we may treat them as independent trials. Therefore, we may multiply the resulting likelihoods and, thus, add the resulting lod scores to build evidence.

An improvement to the two point linkage analysis involves using more markers that flank the disease locus. In multipoint linkage analysis, the markers are all used to simultaneous inform both the recombination parameter and the order of the markers. Multipoint linkage analysis takes advantage of information that (serial) two point analyses would ignore when the count of recombinants is indistinct, which frequently happens when phase is not known with certainty. Unfortunately, in either analysis, even highly significant results pose the problem that the region surrounding the marker is still immensely vast, especially on the scale of genes and their variants. Further, as described, this test lacks information due to the dichotomous nature of the outcome. Fortunately, investigators have developed methods that apply to quantitative outcomes, whether they are traits themselves or endophenotypes underlying the disease.

With the increase in computational power, investigators now have the ability to perform analyses of sufficient complexity to use quantitative outcomes to search for a **quantitative trait locus** (QTL) within hours or weeks. The software may distribute the process over many CPUs or, in the case of simpler analyses, a desktop computer.

A particularly robust and popular tool is the variance-components (VC) based linkage analysis(60). Almasy and Blangero have extended the method to general pedigrees and note that VC methods also provide a reasonable estimate of the magnitude

of the effect of the locus(61). Briefly, this method specifies the expected genetic covariances (see section 1-VI-D [Heritability]) between arbitrary relatives as a function of the IBD relationships (see section 1-VI-B [Identity by descent / Degree of relation]) at a QTL. Using two models that differ only by the inclusion of the QTL, one can test whether the additive genetic variance due to the QTL is different from zero by comparing the likelihoods of the two models.

Hopper and Mathews have extended multivariate normal models to pedigrees allowing the determination of genotype-specific means while accounting for non-independence of the data due to the family structures(62). Incorporation of these genotype-specific means into a linkage analysis potentially provides useful information and potentially more power. If the QTL is the only functional variant in the region, then the result should be that the additive genetic variance due to that QTL will be reduce to zero, otherwise there will be more power to detect the other functional variant(63).

1-VII. Determinants of FVIII:C Levels

1-VII-A. Introduction

The two broad aspects of the FVIII protein that may affect activity levels are function and concentration. The range of mechanisms that may affect activity is

therefore vast. Transcription rate, translation rate, mRNA stability, post-translational modification rate, secretion rate, stability of FVIII, activation rate, stability of FVIIIa, and clearance rate all affect the amount of circulating protein. Differences in performance functionality may occur if there are two or more variant protein molecules or if reaction conditions vary. Differences in the latter may directly change the FVIII or FVIIIa protein, such as altering its conformation, blocking the binding sites, or affecting the amount of the B-domain. However, it is possible that these differences may not affect FVIII or FVIIIa, but instead affect the components with which they interact, such as increases the number of platelet binding sites. With the exception of the limited *in vitro* studies of Saenko, it is still not clear which of these aspects change with different exposures. In most circumstances, investigators only measured FVIII:C levels leaving the amount of protein (FVIII:Ag levels) open to conjecture.

1-VII-B. Environmental

Of the several environmental factors investigated to date, all are associated with increased FVIII:C levels with the exception of those that may generate inhibitors, i.e. anti-bodies that suppress or inhibit FVIII cofactor function. *In vivo* evidence from human populations has yet to reveal what the nature of the induced changes may be, i.e. increased activation, conformation changes, increased expression, increased secretion, decreased clearance, etc.

To standardize measurement of blood chemistry, investigators or health professions phlebotomize patients in the morning after a fast. The MAS protocol of

Balleisen et al., for example, required that "[b]lood sampling [be] performed between 7 and 9 a.m. in [a] fasting state"(64). Three potential factors influence the decision to implement this protocol: a circadian rhythm, an influence of diet, and the effects of exertion or stress.

Haus et al. found a statistically significant circadian rhythm using a cosinor analysis, with an acrophase at 08:20, in an investigation of 5 men and 5 women, whose mean age was 31 (SD = 11) years(65). The study design of Haus et al. was a Latin Square of size three in which they sampled blood at two time points that were 12 hours apart per investigational day. Iversen et al. included 8 control subjects, age 33 $\pm$ 1.5 years, in their study of circadian variations over a single 24-hour period in tetraplegic patients(66). Iversen et al. collected blood every 2 hours while the control subjects mostly rested or performed low-activity tasks. Apparently, there is a sole peak at 8 am of perhaps modest size but the Iversen et al. do not state whether this is statistically significant, though they stated that in the six tetraplegic patients FVIII:C "remained virtually constant and at a lower level ($P < 0.05$)." van Diest et al., however, found no diurnal variations in FVIII:C levels among either of patients with vital exhaustion (n = 29) or the controls (n = 30), mean age 55 $\pm$ 5 years(67).

Salomaa et al. tested the effect of three test meals given one week apart on 10 men from their research staff who did not smoke, were not taking medication, whose mean age and BMI was 40.8 $\pm$ 8.4 and 23.7 $\pm$ 2.7, respectively(68). After a twelve-hour fast, the men had a blood sample drawn at 8 am, ate the test meals within 10 minutes, had blood drawn every 2 hours for a total of six post-meal draws, ate a light fat-free lunch, and pursued their usual work activities without strenuous physical activities. The men

received the meals in random order. The three meals were a fat-free meal, a saturated fat meal, and an unsaturated fat meal. There seems to have been no effect of the meals on FVIII:C levels.

Intense or prolonged exertion, such as exercise increase FVIII:C levels, by as much as 200%-400%(69). Kopitsky et al. suggest that this is not a result of a release of VWF/FVIII but rather a result of in vivo activation of FVIII by thrombin(70). van den Burg et al. showed that FVIII:C continued to increase after their subjects achieved their maximal heart rate through their last timepoint in recovery(71). Jilma et al. have demonstrated that partial blockade of nitric oxide synthase attenuates the exercise-induced rise in FVIII:C levels(72). Jern et al. reported that FVIII:C levels significantly increased from 125% to 217% after 20 minutes of mental stress in 22 young, non-smoking males(73). Consistent with this, Chang et al. reported that FVIII:C levels were significantly related to job strain in Korean male workers, using three related multiple regression models(74).

It appears that, at best, demonstration of either a circadian rhythm or postprandial effect requires further research. From these studies, the conclusion might be that this standardization, while providing obvious benefit to the measurement of other values that investigators frequently make on the same sample, such as glucose or cholesterol, might not affect FVIII:C levels. Thus, from a perspective of an investigation into FVIII:C levels, the implementation of such a protocol might not introduce variability.

In the case of diet, an acute postprandial effect would be obviously important, while dubious, but there may still be a more prolonged or chronic effect. Allman-Farinelli et al. compared the effects of two low-fat diets with different ratios of the acids

α-linolenic and linolenic(75). Following a two week run-in diet, the 30 men aged 18 to 35 years began six week diets to which the investigators randomized them. Allman-Farinelli et al. reports no change in FVIII:C levels. Elkeles et al. studied the effects of a carbohydrate-reduced, fat modified diet on eleven men whose mean age was 50 years. The effect of this treatment significantly reduced weight, blood glucose, triglyceride levels, cholesterol levels, and FVIII:C, but not FVIII:Ag, levels. Elkeles et al. had no control group and they made no adjustment for covariates.

In an early report, Meade et al. investigated the effect of oral contraceptive use on FVIII:C levels in 234 women enrolled in a prospective study of thrombosis and a subset of 20 of these women with age-matched controls(76). Meade et al. reported no associations. Kadir et al. investigated the effect of oral contraceptive use on FVIII:C levels in a cross-sectional (N = 123) and longitudinal study (N = 39)(77). Kadir et al. failed to find differences in the cross-sectional study, but report a mild cyclical effect in the longitudinal study among pill users with a reported range of 103 to 113 U/dL. It is important to note that the study is small and that Kadir et al. may have been subject to misclassification of the stages of the cycles, i.e. sampling time variability may have caused overlap between the days. Balleisen et al. also failed to find a significant effect of oral contraceptive use from a multiple regression analysis of approximately 1,300, 26% of whom used oral contraceptives in the MAS study(64). Lowe et al. investigated the effect of oral contraceptive use on FVIII:C levels in a third Glasgow MONICA survey in which they had obtained measurements on 146 women, 56 of whom were taking oral contraceptives, who were between 25-34 years of age(78). The crude results of Lowe et

al. suggest that oral contraceptive use significantly increased FVIII:C levels when compared to non-users, 135 versus 112, respectively.

There has been intense interest in thrombotic events that appear to be related to air travel. Several aspects to air travel may increase the risk of thrombosis. First, the seats confine the passengers limiting movement which may lead to stasis in the extremities. Second, the passenger may not be adequately hydrated. Third, the partial pressure of carbon dioxide increases while that of oxygen decreases. Schobersberger et al. have reported an interesting effect on FVIII:C levels on 20 patients who embarked on an approximate 9 hour trans-Atlantic flight from Vienna to Washington, D.C. with a two-night stay before the return trip(79). The passengers all traveled from and to Innsbruck via an hour plane trip and consumed at most the equivalent of a half liter beer during the flights. The were six sampling points: T1, Baseline 48 hours before the trip; T2, between 5-8 hours into the departing flight; T3, between 5-8 hours into the returning flight; T4, post-flight arrival into Vienna; T5, the morning of following day; and T6, the morning of third day after arrival. The median FVIII:C levels at these timepoints were 108%, 113%, 106%, 126%, 160%, and 118%, respectively. With the exception of the T3 value (106%), all differed significantly from the T1 baseline value (108%). Importantly, there was a significant decrease in VWF only between T1 and T2, 108% and 98%, respectively. The value of VWF at T3 (99%), the only timepoint at which the value of FVIII:C did not differ significantly from T1, likely represents a true change, but the sample size likely obscured it. Therefore, had the analysis accounted for VWF, then there would likely have been a significant difference at every timepoint for FVIII:C.

In the ARIC study, Conlan et al. reported significantly lower sex-, age-, and race-adjusted FVIII:C levels in current smokers versus former smokers and non-smokers, 126, 131, and 130, respectively(80). However, among smokers, the sex-, age-, and race-adjusted FVIII:C levels did not differ among the categories less than 15, 15 to 24, and more than 24 cigarettes per day. The effect of dichotomous current smoking status remained significant in a multiple regression model(80). Balleisen et al. reported that the number of cigarettes did not influence FVIII:C levels in the MAS study(64). Chang et al. also reported a significant effect of smoking on FVIII:C levels in multiple regression analyses(74). Geffken et al. reported that persons in the CHS study who never smoked had higher FVIII:C levels than both former smokers and current smokers, who had the lowest levels(81). It is not clear whether the results of Geffken et al. are crude or adjusted, however. In a study of 151 singleton newborns, van der Salm et al. found no difference in the FVIII:C levels obtained from cord blood of infants whose mothers smoked at least 15 cigarettes per day compared to the infants whose mothers did not smoke(82). van der Salm et al. did find that the birth weight was significantly lower in the former group, as expected. To explain the apparent lack of a difference, van der Salm et al. hypothesized that macrophages that contact smoke in the lungs may produce interluekin-6, which may stimulate FVIII:C levels, a known acute phase reactant. Green et al. investigated FVIII:C levels from the 1,724 CARDIA study that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 sample(83). Green et al. reported no correlation between FVIII:C levels and smoking status; further, the mean FVIII:C level of smokers was significantly higher than that of non-smokers only for white males.

Korkmaz et al. reported the FVIII:Ag levels were significantly elevated in a group of 49 patients with familial Mediterranean fever as compared to both a control group and the same patients at least two weeks after the attack(84). Reitsma et al. studied the effect of endotoxin on FVIII:Ag levels in 8 male volunteers 19-29 years of age(85). Compared to baseline measurements, the antigen levels of FVIII and VWF were increased up to a maximum of four times the pretest levels after endotoxin administration. After 24 hours, the FVIII:Ag levels were still three time the pretest levels, however, FVIII:C levels apparently return to normal(85).

Gibbs et al. have investigated the effect of elective abdominal aortic bypass surgery in 16 men and 3 women whose mean age was 68 (SD 8) years on the coagulation system(86). They reported that FVIII:C levels were significantly increased by 39%, 95%, and 76% on post-operative days 2, 4, and 6 compared to pre-operative day 0.

Table 1-10 summarizes these findings. Evidence of an effect of smoking and oral contraceptive use is weak, despite anecdotal expectations. The support for either a circadian or postprandial effect, both of which could bias measurements of basal FVIII:C levels if measured at different times of day or levels of fasting, is dubious. The effect of exertion or stress and surgery, perhaps through a sympathoadrenal response, is strong. Air travel and the many possible factors related to cabin conditions also seem to increase FVIII:C levels.

Table 1-10. Environment factors investigated for their potential effects on FVIII levels

Exposure	Effect on FVIII:C levels	Effect on FVIII:Ag levels	Investigators
Circadian Variation	No effect, except Haus et al.	Unknown	Iversen et al.(66), van Diest et al.(67), Haus et al. (65)
Eating	No effect	Unknown	Salomaa et al.(68)
Exercise/Exertion	Increase	Unknown	Bourey and Santoro (69)
Mental Stress	Increase	Unknown	Jern et al.(73), Chang et al.(74)
3 Low-fat diet	No effect	Unknown	Allman-Farinelli et al.(75)
Carbohydrate-reduced, fat modified diet	Decrease	No effect	Elkeles et al.(87)
Oral Contraceptives	No effect	Unknown	Meade et al.(76), Kadir et al.(77), Balleisen et al.(64), Lowe et al.(78)
Air Travel	Increase	Unknown	Schobersberger et al.(79)
Smoking	Decrease, weak effect, see text	Unknown	Conlan et al.(80), Geffken et al.(81), Green et al.(83)
Elective abdominal aortic bypass surgery	Increase	Unknown	Gibbs et al.(86)

1-VII-C. Endogenous

The distinction between an endogenous and environment factor is subtle. One distinguishing aspect will be the internal nature of an endogenous factor, even if (external) environmental factors strongly influence it. Whether it is a component of the immediate reaction or a character of the body, the investigator will be unable with ease to create a separation or distinction between the person and the endogenous factor, if at all.

AGE

Conlan et al. reported on FVIII:C levels in the ARIC study(80). There were 5,022 men and 6,363 women, between 45 and 64 years of age, with measured covariates and FVIII:C levels. In a multiple linear regression, Conlan et al. found that for each 5-year increase in age, mean FVIII:C levels significantly increased 5.0 and 4.7 IU/dL in men and women, respectively.

Tracy et al. reported on FVIII:C levels in the CHS study(88). There were 2,221 men and 2,918 women, 65 years of age and older, with measured FVIII:C levels. In a linear regression, Tracy et al. found that for each 10-year increase in age, FVIII:C levels significantly increased 7.3 and 10.9 IU/dL, in men and women, respectively.

Balleisen et al. reported on FVIII:C levels in the MAS study(64). They initially recruited and reported on 2,880 men and 1,306 women, who were between 16 and 74 years of age. The mean age was 40 years in men and 37 years in women. In a multiple linear regression, Balleisen et al. found that for each 1-year increase in age, FVIII:C levels significantly increased 0.24 and 0.28 IU/dL, in men and women, respectively. The plot of FVIII:C levels versus age, however, clearly shows a more quadratic relationship.

Lowe et al. reported on FVIII:C levels in the third Glasgow MONICA Survey(78). They had FVIII:C level measurements on 714 men and 789 women, who were between 25 and 74 years of age. Lowe et al. reported that FVIII:C levels increased with age and found that FVIII:C levels significantly correlated with age with $\rho = 0.28$ and $\rho = 0.36$, in men and women, respectively.

Lindeberg et al. included FVIII:C levels in the Kitava study(89). The hemostatic component included 76 Kitavan males, 33 Kitavan females, 66 Swedish males, and 23

Swedish females. The mean age (SD) of the Kitavan males and females was 49 (18) and 53 (15), respectively. The mean FVIII:C levels (SD) of the Kitavan males and females was 132 IU/dL (43) and 122(33), respectively. Lindeberg et al. reported that FVIII:C levels significantly correlated with age with $\rho = 0.39$ and $\rho = 0.40$, in men and women, respectively.

In the 1,724 subjects that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 CARDIA sample study, Green et al. found no correlation between age and FVIII:C levels among the race-sex stratifications(83). Among the studies to examine age and FVIII:C levels, the CARDIA study had the youngest populations. In the CDC Menorhaggia study, Miller et al. found no effect of age(90). The 18-45 year age range of subjects in the CDC Menorrhagia study was similar to that of the CARDIA study.

Yarnell et al. reported on FVIII:C levels from the first re-examination of in the Caerphilly study(91). They obtained a blood sample from 2,223 men between 49 and 65 years of age. Yarnell et al. reported that FVIII:C levels significantly different in the three age groups 49-54, 55-60, and 61-65, which had geometric mean levels of 90.2 IU/dL, 95.4 IU/dL, and 98.2 IU/dL, respectively.

FVIII:C levels clearly increase with age. It may be that this effect is more extreme in the later years of life. Notably, there it may not be present in young to middle aged adults. The activity levels of FIX, the enzyme for which FVIII is a cofactor, also increase with age(92). In this case, the hormonal influences on elements of the promoter may play a role since a form of hemophilia B resolves at puberty(93). In addition, Kurachi et al have identified elements in the promoter of mice that are associated with

age-related increases(94). Whether this mechanism might be among those acting on FVIII:C levels is unknown. In addition, these analyses failed to consider two factors to be discussed below, von Willebrand Factor and ABO blood group, which have considerable effect on FVIII:C levels. **Table 1-11** summarizes the effect of age on FVIII:C levels in these studies.

Table 1-11. The effect of age on mean FVIII:C levels

Increment	Result	Source	Investigators
5 years	5.0 and 4.7 IU/dL, in men and women respectively	ARIC	Conlan et al.(80)
10 years	7.3 and 10.9 IU/dL, in men and women respectively	CHS*	Tracy et al.(88)
1 year	0.24 and 0.28 IU/dL, in men and women respectively	MAS	Balleisen et al.(64)
N/A	$\rho = 0.28$ and $\rho = 0.36$, in men and women, respectively	3rd Glasgow Monica Survey*	Lowe et al.(78)
N/A	$\rho = 0.39$ and $\rho = 0.40$, in men and women, respectively	Kitava*	Lindeberg et al.(95)
N/A	Increased with age	Caerphilly*	Yarnell et al.(91)
N/A	No correlation	CARDIA*	Green et al.(83)
N/A	No statistical difference	CDC Menorrhagia	Miller et al.(90)

*Crude or unadjusted results

SEX

In the ARIC study, Conlan et al. reported that among the among the age groups 45-54 and 55-64, both black women and white women had significantly higher FVIII:C levels than black men and white men, respectively(80). **Table 1-12** presents stratified results. Conlan et al. performed separate multiple linear regression analyses for each sex, however.

Table 1-12. Mean FVIII:C levels (SD, IU/dL) by race, sex, and age from ARIC(80)

Age	Black men (n = 1,149)	Black women (n = 1,791)	White men (n = 4,048)	White women (n = 4,757)
45-54	131 (36)	140 (43)	116 (29)	121 (32)
55-64	145 (41)	151 (41)	125 (32)	131 (33)

In the CHS study, Tracy et al. reported that white women (N = 2,775) had significantly higher FVIII:C levels than white men (N = 2,125)(88). There were 96 black men and 143 black women in the analyses, but their levels did not differ. Tracy et al. stratified by sex and the age groups 65-69, 70-74, 75-79, 80-84, and 85 and older. They report that "in each of the age groups examined, women had higher values on average than did men" but it is ambiguous whether this was significantly different.

In the MAS study, Balleisen et al. reported that among the age groups less than 20 (M = 17, F = 45), 20-29 (M = 302, F = 310), 30-39 (M = 496, F = 168), 40-49 (M = 551, F = 210), 50-59 (M = 349, F = 173), and 60 and over (M = 59, F = 15) there were no significant difference between the FVIII:C levels of males and females(64). Jeremic et al. also failed to find a significant effect of sex in a multiple regression analyses of FVIII:C levels in 1,016 regular blood donors when controlling for age and ABO blood groups(96). In the 1,724 subjects that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 CARDIA sample study, Green et al. found no race-stratified differences in the mean FVIII:C levels among the males and females(83).

The prevailing belief is that sex is an important factor to consider in an investigation of FVIII:C levels. Given the conflicting results and the lack of analyses that accounted for important covariates, the reported effect of sex on FVIII:C levels may be

spurious. **Table 1-13** summarizes the associations of sex with FVIII:C levels in these studies.

Table 1-13. The associations of sex with FVIII:C levels

Result	Source	Investigators
Women have higher levels	ARIC[+]	Conlan et al.(80)
Women have higher levels	CHS[+]	Tracy et al.(88)
No statistical difference	MAS[+]	Balleisen et al.(64)
No statistical difference	Blood donors[+]	Jeremic et al.(96)
No statistical difference	CARDIA[+]	Green et al.(83)

[+]Stratified results

RACE

In the ARIC study, Conlan et al. performed separate multiple linear regression analyses for each sex in which black race was associated with a significant increase of 17.0 IU/dL and 15.0 IU/dL among men and women, respectively(80). In the CHS study, Tracy et al. reported that blacks had significantly higher mean FVIII:C levels than whites, 138 IU/dL and 121 IU/dL, respectively(88). In the 1,724 subjects that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 CARDIA sample study, Green et al. within the blood groups B, AB, and O, blacks had significantly higher mean FVIII:C levels than whites(83). In the cross-sectional study (N = 123) of oral contraceptive use that Kadir et al. performed, the ABO blood group adjusted mean FVIII:C levels of black women were significantly higher than those of white, Mediterranean, and other women(77). In the investigation among the controls of the

CDC Menorrhagia study, Miller et al. reported a statistically significant difference in the crude mean FVIII:C levels among black and white women(90). When Miller et al. stratified by O versus Non-O blood type, there persisted a statistical difference in the crude FVIII:C levels among races. When Miller et al. adjusted for VWF:Ag, however, these effects were no longer significant. Iso et al. surveyed Japanese men aged 34-55 years living in a Japanese farming community (N = 29), a Japanese urban area (N = 34), an American urban area (N = 36), and a Caucasians living in a suburban area (N = 35)(97). Iso et al. found the mean FVIII:C levels to be 81, 90, 97, and 96, respectively, with the overall means significantly different. In the Kitava study, Lindeberg et al. found that sex-stratified mean FVIII:C levels were significantly different among the Kitavan and Swedes. They did not report on the age of the Swedes, but that they were in "three different age cohorts (46, 56, and 66 years old)." The results of Iso et al. and Lindeberg et al. (Kitava study), however, do not offer much insight since the numbers are small and they are crude results.

Apparently, FVIII:C levels differ among races. Interestingly, Iso et al. reported that Japanese men living in an urban area of the United States had similar levels to Caucasians living in a nearby suburban area, but different from those of Japanese men in a living in both an urban and rural area of Japan(97). It is not possible to evaluate from this study, but it suggests that perhaps race is simply a confounding factor. The ARIC study, in which the analyses included important variables such as age and BMI, but not VWF, still found a significant effect of race(80). It appears that FVIII:C levels at least differ among blacks and whites. **Table 1-14** summarizes the effects of race on FVIII:C levels in these studies.

Table 1-14. The associations of race with FVIII:C levels

Groups	Results	Source	Investigators
Black vs White	Black race associated with an increase of 17.0 IU/dL and 15.0 IU/dL among men and women, respectively	ARIC	Conlan et al.(80)
Black vs White*	Black race associated with an increase	CHS	Tracy et al.(88)
Black vs White+	Black race associated with an increase within blood groups	CARDIA	Green et al.(83)
Black versus white, Mediterranean, and other+	Black race associated with an increase		Kadir et al.(77)
Japanese versus Caucasian*	See text		Iso et al.(97)
Black versus white, stratified by ABO blood type+	No statistical difference	CDC Menorrhagia	Miller et al.(90)

*Crude or unadjusted results
+Stratified results

BMI

In the analysis of the baseline data of the ARIC study, Conlan et al. reported than mean FVIII:C levels adjusted for age, race, and sex, increased with the BMI (**Table 1-15**)(80). In the multiple regression analyses of these data that Conlan et al. separately performed in the 5,022 men and 6,363 women in the ARIC study, each 5 kg/m^2 increase

Table 1-15. The association between BMI and mean FVIII:C levels adjusted for age, sex, and race in the ARIC study(80).

BMI category (kg/m^2)	Adjusted mean FVIII:C levels in MEN	Adjusted mean FVIII:C levels in WOMEN
Less than 23.97	128	128
23.97 – 26.83*	127	132
26.83 – 30.35*	128	135
30.35 and over	134	142

*Inclusive of the lower bound, exclusive of the upper bound

in BMI significantly mean FVIII:C levels by 1.44 IU/mL and 2.82 IU/mL in men and women, respectively(80).

Cushman et al. found that the age-adjusted mean FVIII:C levels of the 1,398 obese subjects, that is subjects whose BMI was greater than 130% of their ideal BMI, was significantly higher than 3,626 non-obese subjects in the CHS study, 128 IU/dL versus 120 IU/dL, respectively(98). Cushman et al. also performed multiple regression analyses separately for 2,150 men and 2,874 women and found that obesity was not significantly associated with the mean FVIII:C levels of women or men(98). In the multiple regression analysis of the baseline data that Balleisen et al. separately performed in the men and women in the MAS study, the Broca-Index (height in cm minus 100) was not significantly associated with FVIII:C levels(64). In the 1,724 subjects that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 CARDIA sample study, Green et al. found significantly correlations between FVIII:C levels and BMI in all race-sex categories but black male: $\rho_{bf} = 0.09$, $\rho_{wf} = 0.26$, and $\rho_{wm} = 0.09$(83). In the Caerphilly study, Yarnell et al. found a significant U-shaped, i.e. quadratic effect, of BMI on adjusted FVIII:C levels(91).

Cigolini et al. further investigated a correlate of BMI, visceral fat accumulation, in 52 healthy 38 year old men(99). They dichotomized the subjects using the median value of their visceral fat area (91 cm^2). As expected, there was a significant difference in BMI (24.4 versus 26.8 kg/m^2), 2-hour glucose tolerance test (4.72 versus 5.60 mmol/L), and 2-hour insulin (24.2 versus 43.0 μU/mL). There was also a significant difference in the crude mean FVIII:C levels (74.5 versus 89.0)(99).

Of the reviewed studies, the only study that used an alternative to BMI was the study to not find a significant difference. This study, the MAS study, was among workers, who were likely more able-bodied and of a narrow range of BMI than the community drawn subjects of the other four studies.

Table 1-16. The effects of BMI on mean FVIII:C levels

Increment	Results	Source	Investigators
5 kg/m^2	1.44 in males, 2.82 in females	ARIC	Conlan et al.(80)
Obese versus non-obese	No statistical difference	CHS	Cushman et al.(98)
N/A	120 in non-obese, 128 in obese[+]	CHS	Cushman et al.(98)
1 unit	No statistical difference[#]	MAS	Balleisen et al.(64)
N/A	Significantly correlated except in black males[*]	CARDIA	Green et al.(83)
1 unit	U-shaped trend	Caerphilly	Yarnell et al.(91)
Visceral fat area	74.5 in lower group, 89.0 in higher group	N/A	Cigolini et al.(99)

[*]Crude or unadjusted results
[+]Stratified results
[#]Broca-index, not BMI

The ARIC study included lipids among the covariates, which suggests that BMI affects FVIII:C levels through different unknown mechanisms(80). **Table 1-16** summarizes the associations of BMI with FVIII:C levels in these studies.

Blood glucose/insulin/diabetes

In the analysis of the baseline data of the ARIC study, Conlan et al. reported than mean FVIII:C levels adjusted for age, race, and sex was significantly higher in diabetics and increased with serum insulin levels (**Table 1-17**)(80). In the multiple regression

analyses of these data that Conlan et al. separately performed in the 5,022 men and 6,363 women in the ARIC study, diabetes significantly increased mean FVIII:C levels by 14.0 IU/dL and 18.0 IU/dL , in men and women, respectively(80). In addition, each 30 μU/mL increase in serum insulin significantly increased mean FVIII:C levels by 1.02 IU/dL and 2.41 IU/dL , in men and women, respectively(80).

Table 1-17. The association between mean FVIII:C levels adjusted for age, sex, and race, and diabetes and serum insulin in the ARIC study(80).

Diabetes	**Adjusted mean FVIII:C levels in MEN**	**Adjusted mean FVIII:C levels in WOMEN**
No	127	132
Yes	143	158
Serum Insulin (μU/mL)		
Less than 6	123	122
6 – 9*	124	129
9 – 14*	127	133
14 and over	134	140

*Inclusive of the lower bound, exclusive of the upper bound

Cushman et al. found that the age-adjusted mean FVIII:C levels significantly increased with both glucose and insulin level (**Table 1-18**) in the CHS study(98). Cushman et al. also performed multiple regression analyses separately for 2,150 men and 2,874 women and found that each 35 mg/dL increase in glucose was significantly associated with a 5.0 IU/dL increase and a 6.6 IU/dL increase in mean FVIII:C levels in men and women, respectively(98). In addition, each 25 μU/mL increase in serum insulin was significantly associated with a 0.5 IU/dL increase in mean FVIII:C levels in women and an significant, but omitted, increase in men (98).

Table 1-18. The association between age-adjusted glucose and insulin levels and mean FVIII:C levels in the CHS study(98)

Glucose (mg/dL)	**Adjusted mean FVIII:C levels**
less than 96	117
96 – 107*	120
greater than 107	130
Insulin (μU/mL)	
less than 11	116
11 – 15	122
greater than 15	130

*Inclusive of the lower bound, exclusive of the upper bound

In the analysis of the baseline data of the MAS study, Balleisen et al. reported that the estimated correlations between FVIII:C levels and blood glucose levels were not significantly different from zero in the 2,880 men and 1,306 women(100). In the multiple regression analyses that Balleisen et al. separately performed in the men and women in the MAS study, blood glucose concentration was not significantly associated with FVIII:C levels in men or women. The mean (SD) blood glucose level in men and women were 98.6 (14.4) mg/dL and 89.5 (11.6) mg/dL, respectively.

Temporality of effects remains an important consideration in the relationship between diabetes and high FVIII:C levels. It could be that high FVIII:C is either a component cause of diabetes or is a confounder. A perturbed endothelium, a hallmark of the vessel damage that accompanies diabetes, could result increased FVIII:C levels(101).

Bern et al. study four groups of diabetic patients and a group of non-diabetic subjects to investigate the association between atherosclerosis and FVIII:C levels(102). The groups of diabetes patients were: juvenile-onset insulin-dependent diabetics without retinopathy, juvenile-onset insulin-dependent diabetics with proliferative retinopathy but without peripheral vascular disease (PVD), insulin-dependent diabetics with

atherosclerosis, and patients with PVD but with no personal or family history of glucose-intolerance who had normal Hb1Ac levels. Each group had 11 men. The mean FVIII:C levels of each diabetic group was significantly higher than that of the control group, but not different from that of the remaining diabetics groups.

Hughes et al. have investigated FVIII:C levels in 13 men and 24 women, aged 45-67 years, who were newly diagnosed with type 2 diabetes at the time of diagnosis and following two to three months of therapy with diet and oral hypoglycaemic agents(103). At diagnosis, the mean (SD) FVIII:C levels were 144 (44) IU/dL, significantly higher than the mean (SD) FVIII:C levels obtained following therapy 107 (21) IU/dL. The corresponding HbA1c levels were 10.7% (2.2%) and 7.7% (1.2%), respectively.

Herlihy et al. studied 13 pairs of siblings of type II diabetic patients who were discordant for fasting plasma glucose levels (normoglycemic and hyperglycemic) and 13 controls matched on age and sex(104). The mean FVIII:C levels did not differ between the normoglycemic and hyperglycemic siblings, whereas only waist circumference and markers of insulin sensitivity, such as fasting insulin levels, were significantly different.

It appears that glucose levels and diabetes status are each associated with higher FVIII:C levels. The results of Bern et al. suggest that atherosclerosis is associated with increased FVIII:C levels in both diabetic and non-diabetic patients, but they do not address temporality or elucidate whether other sequelae of diabetes other than angiopathy may increase FVIII:C levels. The results of Hughes et al. provide insight into this quandary since their study patients were free of clinical angiopathy and therapy was associated with a concomittal decrease in glucose levels and FVIII:C suggesting that glucose levels affect FVIII:C levels(103). The studies to date, do not permit evaluation

of the possibility that high FVIII:C levels precede the development of diabetes through some undetermined pathological consequences or as a confounder.

LIPIDS

Cushman et al. reported that the age-adjusted mean FVIII:C levels significantly increased with total cholesterol, LDL cholesterol, and triglyceride levels and significantly decreased with HDL cholesterol levels (**Table 1-19**) in the CHS study(98). Cushman et al. also performed multiple regression analyses separately for 2,150 men and 2,874 women and found that each 35 mg/dL increase in LDL cholesterol was significantly associated with a 1.6 IU/dL increase in mean FVIII:C levels in men only. No significant

Table 1-19. The age-adjusted mean FVIII:C levels by cholesterol level in the CHS study(98)

Total cholesterol (mg/dL)	**Adjusted mean FVIII:C levels**
Less than 197	121
197 – 229	120
Greater than 229	127
HDL cholesterol (mg/dL)	
Less than 46	124
46 – 58	122
Greater than 58	120
LDL cholesterol (mg/dL)	
Less than 117	121
117 – 146	120
Greater then 146	125
Triglycerides (mg/dL)	
Less than 103	118
103 – 149	122
Greater than 149	126

association was found between FVIII:C levels and either HDL cholesterol or triglycerides in either men or women(98).

In the analysis of the baseline data of the MAS study, Balleisen et al. reported that the only significant correlations between FVIII:C levels and measured lipids were found among women for cholesterol ($\rho = -0.070$) and LDL ($\rho = 0.067$)(100). In the multiple regression analyses that Balleisen et al. separately performed in the men and women in the MAS study, LDL levels were significantly associated with FVIII:C levels only women, but the estimated coefficient was –0.106(100).

In the 1,724 subjects that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 CARDIA sample study, Green et al. found significantly correlations between FVIII:C levels and triglyceride levels in all race-sex categories but black male: $\rho_{bf} = 0.13$, $\rho_{wf} = 0.12$, and $\rho_{wm} = 0.21$(83). Additionally, in white women, there were significant correlations between FVIII:C levels and both cholesterol and LDL, $\rho_{wf} = 0.13$ and $\rho_{wf} = 0.10$, respectively.

The results of these studies do not offer much support for an effect of lipids on FVIII:C levels. At best, one might conclude that further research, especially that focuses on this aspect, needs to establish the association. Khrenov et al. reported that both LDL and HDL are able to support the assembly of the Xase complex in a manner similar to the phospholipid surface(105). Whether this has physiologic relevance is dubious and even among the results supporting an association, the estimated effect sizes are rather unimpressive.

PREGNANCY

Sie et al. measured VWF associated molecules, including FVIII:C levels, cross-sectionally at one of nine three-week time points in 306 pregnant women who were undergoing routine prenatal examinations(106). These women were normotensive and without a history of diabetes, renal disease, autoimmune disease, or vascular complications during a previous pregnancy. From among the initial enrolled cohort of 406 women, the investigators excluded 100 women: 15 of whom had spontaneous abortions, 34 of whom delivered babies below the 10^{th} percentile of weight adjusted for age and sex, 33 of whom delivered babies above the 90^{th} percentile of weight adjusted for age and sex, 10 of whom had non-singleton births, and 10 of whom they lost to during follow-up. There were 150 women with non-OO blood groups, 140 with the OO blood group, and 16 for whom it was undetermined. The median age was 29 years. Sie et al. found that the logarithm of FVIII:C levels of the mothers correlated with gestational age ($\rho = 0.35$). Unfortunately, they did not perform multiple linear regression nor stratify on ABO blood group. **Table 1-20** presents the FVIII:C levels of mothers that Sie et al. cross-sectionally obtained across gestational age. The unadjusted ratio VWF:Ag to FVIII:C did not increase with gestational age among the entire cohort(106).

Stirling et al. measured FVIII:C levels longitudinally in 80 women who had prenatal visit at the Northwick Park Hospital and were either 1) primigravidae and clinically well or 2) multigravidae, clinically well, and whose previous pregnancies had been free of complications(107). The investigators obtained a blood sample at the first prenatal visit, which may have occurred between the 9^{th} and 27^{th} week, and visits that occurred at approximately the 20^{th}, 28^{th}, 34^{th}, and 38^{th} weeks of pregnancy, post-delivery,

Table 1-20. The FVIII:C levels of mothers obtained cross-sectionally across gestational age as found by Sie el al.(106)

Gestational Week	n	Mean FVIII:C levels of the mothers (IU/dL)
14 – 19	37	128
19 – 21	21	144
22 – 24	27	153
25 – 27	34	162
28 – 30	38	172
31 – 33	33	183
34 – 36	47	194
37 – 39	51	206
40 – 41	18	218

which occur between 6 hours and 4 days after delivery, and at the first post-natal visit, which occur between 2 weeks and 12 months in only 61 women. The investigators represent time as gestational age, which they calculated using the dates of last menstruations and ultra-sound results that occurred between the 14^{th} and 18^{th} weeks. There were 59 white women, 4 black women, 6 Asian women, and 3 Chinese women and the mean age was 27 years. The FVIII:C levels of the mothers are very similar to the results of Sie et al. above. The post-delivery mean FVIII:C levels remained elevated and the levels did not begin to decline until 2-7 weeks postnatally.

Sanchez-Luceros et al. measured FVIII:C levels cross-sectionally in 248 women who were primigravidae or multigravidae with no history of complications, no personal or family history of bleeding, thromboembolism, hypertension, spontaneous abortion, or other systemic disease(108). They investigators included only women who lacked bleeding and thrombosis complaints and who had normal menstrual losses. The mean age was 25 years. Sanchez-Luceros et al. reported similar increases in VWF:Ag as did both Sie et al. and Stirling et al. and they reported an increase in FVIII:C levels with

gestational age, but the magnitude of the increase is much less (**Table 1-21**). There were 150 women with the OO blood group.

Table 1-21. The FVIII:C levels of mothers obtained cross-sectionally across gestational age as found by Sanchez-Luceros et al.(108).

Gestational Week	n	Mean FVIII:C levels of the mothers (IU/dL)
6 – 11	23	83
12 – 16	36	90
17 – 23	31	87
24 – 28	28	91
29 – 35	39	98
36 – 40	27	89
Early puerperium	27	121
Late puerperium	37	87

These three studies support an increase in FVIII:C levels of the mother with gestational age, though the study of Sanchez-Luceros provided unexpectedly modest increase in FVIII:C levels. Pregnancy is a hypercoagulable state(109), presumably an adaptation to stem the loss of blood in the event of delivery or spontaneous abortion. From implantation forward, the placenta challenges the circulatory system of the mother, with the possible effect of perturbing the endothelium, which could lead, in part, to the observed increases in FVIII:C levels.

MENSTRUATION

Miller et al. investigated the variation in FVIII:C levels with stage of the menstrual cycle in 90 controls and 85 cases from the CDC Menorrhagia study who

reported that the onsets of their last menstruation periods were fewer than 31 days prior to blood sampling(110). In this cross-sectional analysis, the investigators estimated the phase of the cycles by subtracting the dates of the last menstruations from the dates that they obtained the blood samples. There were between 19 and 20 subjects in each phase: 1-4 days (menses), 5-7 days (early follicular), 8-11 days (late follicular), 12-16 days (ovulatory), 17-21 days (early luteal), and 22-30 days (late luteal). Further, the distribution of ABO blood type, race, age, oral contraceptive use, and case/control status was not different across the phases. Miller et al. reported that there were significantly different mean FVIII:C levels across the phases among the entire study population and among the controls only.

Kadir et al. investigated the variation in FVIII:C levels with stage of the menstrual cycle in both a cross-sectional study of 123 women referred to gynaecological clinics because of a complaint of heavy periods but regular cycles (23-39 days) and a longitudinal study of 39 normal volunteers(77). The mean age of the women in the cross-sectional study was 39 years, 39% of whom were in the blood group OO. The mean age of the women in the longitudinal study was 26 years old. Kadir et al. reported that there were no significant differences in mean FVIII:C levels across the phases of the menstrual cycles of women in the cross-sectional study with the following groupings: 1-6 days (early follicular), 7-11 (late follicular), 12-16 ovulation, 17 and over (luteal). However, Kadir et al. reported that FVIII:C levels had a significantly cyclical pattern adjusted for age, age dichotomized at the median, blood group (A versus non-A), dichotomous alcohol use, and dichotomous smoking status in the longitudinal study. When Kadir et al.

performed these analyses separated for oral contraceptive users and non-user, only users had a significantly cyclical pattern.

These studies suggest that phase of the menstrual cycle effects FVIII:C levels. The absence of measurements of hormones dictates the need for further research if elucidation of the nature of this effect is to be possible. The results of these studies suggest that, in future studies that involve adult women, it will be important to at least record the date of last menstruation, if not limit collection to a certain phase.

DEPRESSION

Depression is a common psychiatric illness. In lay terms, depression may be chronic sadness or gloom, with feelings of hopelessness, and accompanied by a decrease in physical and mental expenditures. The fourth edition of the "Diagnostic and Statistical Manual of Mental Disorders" classifies "depression" as a mood disorder(111). Of specific interest is Major Depressive Episode (MDE), "a period of at least 2 weeks during which there is either depressed mood or the loss of interest or pleasure in nearly all activities."(111) To diagnose MDE, the physician must observe, during a two week period, five or more of the following symptoms, which must represent a change in previous functioning:

A.

(1) Depressed mood

(2) Diminished interest or pleasure

(3) Significant weight or appetite loss

(4) Insomnia or hypersomnia

(5) Psychomotor agitation or retardation

(6) Fatigue

(7) Feelings of worthlessness or excessive or inappropriate guilt

(8) Impaired ability to think or concentrate

(9) Recurrent thoughts of death

B. Must not meet the criteria for a Mixed Episode

C. The symptoms must cause clinically significant distress of impairment

D. The symptoms should not be due to exposure to a substance (e.g. drugs or medication) or to a medical condition

E. The symptoms are not accounted for by Bereavement

Further, these symptoms must persist throughout most of the day for the duration of the period. One or more episodes of MDE may represent Major Depressive Disorder (MDD). Several large studies of the general, non-institutionalized population have estimated the lifetime prevalence of MDD to be between 13%-16%(112-115). The estimated prevalence in the 12-month period prior to the interview in two of these studies was between 5.3%-6.6%(113, 114). Interestingly, diagnostic laboratory findings for MDE have yet to be identified(111).

As determining a diagnosis of MDD is not conducive to large studies, observational or clinical, Radloff designed the Center for Epidemiologic Study Depression instrument (CES-D), which is in wide used(116). Briefly, the CES-D

instrument consists of twenty questions concerning depressive symptoms. Each question has four possible answers relating to the frequency of the occurrence. Three of the questions are phrased positively to interrupt a possible flow of set responses. Each answer receives a score of 0-3, with the positively phrased questions being scored in reverse. The final CES-D score, with penalties for missing answers, ranges from 0-60, which investigators conventionally dichotomize by designating subjects with scores 16 or greater as depressed or of depressive symptomatology(116). Wong(117) and McDowell(118) provide salient discussions of the CES-D scale, including its use, validity, and reliability.

Kop et al.(119) used a modified 10-item Center for Epidemiological Study Depression (CES-D) instrument treating the score as both a continuous variable and dichotomizing at cut-off of 10 or greater. They found that the continuous CES-D score was significantly associated with FVIII:C levels, but that FVIII:C levels were only weakly associated with the dichotomous categorization. Eskandari et al.(120) also found a positive, significant association between FVIII:C levels and MDD, as measured by the Hamilton Depression Scale. Finally, several investigation have reported an association between depression and other procoagulant factors(119, 121-123), supporting the plausibility of the association with FVIII.

ABO BLOOD GROUP

Jeremic et al. investigated the effects of age, ABO blood group, and sex on FVIII:C levels of 1,016 regular blood donors(96). The investigators grouped the donors into the phenotypic blood groups 1) A: AA or AO, 2) O: OO, and 3) B+AB: BO, BB, or

AB. Using a multiple regression analyses, Jeremic et al. reported that ABO blood group was significantly associated with mean FVIII:C levels. In addition, their results suggest that the interaction between ABO and age may also have been significantly associated with mean FVIII:C levels ($p = 0.05$). The least squares constants for blood groups A, O, and B+AB were 28, -57, and 28, respectively.

McCallum et al. investigated the effects of ABO blood group in 136 female blood donors aged 18 to 50 years(124). The investigators grouped the donors into phenotypic blood groups 1) A: AA or AO, 2) O: OO, 3) B: BO or BB, and 4) AB. McCallum et al. reported an analysis of variance of the logarithm of FVIII:C levels suggested that the mean FVIII:C levels of donors with the O phenotype was significantly lower than those of donors with the A, B, and AB phenotypes (**Table 1-22**). In response, Mohanty et al. attempted to replicate this finding in 50 male and 46 female normal individuals(125). Mohanty et al. reported that the mean FVIII:C levels of persons with phenotypic group O and B were similar and that these levels were significantly different from those with phenotypic group A and AB. McCallum. et al used a one-stage assay(124), whereas Mohanty et al. used a two-stage assay(125). In addition, Mohanty et al. did not report the mean age nor range of their subjects.

Green et al. investigated FVIII:C levels from the 1,724 CARDIA study that were from Chicago and Minneapolis and aged 23 to 35 years old in the 1990 to 1991 sample(83). Green et al. reported that mean FVIII:C levels of subjects with the A, AB, and B phenotypic blood groups were significantly higher than those of subjects with the O phenotypic blood groups for each of the sex-race (black/white) groups, except for

Table 1-22. The ANOVA results of the log transformed FVIII:C levels of 136 female blood donors(124).

Blood group phenotype	Mean FVIII:C levels (IU/dL)
O	92.6
A	126.3
B	135.2
AB	139.5

black men with the A phenotype and white men with the AB phenotype (**Table 1-23**)(83).

Orstavik et al. investigated FVIII:Ag levels in 74 pairs of monozygotic twins and 84 like-sexed pairs of dizygotic twins who were either 33-39 years or 57-62 years old(126). The investigators determined antigenic ABO phenotype. To overcome the lack of independence, the investigators randomly selected one twin from each pair and included the 17 individuals from the pairs that had samples on only one of the two twins. The resulting sample included 60 individuals with the O phenotype, 68 individuals with the A_1 phenotype, 20 individuals with the A_2 phenotype, 15 individuals with the B phenotype, and 4 individuals with the A_1B phenotype. Orstavik et al. found that the association between FVIII:Ag levels and ABO phenotype no longer persisted after they adjusted for VWF, i.e. FVIIIR:Ag(126).

Clearly, persons with an OO genotype (O phenotype) have mean FVIII:C levels that are lower than persons with non-OO genotypes. Some investigators chose to represent this factor with respect to A and non-A phenotypes. The findings of Mohanty et al. suggest that the later representation may be more conservative, if one chooses to use such restrictive categorizations such as dichotomization.

Table 1-23. FVIII:C levels by ABO phenotype, sex, and race in the CARDIA study(83).

	Mean (SD) FVIII:C levels (IU/dL)			
ABO phenotype	**Black women**	**Black men**	**White women**	**White men**
A	109 (40) n = 70	102 (41) n = 51	102 (31) n = 137	93 (34) n = 145
AB	139 (59) n =10	124 (33) n = 9	98 (34) n = 17	91 (38) n = 14
B	126 (39) n = 56	129 (40) n = 47	98 (32) n = 46	95 (32) n = 52
O	86 (34) n = 135	87 (27) n = 84	75 (25) n = 158	77 (25) n = 160

VWF

Cadroy et al. investigated 16 hemostatic values in a group of 58 males (mean age 40 years) and 42 females (mean age 32 years) they randomly selected among blood donors(127). They had data only on age and sex. They noted that FVIII:C levels were not correlated with age, but that males had a higher mean FVIII:C level than that of females. Cadroy et al. reported that the estimated beta coefficient for FVIII:C levels was 0.69 ($p < 0.01$), but did not list the other variables included in the model(127).

Morange et al. investigated FVIII:C levels in a study of the members of 100 unrelated families (parents and at least two offspring 10 years or older) who were free of acute and chronic diseases and underwent a free health examination(47). Morange et al. reported the correlations coefficients between FVIII:C and VWF levels in fathers, mothers, sons, and daughters as 0.80, 0.70, 0.67, and 0.73, respectively. The members of the first two classes represented independent units since none of the families were related.

Using the controls from the CDC menorrhagia study, Miller et al. investigated the effects of measures of VWF, race and ABO blood group on FVIII:C levels(90). In a

step-wise multiple regression analyses, VWF:Ag accounted for 0.54 of the total variance of FVIII:C levels ($p = 0.0001$). Including VWF:Ag levels in regression analyses of FVIII:C levels removed the effects of both ABO blood group and race. Miller et al. reported that the correlation coefficient between FVIII:C levels and VWF:Ag was 0.77, which was presumably significant(90).

VWF is a major determinant of FVIII:C levels. Unexpectedly, though, a dearth of epidemiologic studies investigating the relationship between FVIII:C and VWF levels exists. This is surprising since deficiencies of VWF, the carrier of FVIII, are not uncommon. VWF circulates in multimeric forms of heterogenous sizes. The stoichiometry is one FVIII molecule per 50 VWF monomers *in vitro*(128). The degree of multimerization does not affect the equilibrium binding constant *in vitro*, though the highest molecular weight VWF multimers are associated with the highest hemostatic efficacy(129). It is pertinent to explore this relationship more fully in population based studies.

FIX

Cadroy et al. investigated 16 hemostatic values in a group of 58 males (mean age 40 years) and 42 females (mean age 32 years) they randomly selected among blood donors(127). They had data only on age and sex. Cadroy et al. reported that FVIII:C levels were significantly correlated with FIX:C levels ($p < 0.01$)(127).

Lowe et al. reported on FVIII:C levels in the third Glasgow MONICA Survey(78). They had FVIII:C level measurements on 714 men and 789 women, who were between 25 and 74 years of age. Lowe et al. reported that FVIII:C levels

significantly correlated with FIX:C levels with $\rho = 0.62$ and $\rho = 0.64$, in men and women, respectively.

van Hylckama Vlieg et al. investigated FIX:C levels in 153 women who were control subjects of the LETS study between 15 and 49 years old(130). They performed (linear) regression with FIX:C levels as the dependent variable. van Hylckama Vlieg et al. reported that the coefficient for FVIII:C levels was 0.18 (0.12-0.24), though it is not clear whether this is a multiple linear regression.

FIX is the enzyme for which FVIII is the cofactor. The are no reported functions of either protein except as the tenase complex. An organism that coordinates production of these proteins should conserve resources. Whether this indeed occurs is unknown, especially since many studies fail to either measure both of them or investigate the associations between them.

1-VIII. FVIII:C levels and thrombosis

Koster et al. first reported in 1995 that high FVIII:C levels were associated with deep vein thrombosis(DVT)(131). They inferred this conclusion using a subset of the Leiden Thrombophilia Study (LETS), a case-control study(132). From three thrombosis clinics in the Netherlands, LETS recruited 474 consecutive patients with an objectively confirmed DVT, who were younger than 70 years, without malignancy, and experiencing their first clinically recorded event. They asked each patient to recruit their own healthy

control that was the same sex, within 5 years of age, not related, did not have a history of thrombotic events, did not take anti-coagulants in the preceding three months, and had no known malignancy. When a patient failed to provide a control, a partner of another patient served as the control subject to the same constraints. This occurred for 225 (47%) of the patients.

Restricting to the period between January 01, 1988 and December 31, 1992, Koster at el. obtained 374 patients and their controls, 126 of whom were partners of the other patients(131). After enforcing exclusion criteria, such as anti-coagulant treatment, there were 121 male and 180 female case-control pairs, with 42% of the controls being partners of other patients. They performed two conditional logistic regression analyses. In one, they treated FVIII:C levels as a continuous variable and in the other they categorized FVIII:C levels into quartiles (<1000, 1000-1249, 1250-1499, and >1500 IU/L). With the first quartile as the reference category, they reported adjusted ORs of 2.3 (1.3 – 3.8), 3.0 (1.6 – 5.7), and 4.8 (2.3 – 10.0) for the second, third, and forth quartiles, respectively, with a significant test for trend ($p < 0.001$)(131). They adjusted for O versus non-O blood groups, von Willebrand Factor. They reported that BMI, diabetes mellitus, and smoking did not change the results. Unfortunately, they did not provide details of variables that represented these factors nor the modeling strategies. When they treated FVIII:C levels as continuous, they reported that each 10 IU/L increase was associated with a 1.019 (1.013 – 1.025) increase in the OR. However, the details were lacking and it is unclear whether these were adjusted results. It would be rather interesting to know if they investigated higher order and interaction terms or splines.

Patel et al. reported similar findings case-control study of blacks of African or Caribbean descent matched on age and sex(133). When they considered exposure as FVIII:C levels greater than or equal to 250 IU/L (the 90th percentile in the "normal blacks") the crude OR was 12.5 (4.0 – 40.6). When they considered exposure as FVIII:C levels greater than or equal to 150 IU/L the crude OR was 7.0 (3.5 – 14.0). The reported methods did not indicate that they performed a conditional logistic regression nor accounted for covariates.

At least two groups have investigated the role of high FVIII:C levels in recurrent venous thrombosis. Kyrle et al. followed 360 patients for an average of 30 months after anti-coagulation treatment ended for their first clinically recorded venous thromboembolism (VTE). These patients were members of a larger cohort of 1259 patients collected between July 1992 and December 1999(134). They reported that the hazard ratio increased by 1.07 (1.02 –1.12) per 10 IU/L increase in FVIII:C levels, adjusted for age, sex, FVL, Prothrombin G20210A, and duration of oral anticoagulant treatment following the event. Cristina et al. analyzed 924.4 person-years from 564 patients who had ended anti-coagulant treatment for a first VTE and had no liver failure, renal failure, or cancer(135). Their results from 53 recurrent cases were not as strong. Only when they analyzed the subset of data from patients whose index event was idiopathic (secondary to an environmental exposure such as pregnancy, trauma, or surgery) did their results become interesting, though they likely suffer a lack of power due to small sample size. With the referent group being the patients with FVIII:C levels less than 1.47 IU/mL, the risk ratio for the group FVIII:C levels above 2.98 IU/mL in the chromogenic assay and above 2.94 IU/mL in the clotting assay was 5.4 (1.8 – 16.8) and

6.2 (1.6 – 24.5), respectively. They adjusted for age, sex, duration of anticoagulation, and thrombophilic alteration status(135).

From a simplistic view of the reaction, an increase in FVIII:C could be expected to increase the rate of clot formation. This is not a unique aspect of FVIII:C, but is consistent with Le Chatlier's principle which states that adding product or reactants to a reaction at equilibrium drives the reaction towards the unperturbed component or components. Therefore, we might expect this to be approximately true of any coagulation pathway component, noting that equilibrium is not likely physiologically achieved since the blood carries away the cleaved products from each other or they become incorporated into downstream products.

Though blood flow in the deep veins may be slower at times to the point of occasional stasis during immobility, thrombosis in this location is not likely to be fundamentally unique. Thus, we might expect that high FVIII:C levels might have a role in (inappropriate) thrombotic events in other vessels besides the deep veins, including arteries. This might suggest the involvement of high FVIII:C levels in myocardial infarction and stroke. Indeed, Ananyeva et al. reported that translocation of phosphatidylserine to the outer surface of oxLDL-treated smooth muscle cells (SMC) and macrophages provided additional FVIII binding sites *in vitro*(136). This may have great impact after disruption of the endothelial layer exposing SMC and macrophages.

These data appear convincing, and supported by the central role of FVIII in the intrinsic coagulation pathway, it appears that high FVIII:C levels may cause thrombosis. However, there needs to be considerable refinement of the study and representation of FVIII:C levels. As section **1-VII** details, FVIII:C levels are broadly variable and certain

environmental conditions are associated with spikes that more than double FVIII:C levels. In addition, it is not rare to encounter patients with levels above 250 or even 300 IU/L. In these case-control studies, the investigators were careful draw blood well after the occurrence of any event might have transiently raise the levels. Further, Kamphuisen reported that acute phase reactions did not cause elevation of FVIII:C levels in the LETS study(137). It appears that these FVIII:C levels, therefore, reflect basal levels.

Three important questions to address remain. The first concerns the levels of the other components in the coagulations system. Since no investigation is complete without measurement of VWF, investigators frequently included this protein as a confounder. However, it is possible that high basal levels of FVIII:C are accompanied by high basal levels of other proteins, whether procoagulant, anti-coagulant, or profibrinolytic, for which the investigators may not have controlled, if they had even measured them.

The other two questions concern FVIII directly. The second is whether the amplitude and duration of a spike might relate to basal levels. This may be important since nothing is known about the FVIII:C levels at the time of the thrombosis. The third is whether there might exist refined and more stable time dependent variables that describe FVIII:C levels more accurately or provide better classifications of thrombotic events and their causes. In general, the answers to these last two questions likely involve meticulous research that consumes vast resources, in addition to access to an immense amount of detailed person-years.

An obvious candidate to satisfy these criteria is genetic data. As mentioned in section **1-VI-D** the heritability of FVIII:C levels suggest a moderate genetic component, yet both linkage analyses and candidate gene studies have failed to provide evidence of

the existence of genetic variants that affect the levels. It seems reasonable that variations within the structural *F8* gene may affect the amount or function of the FVIII protein. If such variants exist, then investigation into FVIII are incomplete, even though the effect is likely to be modest since linkage analyses have failed to produce signals in the region of the gene. Thus, we have a strong impetus to resequence the *F8* gene in a population such as the GAIT project that has an estimated $h^2 = 0.4$(138).

1-IX. Goals of the dissertation

In summary, FVIII:C levels are central to adequate hemostasis; insufficient activity results in potentially fatal bleeding, whereas high levels my cause thrombosis. The variability of FVIII activity is quite large, and many factors have been reported to be associated with FVIII:C levels. Nevertheless, some of the reported associations may be inaccurate or spurious, perhaps due to inadequate control of confounding. Further most of the reported risk factors are either environmental or endogenous in nature, and few specific genetic factors have been identified. Therefore, further study and understanding of these associations is needed.

This dissertation has three goals meant to refine and further our understanding of the determinants of FVIII:C levels in populations without obvious FVIII-related coagulopathies. The first goal is to identify and describe the DNA variants in the classically considered functional regions of *F8*. Importantly, a subset of the subjects in this discovery phase is members of a family-based study that included measurement of

FVIII:C levels, in addition to extensive covariate data. The second goal of this dissertation is to genotype the remaining subjects of this study that were not in the first phase and to determine associations of these variants with FVIII:C levels. Importantly, a number of environmental and endogenous factors are available so that analytical adjustment is not only possible, but will be rather extensive relative to previous reports on FVIII:C levels. The third goal of this dissertation is to investigate the association between FVIII and depression, a common mental illness, again while controlling for a relatively extensive array of environmental and endogenous factors.

Chapter 2

A sequence variation scan of the coagulation Factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels.

This chapter of the dissertation has been published in the journal Blood:

Viel KR, Machiah DK, Warren DM, Khachidze M, Buil A, Fernstrom K, Souto JC, Peralta JM, Smith T, Blangero J, Porter S, Warren ST, Fontcuberta J, Soria JM, Flanders WD, Almasy L, and Howard TE**.** A sequence variation scan of the coagulation Factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels. Blood 2007;109:3713-3724.

A sequence variation scan of the coagulation factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels

Kevin R. Viel,[1,2] Deepa K. Machiah,[1] Diane M. Warren,[3] Manana Khachidze,[1] Alfonso Buil,[4] Karl Fernstrom,[1] Juan C. Souto,[4] Juan M. Peralta,[3,5] Todd Smith,[6] John Blangero,[3] Sandra Porter,[6] Stephen T. Warren,[7] Jordi Fontcuberta,[4] Jose M. Soria,[4] W. Dana Flanders,[2] Laura Almasy,[3] and Tom E. Howard[1]

[1]Department of Pathology and Laboratory Medicine and [2]Department of Epidemiology, Emory University, Atlanta, GA; [3]Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, TX; [4]Unitat d Hemostàsia i Trombosi, Departament d Hematologia, Hospital de la Santa Creu i Sant Pau, Barcelona, Spain; [5]Centro de Investigacion en Biologia Celular y Molecular, Universidad de Costa Rica, San José, Costa Rica; [6]Geospiza, Seattle, WA; and [7]Department of Human Genetics, Emory University, Atlanta, GA

Plasma factor VIII coagulant activity (FVIII:C) level is a highly heritable quantitative trait that is strongly correlated with thrombosis risk. Polymorphisms within only 1 gene, the ABO blood-group locus, have been unequivocally demonstrated to contribute to the broad population variability observed for this trait. Because less than 2.5% of the structural FVIII gene *(F8)* has been examined previously, we resequenced all known functional regions in 222 potentially distinct alleles from 137 unrelated nonhemophilic individuals representing 7 racial groups. Eighteen of the 47 variants identified, including 17 single-nucleotide polymorphisms (SNPs), were previously unknown. As the degree of linkage disequilibrium across *F8* was weak overall, we used measured-genotype association analysis to evaluate the influence of each polymorphism on the FVIII:C levels in 398 subjects from 21 pedigrees known as the Genetic Analysis of Idiopathic Thrombophilia project (GAIT). Our results suggested that 92714C>G, a nonsynonymous SNP encoding the B-domain substitution D1241E, was significantly associated with FVIII:C level. After accounting for important covariates, including age and ABO genotype, the association persisted with each C-allele additively increasing the FVIII:C level by 14.3 IU dL^{-1} ($P = .016$). Nevertheless, because the alleles of 56010G>A, a SNP within the 3′ splice junction of intron 7, are strongly associated with 92714C>G in GAIT, additional studies are required to determine whether D1241E is itself a functional variant. (Blood. 2007;109:3713-3724)

Introduction

Factor VIII (FVIII) circulates, bound to von Willebrand factor (VWF), as an inactive heterodimer with domain configuration A1-A2-B (heavy chain) and A3-C1-C2 (light chain). Activated FVIII (FVIIIa), which is released from VWF as a heterotrimer with A1, A2, and A3-C1-C2 subunits, binds the protease FIXa to form intrinsic Xase whose sole function is to catalyze cleavage activation of FX. Intrinsic Xase is regulated by several processes including FVIIIa inactivation through A2-subunit dissociation and cleavage by activated protein C (aPC). Although FVIII coagulant activity (FVIII:C) is deficient in hemophilia-A patients, individuals with elevated levels have an increased risk for both venous[1-3] and arterial thrombosis.[2,4]

FVIII levels in nonhemophilic populations are broadly variable, spanning a greater than 5-fold range.[5,6] Age,[7-12] sex,[7,9] oral contraception (OC),[10,13] smoking,[7,14,15] body-mass index (BMI),[7,12,16,17] diabetes mellitus (DM),[7,16] ABO blood type,[17-22] and plasma levels of total cholesterol (TC),[16,17] low-density lipoprotein (LDL),[16,17,23] triglyceride (TG),[16,17] VWF,[24-29] and FIX coagulant activity (FIX:C)[10,24,30] all have ample evidence supporting their role as correlates of FVIII:C. Family studies, which have yielded heritability estimates ranging between 40% and 57% and 20% and 61%, respectively, for the levels of FVIII:C[21,31] and FVIII antigen (FVIII:Ag),[32-34] also demonstrate a substantial genetic contribution to FVIII variance. Furthermore, the finding by Souto et al[35] of a strong genetic correlation between FVIII:C level and thrombosis ($\rho_g = 0.689$; $P < .001$) suggests that a subset of polymorphisms underlying FVIII variability pleiotropically influence thrombosis risk. Supporting this, Soria et al[36] discovered a region on chromosome 18 (chr18) that contributes to the variance in both FVIII:C and aPC resistance (aPC-R), an established pathophysiologically related thrombosis risk factor. However, this putative FVIII quantitative-trait locus (QTL) and 2 other potential determinants localized to regions on chr5 and chr11 in the first reported genome-wide screen for which FVIII:C was the primary phenotype[37] have yet to be confirmed. In a family-based study of the ABO structural gene, the only well-established FVIII QTL,[17-22] Souto et al[22] applied combined tests for linkage and association to demonstrate that the polymorphisms encoding allelic glycosyltransferases responsible for ABO blood-group antigens also directly influence FVIII:C levels and represent the only known FVIII quantitative-trait nucleotides (QTNs).

While approximately 1000 loss-of-function FVIII gene (*F8*) alleles underlie the heterogeneity observed for FVIII:C in hemophilia A,[38] only 1 *F8*-based QTN contributing to the FVIII

Submitted June 9, 2006; accepted November 3, 2006. Prepublished online as *Blood* First Edition Paper, January 5, 2007; DOI 10.1182/blood-2006-06-026104.

A portion of this work was presented initially at the 45th annual meeting of the American Society of Hematology, San Diego, CA, December 8, 2003.[39]

The online version of this article contains a data supplement.

The publication costs of this article were defrayed in part by page charge payment. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

variability in nonhemophilic individuals has been identified[39] and confirmed[40] despite numerous studies.[25,36,37,39,41-45] The X-linked *F8*[46] is thought to be less variable than other genes,[6] in part because no polymorphisms have been found in prior investigations of this candidate FVIII determinant.[25,41-43,45] Mansvelt et al[41] identified no promoter variants in subjects with high FVIII levels. Regulatory regions required for *F8* expression in vivo remain poorly characterized despite findings from in vitro assays demonstrating maximal transcription with less than 500 bp of the 5′ genomic sequence.[47,48] Morange et al[25] also found no variants in *F8* regions encoding residues implicated in FVIII clearance by the lipoprotein receptor–related protein (LRP). Furthermore, studies of idiopathic thrombophilia subjects, with or without aPC-R, have not identified polymorphisms within *F8* regions encoding aPC-cleavage sites,[42,43,45] in contrast to the aPC-R variants frequently found within homologous *F5* gene regions in thrombosis patients.[49,50] Finally, while a subset of thrombophilic individuals with high prothrombin levels carry the prothrombin gene 3′ untranslated region (UTR) mutation 20210G>A,[51,52] Mansvelt et al[41] found no polymorphisms within this region of *F8* in thrombosis patients with elevated FVIII levels. However, these studies scanned only small segments of *F8* and often used screening methods less sensitive than DNA sequencing.[41]

Since loss-of-function hemophilic variants have been identified within every exon and junctional intronic sequence, multiple *F8* regions are likely essential for FVIII expression and/or function in vivo.[38] Future attempts to identify FVIII QTNs in *F8* should examine every functional region by direct sequencing in a large sample of unrelated healthy subjects. While 98 predominantly unknown *F8* polymorphisms were recently identified in a scan that met most of these criteria, subsequent analyses for associations with FVIII:C were not possible, as no phenotypic data were available for these subjects.[53] To determine whether *F8* is a determinant of FVIII variability, we resequenced all known functional regions within a collection of 222 X chromosomes from 137 unrelated subjects. Next we genotyped the polymorphisms identified in the 398 subjects of the Genetic Analysis of Idiopathic Thrombophilia project (GAIT).[31] Finally, we used association analysis to determine if GAIT subjects have genotype-specific differences in mean FVIII:C levels after adjusting for relevant covariates.

Patients, materials, and methods

Reagents and instruments

HotStarTaq-Master Mix was from QIAGEN (Valencia, CA); AmpPure and CleanSEQ were from Agencourt Biosciences (Beverly, MA); polymerase chain reaction (PCR) and DNA-sequencing oligonucleotides (oligos) were from Invitrogen (Carlsbad, CA); and BigDye-Terminator (v3.1 or v1.1) Cycle-Sequencing Kits and ABI Prism-3100 and -3700 automated sequencers were from ABI (Foster City, CA). SeqMan multiple-sequence alignment program was from DNASTAR (Madison, WI); PHRED (v020425.c)[54,55] was a gift from the University of Washington; and FINCH DNA-Sequencing System (v.2.09) was from Geospiza (Seattle, WA). Reaction conditions and thermal-cycling parameters for all PCR and DNA sequencing are available upon request.

Subjects and phenotypes

(1) GAIT study subjects. All procedures were approved by the institutional review board (IRB) of the Hospital de la Santa Creu i Sant Pau (Barcelona, Spain). Adult subjects gave informed consent for themselves and their minor children. Coded archived DNA samples from these subjects as well as previously obtained data for measured variables, including FVIII:C levels, were used. (2) Emory study subjects. All studies performed on these DNA specimens were approved by the IRB of Emory University (Atlanta, GA). Archived coded DNA samples from these subjects, which were collected for an unrelated study, were the only biologic specimens used. (3) Coriell study subjects. These subjects represent a collection from the National Institute of General Medical Sciences (NIGMS) of commercially available genomic-DNA samples from unrelated healthy subjects and different ethnic backgrounds that we obtained through the Coriell Cell Repository.

Studies on samples from the 3 groups of subjects referred to as GAIT, Coriell, and Emory were approved by the Emory University IRB. GAIT is composed of 398 Spanish white (SW) individuals from 21 extended pedigrees.[31] The recruitment, sampling, and phenotypes measured in GAIT, including FVIII:C levels and additional measured variables, have been extensively described.[31,35] Coriell designates a collection of genomic DNAs from NIGMS human cell panels representing 45 unrelated healthy subjects, including 10 female African American (AA), 3 female and 7 male Chinese (Ch), 5 female and 2 male Japanese (J), 2 female and 3 male Mexican Indian (MI), 3 female and 2 male South American Andean (SAA), and 5 female and 5 male non-J/non-Ch Southeast Asian (SEA) individuals, which were obtained from Coriell (Camden, NJ); catalog numbers available upon request. Emory designates genomic DNAs from 24 unrelated nonthrombotic male subjects, a subset of an archived collection originally obtained for an unrelated study,[56] of which 18 are white American (WA) and 6 AA. The variation discovery group (VDG), which contained 137 unrelated subjects (85 female, 52 male) from 7 racial groups, included all Coriell and Emory individuals and a subset of 68 GAIT founders (58 female, 10 male) with FVIII:C levels at the extremes. In the VDG, "white" designates both SW and WA subjects.

F8 reference sequence

We used the May 2004 human genome (hg) assembly to download sequence data for a 286-kb stretch of DNA from Xq28.1 that contained the entire *F8* (~186 kb) plus 50-kb segments of contiguous DNA flanking both the 5′ and 3′ ends.[57] Referred to as hg17, this was the first version (National Center for Biotechnology Information [NCBI] Build-35) containing all 26 exons of *F8*; specifically, it included exons 21 and 22, which were missing from earlier releases. This sequence is contained in GenBank accession NG_005114 and can be used to cross-reference oligos (Table 1) and polymorphisms (Table 2). To define polymorphisms with respect to the *F8* transcription unit (TU), we used the reverse compliment of these sequences, assigned 1 to the start site mapped by Mansvelt et al[41] and −1 to the base immediately 5′ to it (Figure 1). In this report, we use "hg17" to indicate that the nucleotide (nt) numbering follows this convention.

F8 variation scan

To identify potential FVIII determinants, we scanned the known functional regions of *F8* (including all exons, 50 to 100 bp of each junctional intronic region, approximately 1.2 kb of contiguous promoter sequence, and approximately 300 bp of flanking 3′ genomic DNA) by directly sequencing 500- to 600-bp amplicons that, where necessary to cover extended regions, were overlapping (Table 1). Table S1 (available on the *Blood* website; see the Supplemental Materials link at the top of the online article) lists the lengths and hg17 nt boundaries of these amplicons and amino acids encoded by the exonic regions examined. Based on the July 2003 hg assembly, we initially generated 39 amplicons. Due to the presence of gaps, amplicons 29 and 30 (designated 29A and 30A) lacked exons 21 and 22 (Figure 1). Although absent from this scan, we included amplicons 29B and 30B, containing exons 21 and 22, respectively, in the genotyping phase described below under "Genotyping." All amplicons, which were generated using genomic DNAs from VDG subjects, and their hg17 nt positions are listed in Table 1.

Agencourt Biosciences performed all initial amplification and cycle-sequencing reactions for the variation scan, in which each amplicon was examined on both strands. Water-negative controls for each amplicon were included and evaluated identically. Agencourt used ABI-3700 sequencers

Table 1. *F8* regions resequenced to identify and genotype common polymorphisms

Amplicon*	Forward primer	Reverse primer	Nucleotides†
01	5′ -TTATCAAAGGGGCTTCTTGC	5′ -CATGCCCTTTCTCCTGACC	−1214 to −573
02	5′ -AGCAAGTGTTGAGGTCCAGG	5′ -TGAAGTAGCAAAAGGGAGGC	−683 to −56
03	5′ -CTTCTCCATCCCTCTCCTCC	5′ -CAGAAATGTTTCTTTGGGGC	−168 to 441
04	5′ -CTTCAAATTTGCCTCCTTGC	5′ -AGACCAAGCAGAGGAAGACG	23012 to 23436
05	5′ -AATCTTGCCTCAGAGCAACC	5′ -GAAAAGCAATTCCTAGGGGG	25450 to 26066
06	5′ -GGGCAACAGAGTGAGACTCC	5′ -TTCTGGAACTCAGCTCCTCC	29413 to 30016
07	5′ -GGAGACCTGACATCAAAGCC	5′ -AACCCCATCTCCTTCATTCC	35266 to 35608
08	5′ -TAAGGTGTGAGCACACTGGG	5′ -CGATGAGTTCTGTTCTGAGCC	37821 to 38404
09	5′ -ATGGTGATTGGTGACCTTGG	5′ -GGAAACTAGGGGATCTTGGC	52986 to 53544
10	5′ -GTCTTGCTCCTGCTTTCACC	5′ -TACCCTTGCCATTTGATTCC	55802 to 56428
11	5′ -CTGCTGAAGAGGAGGACTGG	5′ -ATGTCCATTGGAGACAAGGC	56244 to 56855
12	5′ -GATTGTGGTATCTGCAGGGG	5′ -CAACAGCTGGAGAAAGGACC	61345 to 61753
13	5′ -TGACACTTTCACAGTCAACCG	5′ -CAGCAGGCACGTTTACTACG	65333 to 65920
14	5′ -CAGTCACCCTCTTGTCCTGG	5′ -GGGAATTAAAAGGGAGAGGG	68456 to 69067
15	5′ -CCTGGGAATAAGATAATGGGC	5′ -AAATGCTGGTGAGGATGTGG	74699 to 75338
16	5′ -ACAGCAGCAATGCAAAAACC	5′ -TCTATTGCTCCAGGTGATGG	90867 to 91468
17	5′ -ATGCTCTTGCGACAGAGTCC	5′ -AACAAAGCAGGTCCATGAGC	91365 to 91942
18	5′ -TTGGCAAAAAGTCATCTCCC	5′ -CTAATTGCTTTGGACTGGGG	91756 to 92379
19	5′ -CCACCAGATGCACAAAATCC	5′ -TTTGCTTGGTTTGATTTCCC	92247 to 92850
20	5′ -GAAGGTTCATATGAGGGGGC	5′ -ATGACTGCTTTCTTGGACCC	92700 to 93290
21	5′ -TCTGACCAGGGTCCTATTCC	5′ -CATGATTGCTTTCACAAGCG	93200 to 93816
22	5′ -ATTGGATCCTCTTGCTTGGG	5′ -TGTCCCTGATTCCTCTACCC	93674 to 94323
23	5′ -ATGCAAAATGCTTCTCAGGC	5′ -AAAAGCTTGTTCAAAATAAATGG	116050 to 116647
24	5′ -TCTGTACCACTTCTTCCAGGG	5′ -TTTATGCCAGTCCAACCTGC	117559 to 118125
25	5′ -TATTTTTGGAAGGTGGGAGG	5′ -CGAATCCTTTGATCCTGAGC	118087 to 118699
26	5′ -TTGATGAGACCAAAAGCTGG	5′ -AGAGCATGGAGCTTGTCTGC	118318 to 118933
27	5′ -AAGCACTTTGCATTTGAGGG	5′ -TGGAGATCTTCGAGCTTTACC	120414 to 120947
28	5′ -GGACCCCAGTTTCTTCAGC	5′ -AGTGGGAAGTGGAGAGGAGG	121066 to 121510
29B	5′ -GAATTTAATCTCTGATTTCTCTAC	5′ -GAGTGAATGTGATACATTTCCC	122740 to 122902
30B	5′ -TAAAAATAGGTTAAAATAAAGTG	5′ -TTTAAATGACTAATTACATACCA	126453 to 126668
29A	5′ -TCAGGGTTGGTTACTGGAGC	5′ -ACACTACCATGGTCTTGGGG	158245 to 158735
30A	5′ -AGTCAGTGGGCCTGTTATGG	5′ -GTCCCTAGCTCTTGTTCCCC	158526 to 159105
31	5′ -TGGGCAGATAGGGATAGTGG	5′ -TTTGTGCGTTTCTCAACAGC	158833 to 159409
32	5′ -TTCCCACTTCTTCTTGGTGC	5′ -TGGGCATTTAGGTTGACTCC	159304 to 159934
33	5′ -TCATGCCACTACACTCCAGC	5′ -CTGCCCATAACCAAACTTCC	160633 to 161150
34	5′ -GGGTGACAGAGCAAGACTCC	5′ -AAAAGGCTTGGGAATCAAGG	161982 to 162549
35	5′ -AGATGTCCCAGATGCGTAGG	5′ -GCTTTCATGCAGGTTTCTCC	184818 to 185411
36	5′ -TATTTTCTGCAGCTGCTCCC	5′ -CTTTCAACAATTGCATCCTCC	185329 to 185941
37	5′ -GAGGGGCACATTCTTATCTCC	5′ -TCATAGTGAAGGGGTCAGGC	185733 to 186382
38	5′ -CACCACACAATAGGATCCCC	5′ -GTCAATGGGAAAAGAATGCC	186289 to 186832
39	5′ -CAATCCACAAATGATGCAGG	5′ -AGTGCCAGGATTACAGGCAT	186639 to 187259

*To include all known functional *F8* regions in the variation scan, 41 distinct amplicons of the structural locus were PCR amplified from genomic-DNA samples and resequenced directly. The 11 amplicons indicated in italics were directly sequenced to genotype the entire GAIT cohort for the 12 *F8* variations that were polymorphic in white individuals.

†Numbering for nucleotides corresponds to the hg17 reference sequence for *F8*.

and the PHRAP programs PHRED, CONSED, and POLY-PHRED to identify variants.[54,55] The quality of Agencourt's sequence chromatograms, which we used to both detect and genotype polymorphisms, were assessed in-house by (i) uploading chromatograms, including those for a blind genomic-DNA replicate from an Emory subject, into the Finch server to determine average PHRED quality (Q) scores; (ii) manually reviewing SeqMan alignments to validate base calls for minor alleles and strand consistency; (iii) performing agarose-gel electrophoresis and multiple-sequence alignments to validate negative controls for amplification and sequencing reactions, respectively; and (iv) manually reviewing multiple-sequence alignments of an approximately 5% random subset of all chromatograms. Sequences that were either of poor quality (average Q < 30)[54,55] or that yielded inconsistent base calls were repeated in-house. Previously unknown polymorphisms were designated as naturally occurring if minor alleles were present in 2 or more subjects. Polymorphisms with minor alleles found in only 1 subject were considered naturally occurring if confirmed upon sequencing a second amplicon derived from an independent PCR.

LD

To evaluate pairwise allelic associations across *F8,* we calculated r^2 and D′, 2 commonly used measures of linkage disequilibrium (LD), for the subset of 12 functional-region SNPs that were variable among the unrelated white individuals in the VDG and those GAIT subjects whose parents were not enrolled for study (Table 2). The results were plotted with SOLAR (http://www.sfbr.org/solar/), a software package for genetic variance-components analysis (Figure 2).[60] D′ and r^2 were calculated for all SNP pairs.[61]

Genotyping

We genotyped remaining GAIT subjects (158 female, 172 male) for the subset of functional-region polymorphisms that were variable in white individuals of the VDG (Table 2) by generating and directly sequencing 1 strand of the 9 amplicons indicated in Table 1. Agencourt performed the initial PCR and sequencing, for which we performed quality control

in-house. To ensure that any polymorphisms identified in exons 21 and 22 could be evaluated by association analysis, 2 additional amplicons, 29B and 30B (Table 1), were generated from every GAIT subject and sequenced on both strands. We made all final genotype calls using PHRED output and a custom-written SAS_8.2 program (http://www.SAS.com) that generates a separate chromatogram for each polymorphism (custom-written program available upon request). By being compact and centering the variant between 2 10-bp flanks, these chromatograms standardized our manual genotype review. We paired forward- and reverse-strand sequences from the same individual to perform manual reviews but did not reveal the identity further than providing the file names, which contained logistic data such as plate, well address, amplicon, and direction. To resolve missing genotypes, we sequenced amplicons generated in-house on an ABI-3100. Genotypic data were analyzed with INFER, a component program implemented in the PESYS suite (http://www.sfbr.org/pages/genetics_projects.php?p=42), for violations of Mendelian inheritance; all disparities were reviewed and either corrected or excluded from further analysis.

Table 2. Characteristics, status, and allelic associations of *F8* polymorphisms with FVIII:C levels

			m-AFs, %†								*F8* variation databases‡		
Region, TU*	Protein*	GenBank*	T	W	AA	Ch	SEA	J	MI	SAA	HAMSTeRS[38,58]	VDR[53]	dbSNP[59]
Promoter													
−825G>A	—	49785830	0.9	NP	7.7	NP	NP	NP	NP	NP	NF	F8-002289	NF
−824G>A	—	49785831	0.9	NP	7.7	NP	NP	NP	NP	NP	NF	F8-002290	NF
−493G>A	—	49785832	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-002621	rs4898404
−385 A_4>A_5	—	49785833	0.9	NP	NP	NP	NP	NP	25.0	NP	NF	NF	NF
−287T>C	—	49785834	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	NF	NF
Intron 2													
25610G>A	—	49785835	0.5	NP	4.5	NP	NP	NP	NP	NP	NF	F8-028722	NF
Intron 3													
25865G>A	—	49785836	0.5	NP	4.2	NP	NP	NP	NP	NP	NF	F8-028977	NF
25885G>C	—	49785837	0.5	NP	4.2	NP	NP	NP	NP	NP	NF	F8-028997	NF
29567C>T	—	49785838	0.5	NP	NP	NP	NP	NP	NP	12.5	NF	NF	NF
Intron 4													
29854T>C	—	49785839	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	NF	NF
Intron 5													
35518C>G	—	49785840	0.9	NP	8.0	NP	NP	NP	NP	NP	NF	NF	NF
Intron 6													
53034A>G§	—	49785841	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF
Exon 7													
53206G>T§	*W255C*	49785842	0.5	NP	NP	NP	6.7	NP	NP	NP	NA	NF	NF
Intron 7													
55938C>A§	—	49785843	0.5	NP	NP	NP	6.7	NP	NP	NP	NF	NF	NF
56010G>A§	—	49785844	9.6	6.4	26.9	7.7	NP	12.5	12.5	25.0	+	F8-059122	rs7058826
Exon 8													
56113G>A§	A343A	49785845	0.5	NP	3.8	NP	NP	NP	NP	NP	+	F8-059225	rs1800289
Intron 9													
61534T>C	—	49785846	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-064646	rs5986899
Exon 10													
61620G>A	*R484H*	49785847	0.9	NP	8.3	NP	NP	NP	NP	NP	+	NF	NF
Intron 13													
75215G>A	—		0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-077273	rs5987069
90948A>G	—	49785848	0.5	NP	NP	NP	NP	NP	NP	12.5	NF	NF	NF
Exon 14													
91317A>G	*R776G*	49785849	0.5	NP	NP	7.7	NP	NP	NP	NP	NF	NF	rs2228152
92555C>T§	H1188H	49785850	0.5	NP	4.3	NP	NP	NP	NP	NP	NF	F8-095667	NF
92714C>G§	*D1241E*	49785851	14.7	7.7	72.7	7.7	NP	12.5	12.5	25.0	+	F8-095826	rs1800291
92798A>C§	S1269S	49785852	7.7	6.3	3.8	7.7	20.0	NP	12.5	25.0	+	F8-095910	rs1800292
92927G>A	K1312K	49785853	0.5	NP	NP	NP	6.7	NP	NP	NP	NF	NF	NF
93401C>G	V1470V	49785854	0.5	NP	4.5	NP	NP	NP	NP	NP	NF	NF	NF
93434G>A	P1481P	49785855	0.9	NP	9.1	NP	NP	NP	NP	NP	NF	NF	NF
Intron 15													
116434C>T	—	49785856	0.5	NP	NP	7.7	NP	NP	NP	NP	NF	NF	NF
Intron 18													
118909T>A§	—	49785857	20.3	14.6	84.6	7.7	14.3	NP	42.9	37.5	+	F8-122021	rs4898352
Intron 19													
120776T>C§	—	49785858	24.5	17.3	75.0	7.7	16.7	NP	37.5	50.0	+	F8-123888	rs4074307
Intron 22													
158352C>T	—	49785859	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF
158368T>C	—	49785860	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	F8-161500	NF
158635C>T	—	49785861	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-161767	rs5987054
158777A>G	—	49785862	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF
158820C>T	—	49785863	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-161952	rs5987053
159087G>A	—	49785864	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF

Table 2. Characteristics, status, and allelic associations of *F8* polymorphisms with FVIII:C levels (continued)

Region, TU*	Protein*	GenBank*	m-AFs, %†								*F8* variation databases‡		
			T	W	AA	Ch	SEA	J	MI	SAA	HAMSTeRS[38,58]	VDR[53]	dbSNP[59]
Intron 23													
159874G>A§	—	49785865	1.4	1.4	3.8	NP	NP	NP	NP	NP	NF	F8-163006	NF
Intron 24													
162013G>T§	—	49785866	3.3	5.0	NP	NP	NP	NP	NP	NP	NF	F8-165145	NF
Exon 25													
162161A>G§	*M2238V*	49785867	1.8	NP	15.4	NP	NP	NP	NP	NP	+	F8-165293	rs17051967
Intron 25													
162475T>C§	—	49785868	4.1	5.6	NP	NP	NP	NP	NP	12.5	NF	F8-165607	NF
3′ UTR													
185156C>T§	—	49785869	0.5	NP	3.8	NP	NP	NP	NP	NP	NF	F8-188288	rs5986887
186341G>A§	—	49785870	0.5	NP	NP	NP	NP	12.5	NP	NP	NF	NF	NF
186506A>G§	—	49785871	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF
186602C>T§	—	49785872	0.5	0.7	NP	NP	NP	NP	NP	NP	NF	NF	NF
186799G>A§	—	49785873	24.8	17.4	76.9	NP	13.3	NP	50.0	50.0	+	F8-189931	rs1050705
3′ DNA													
186987T>G§	—	49785874	0.9	NP	9.1	NP	NP	NP	NP	NP	NF	F8-190119	NF
187064T>C§	—	49785875	1.8	NP	15.4	NP	NP	NP	NP	NP	NF	F8-190196	NF

Except for the mild hemophilic missense mutation W255C, the coding region variants in italics represent ns-SNPs that encode the amino-acid substitutions listed.

T indicates total; W, the 68 SW and 18 WA subjects, which were considered together; HAMSTeRS, Hemophilia A Mutation, Structure, Test, and Resource Site; VDR, Variation Discovery Resource (UW-FHCRC); dbSNP (NCBI); —, polymorphisms located outside of the *F8* coding sequence; NP, not polymorphic; NF, not found; NA, not applicable; and +, a previously known polymorphism.

*Polymorphisms are designated by genic region; nt alleles and position in the TU (start site and adjacent 5′ base are designated as nt 1 and −1, respectively); amino acid alleles and position in the mature plasma protein; and GenBank number for dbSNP submission.

†Estimated from subjects in either the total VDG or each of its racial groups separately. Genotypic data were not complete for all variants and resulted in denominators that varied from the maximum number of distinct X-chromosomes (see "Patients, materials, and methods").

‡Public databases with *F8* polymorphisms.

§Twenty-one SNPs contained in the 9 amplicons generated to genotype GAIT subjects for the subset of 12 that were located in a functional-gene region and variable in at least white individuals.

Covariate selection and genotype association analyses

Measured variables available in GAIT include age, sex, smoking status, ABO genotype, OC status, BMI, DM status, and plasma levels of TC, high-density lipoprotein (HDL), LDL, very-low-density lipoprotein (VLDL), TG, lipoprotein-a, fibrinogen, VWF antigen (VWF:Ag), and FIX:C.[31,35] We used indicator variables to represent ABO genotypes, with OO as the reference level, and modeled separate mean FVIII:C levels for each.[35] In addition, smoking and OC status were represented with dichotomous variables that indicated any use versus the reference level of no use. We selected as covariates all measured variables associated with FVIII:C levels, since these factors could confound the relationship of interest if by chance their data were distributed differentially among the genotypes under evaluation. Since GAIT is a family-based study, we used likelihood-based variance-components analysis, as implemented in SOLAR, to account for the nonindependence between study subjects.[60] Further, we performed bivariate analyses with the endogenous factors and FVIII:C levels to investigate potential common genetic sources of variance.[62]

Since all *F8* polymorphisms were biallelic, containing a major (M) and minor (m) allele, with 3 genotypes in females (M/M, M/m, m/m) and 2 in males (M/Y, m/Y) (Table 2), we combined GAIT subjects by genotype into 3 groups (M/M and M/Y; M/m; m/m and m/Y) and used a dosage-compensation model in which hemizygous males had a gene effect equivalent to females homozygous for the same allele. We represented each group with the value of 1, 0, or −1, respectively, using an additive model in which the mean for heterozygous females is halfway between the means for females homozygous for the M and m alleles, and the means for male hemizygotes and females homozygous for the corresponding allele are equivalent. To evaluate the relationship between *F8* and FVIII:C level, we performed marginal measured-genotype association analysis, as implemented in SOLAR.[63] In the first analysis, we tested each polymorphism separately for genotype-specific differences in mean FVIII:C levels using an initial complex model that incorporated as covariates age, age^2, sex, age × sex, age^2 × sex, ABO genotype, and smoking and OC status. We reanalyzed all polymorphisms found to be suggestively associated with FVIII:C level($P \leq .10$) using a more complex model that included as covariates all potentially confounding measured variables. Finally, we performed additional marginal measured-genotype analyses in which paired endogenous factors that shared common genetic influences with FVIII:C level were excluded singly and jointly. This strategy allows for potential individual actions without enforcing the same mechanistic effect on every polymorphism. For these analyses, we used linear splines with knots at 15 and 50 years (ie, approximately corresponding to puberty and menopause) to represent age,[64-66] allowing a more flexible control of confounding than either a linear or a quadratic representation.

Results

Multiple unknown *F8* polymorphisms identified

To determine whether *F8* is a determinant of FVIII variability, we first identified candidate functional polymorphisms in 222 potentially distinct alleles contributed by the 137 unrelated VDG subjects. We examined 41 amplicons containing 1195 bp of promoter, 7054 bp of the 7056 bp of total coding sequence, all 1953 bp of untranslated exonic sequence, 4551 bp out of the approximately 5000 bp of total junctional-intronic sequence, approximately 2.3% of all deep intronic sequences (172 921 bp total), and 309 bp of 3′-flanking genomic DNA (Figure 1).[46,57] These amplicons contained 19 157 bp and included every known functional region except 2 bp of exon 22, the 5′ splice junction (SJ) of intron 22, and 78% to 90% of the SJ sequences from the 3′ ends of introns 20 and 21 and the 5′ end of intron 21 (Table 1). By analyzing the approximately 188 kb of genomic DNA between the forward and

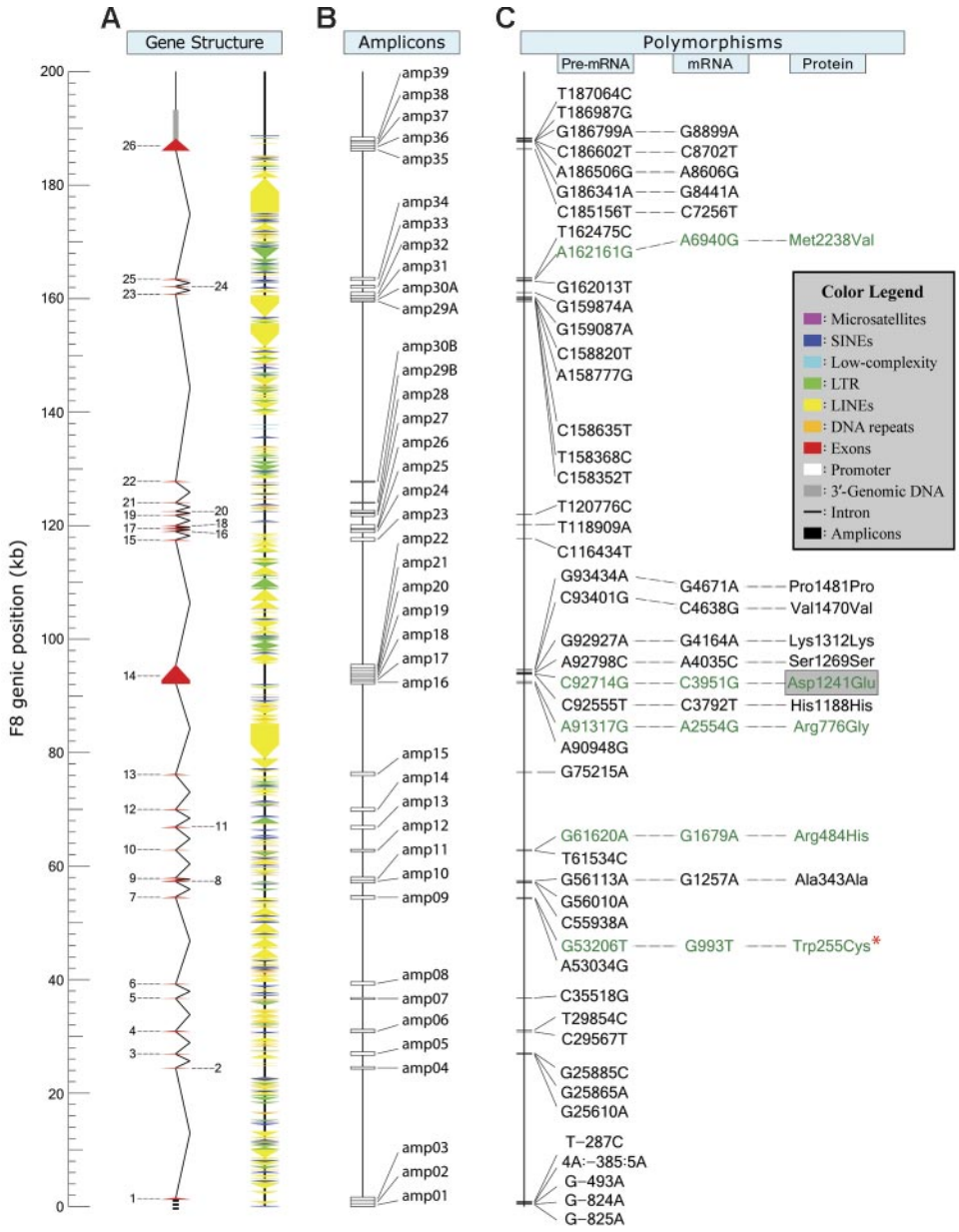


Figure 1. Genomic structure and common polymorphisms of the human F8. (A) The hg17 sequence used as a reference for *F8* spans approximately 286 kb and contains the approximately 186-kb structural locus and 50 kb of contiguous genomic-DNA flanking its 5′ and 3′ ends. *F8* consists of 26 exons (red triangles) and 25 introns (black lines). The approximately 1.2 kb of promoter and approximately 0.3 kb of 3′ genomic segments scanned for polymorphisms are indicated. The numerous repetitive elements (RepeatMasker) are indicated by different-colored triangles. (B) The 41 amplicons (amp) generated for variation detection and genotyping are indicated by open boxes. (C) The 47 variants are designated by their M- and m-alleles and position in the *F8* TU with respect to its major start site, which is indicated as nt 1 and corresponds to base 465469 in the reverse compliment of NG_005114. Promoter SNPs are located upstream of nt −1, which corresponds to base 465468 in the reverse compliment of NG_005114. The 17 exonic SNPs, which include 11 in coding sequence, are also designated by their nt and amino acid position in the polyadenylated mRNA and mature circulating FVIII protein, respectively. Finally, the 4 ns-SNPs and Trp255Cys, which encode the amino acid substitutions indicated, are shown in green. 53206G>T is denoted by a red asterisk, as it represents a known missense mutation encoding Trp255Cys, which was found originally in a male Chinese patient with mild hemophilia A. In the present study, the minor T-allele (Cys^{255}) was found in an asymptomatic heterozygous SEA female.

reverse primers of the first and last amplicons, respectively, with RepeatMasker (http://repeatmasker.org), we found that approximately 31% of *F8* is composed of nonrepetitive sequence, of which approximately 31% was covered in our scan.

We identified 47 variable sites (Figure 1) including 45 SNPs and a 1-bp insertion/deletion polymorphism (INDEL). The m allele of 53206G>T, which was carried by a female SEA subject and found in only 1 of the 222 X-chromosomes examined, encodes W255C, a missense mutation previously identified in Ch patients with mild hemophilia A.[38] These variants, of which 18 were unknown (SeattleSNPs[53]; The Single Nucleotide Polymorphism Database of Nucleotide Sequence Variation[59]; The Haemophilia A Mutation, Structure, Test and Resource Site[58]; Table 2), were located in all functional genic regions. Four coding sequence polymorphisms were nonsynonymous SNPs (ns-SNPs) encoding the amino-acid substitutions R484H, R776G, D1241E, and M2238V (Figure 1). Although when estimated in the combined VDG (Table 2) m-allele frequencies (m-AFs) ranged from 0.5% to 24.8%, most *F8* polymorphisms were singletons with infrequent and/or racially restricted minor alleles.

m-AFs vary with race

As the first step to investigate whether heritable FVIII determinants may vary across populations, a small number of nonwhite subjects from 6 racial groups were also studied. Specifically, we resequenced the same *F8* regions in 51 individuals (31 female, 20 male) including 16 AA, 10 Ch, 10 SEA, 5 J, 5 MI, and 5 SAA subjects. While the m-AFs of a few polymorphisms (eg, 92798A>C) were similar across different races, most varied substantially. Indeed the m-alleles of 38 polymorphisms were found in only 1 racial group. Only 16 of the 46 polymorphisms were variable in either WA or SW subjects, or both, despite examining 144 white X-chromosomes, approximately 6 times more than the number from African

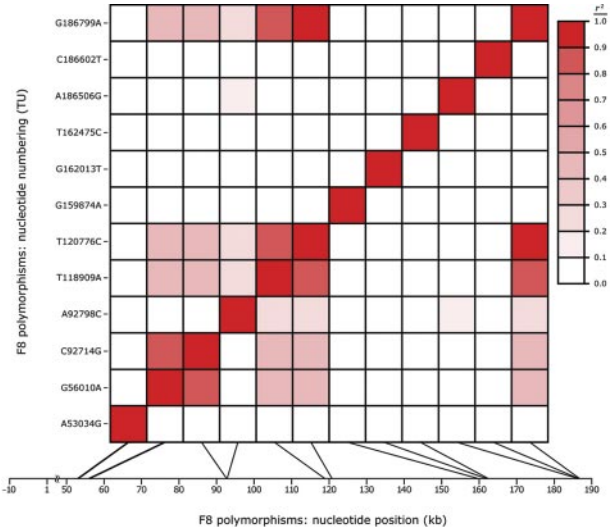


Figure 2. Pattern and degree of LD across *F8*. Allelic associations among the 12 functional region *F8* polymorphisms variable in at least white VDG subjects were estimated pairwise, based on allele frequencies in the GAIT founders, by calculating both D′ and r^2. While only r^2 values are displayed here, similar estimates of LD were obtained using D′. The color legend semiquantitatively illustrates r^2; white indicates linkage equilibrium ($r^2 = 0$), and red indicates complete LD ($r^2 = 1$). In order to reveal the genic location of each variation with respect to both its position in the transcription unit (y-axis) and physical spacing along the gene, based on hg17 (x-axis), the graph is plotted symmetrically such that the information in analogous positions above and below the diagonal pairwise line-of-identity are identical.

Americans, the second most abundantly sampled racial group studied. Moreover the m-alleles for 30 of the 38 potentially racially restricted polymorphisms were found in a nonwhite racial group. Twenty-two of these nonwhite racially restricted polymorphisms were variable only in individuals of African descent; of these, the m-AF was greater than 5% for 8 and greater than 10% for 2 (162161A>G, 187064T>C). Furthermore, 2 of the 4 ns-SNPs, 61620G>A (R484H) and 162161A>G (M2238V), were AA restricted. Because the m-allele of 91317A>G (R776G) was found in a single Ch individual, 92714C>G (D1241E) is the only white ns-SNP. Although 8 SNPs were found in more than 1 racial group, their m-AFs varied widely. For example, the G-allele of 92714C>G was less frequent in the combined VDG and among individuals from each race separately, except those of African descent; in AA subjects it was actually the M-allele with a frequency of approximately 73% (Table 2). A similar pattern of differences in m-AF was observed for 3 additional SNPs (118909T>A, 120776T>C, and 186799A>G).

Low LD across *F8*

To survey LD across *F8* we evaluated the 12 functional-region SNPs that were variable in (at least) the white subset of VDG subjects and/or the GAIT members whose parents were not enrolled in GAIT because the nonwhite populations were insufficiently sampled ($n \leq 16$) to accurately estimate race-specific AF. We found a low overall degree of LD, whether using r^2 (Figure 2) or D′ (not shown), with 6 sites being in linkage equilibrium with all other variants. Nevertheless, 120776T>C and 186799G>A exhibited high LD. The alleles of 92714C>G, which encodes the only white ns-SNP (D1241E), revealed moderate LD ($0.4 \leq r^2 < 0.5$) with 118909A>T, 120776C>T, and 186799A>G and high LD ($r^2 = 0.83$) with 56010G>A (Figure 2).

FVIII:C level is broadly variable in GAIT

Genomic DNA from 394 GAIT subjects (213 female, 181 male) with duplicate FVIII:C measures were available for this study. Data from 2 females, with FVIII:C levels greater than 4 standard deviations (SDs) beyond the mean, were excluded from further analysis. FVIII:C levels in the remaining 392 subjects exhibited a mean (± SD) of 150.7 IU dL^{-1} (± 52.2 IU dL^{-1}) and a greater than 6-fold concentration range (47-338 IU dL^{-1}).

F8 is a modest determinant of FVIII variance

Before investigating the influence of *F8* on FVIII, we attempted to identify variables for which it is necessary to control to avoid a potential bias. Sex, age, smoking status, ABO genotype, DM status, BMI, OC status, and plasma levels of fibrinogen, VWF:Ag, FIX:C, TC, LDL, VLDL, and TG were the measured variables available for this analysis. Although plasma C-reactive protein (CRP) concentration, the most commonly used measure of inflammation, was not assayed in GAIT, fibrinogen level may be a suitable surrogate marker.[67-69] As GAIT subjects are SW, we investigated the relationship between FVIII:C levels and the subset of white *F8* polymorphisms. Although the m-allele for 16 of the 45 SNPs were found in at least 1 white VDG subject (Table 2), 4 (including 158352C>T, 158368T>C, 158777A>G, and 159087G>A) were located more than 500-bp upstream from the 3′ SJ of intron 22, a region not known to be essential for *F8* function in vivo (Figure 1). Because LD across *F8* was irregular and low overall (Figure 2), we genotyped all 12 functional-region white SNPs in GAIT and evaluated their influence on FVIII:C levels separately using marginal measured-genotype association analysis. The results from our initial analyses, in which we modeled as covariates only age, age^2, sex, age × sex, age^2 × sex, ABO genotype, and smoking and OC status, demonstrated significantly different genotype-specific mean FVIII:C levels ($P < .05$) for both 56010G>A, a noncoding SNP located 27 nucleotides upstream from the 3′ SJ of intron 7, and 92714C>G, the only white ns-SNP, which is located in exon 14 and substitutes glutamate for aspartate at residue 1241 in the B domain (Figure 1). Although both SNPs represent potential FVIII QTNs, we are unable to decipher which is the true functional variant from this study because their alleles were strongly associated within GAIT founders (Figure 2).

To explore further the possibility that either 92714C>G or 56010G>A differentially influences FVIII:C level, we performed additional marginal measured-genotype association analyses using more complex models. We chose to present only the results for 92714C>G because in addition to being in high LD with 56010G>A, it encodes the only nonhemophilic amino-acid substitution (D1241E) found in white individuals and had more complete data available (ie, 6 subjects had missing genotypes for 56010G>A). In the final analysis of D1241E (92714C>G) we incorporated as covariates all measured variables associated with FVIII:C level to control for confounding. In a separate bivariate analysis of FVIII:C with both FIX:C and VWF:Ag, we found potential overlap in the sets of (unknown) genes that may affect their trait levels. Thus we present the 4 analyses in Table 3. We excluded FIX:C (model 3), VWF:Ag (model 2), and the pair jointly (model 1), such that they were nested within model 4. Results from model 1 demonstrated a significant association between D1241E and FVIII in which each D-allele additively increased the mean FVIII:C level by 14.3 IU dL^{-1} ($P = .016$).

Table 3. Multicovariate analysis of 92714C>G (D1241E) and FVIII:C levels in the GAIT population

	Model 1		Model 2		Model 3		Model 4	
Covariate*	β†	*P*‡	β†	*P*‡	β†	*P*‡	β†	*P*‡
Age, y	−4.12	.089	−3.43	.178	−3.75	.065	−3.33	.090
Sex	4.43	.364	6.45	.187	13.72	.001	13.41	.001
Age × sex	−1.33	.729	−4.93	.250	−3.37	.322	−3.73	.257
A15§	3.62	.152	3.00	.258	3.44	.102	3.03	.136
A15 × sex	2.12	.596	5.62	.206	3.88	.272	4.14	.225
A50§	3.48	< .001	3.13	< .001	1.44	.013	1.58	.005
A50 × sex	−2.78	.004	−2.30	.015	−1.48	.058	−1.42	.059
OC status‖	−23.82	.056	−23.80	.080	−11.22	.298	−15.10	.148
Smoking status‖	−11.45	.021	−13.48	.008	−5.84	.153	−8.27	.038
AA¶	25.85	.002	27.61	.001	10.36	.135	11.47	.084
AB¶	26.03	.109	35.41	.047	11.23	.444	9.91	.482
AO¶	21.80	< .001	15.78	.006	3.75	.447	2.71	.568
BO¶	26.91	.002	29.53	.001	15.56	.032	16.97	.015
DM status‖	−6.11	.618	−25.36	.101	−4.01	.746	−9.45	.470
Fibrinogen level, IU dL^{-1}	13.56	< .001	9.17	.012	2.55	.400	0.37	.901
BMI, kg m^{-2}	0.22	.718	−0.46	.471	0.85	.085	−0.05	.926
VLDL level, mM	45.58	< .001	22.49	.082	21.12	.039	8.60	.398
FIX:C level, IU dL^{-1}	N/A	N/A	0.56	< .001	N/A	N/A	0.45	< .001
VWF:Ag level, IU dL^{-1}	N/A	N/A	N/A	N/A	0.73	< .001	0.66	< .001
D1241E#	14.33	.016	11.99	.049	7.26	.142	6.10	.199

For model 1, n = 361; for model 2, n = 313; for model 3, n = 307; for model 4, n = 306.

*Endogenous, environmental, and genetic variables included in the marginal measured-genotype association analysis of the relationship between D1241E and FVIII:C levels in GAIT.

†The estimated effect (β) of a given covariate or each D-allele of the FVIII D1241E polymorphism (*F8* 92714C>G) on FVIII:C levels.

‡The *P* value for each β.

§The coefficient for splines with knots at 15 and 50 years.

‖Status (yes/no) indicates either any use (smoking and oral contraception) or the presence of diagnostic criteria for DM.

¶The 4 ABO genotypes, other than OO (reference genotype), found in GAIT subjects.

#D1241E genotypes were grouped and represented with indicator variables as follows: 1 (D/D and D/Y); 0 (D/d); and −1 (d/d and d/Y).

Since data were not available for FIX:C and VWF:Ag levels from all GAIT subjects, including them as covariates resulted in smaller sample sizes for the analyses using models 2-4. However, the results of these analyses, when restricted to the 306 subjects with complete data, showed similar trends, suggesting that the effects of adding FIX:C and/or VWF:Ag were not limited to the loss of subjects excluded due to missing data. For instance, when the 306 subjects in model 4 were reanalyzed using model 1, each D-allele was estimated to additively increase the mean FVIII:C level by 11.8 IU dL^{-1} ($P = .042$).

As expected, TG and TC levels were nearly colinear with VLDL and LDL levels, respectively, in GAIT (not shown). Furthermore, the Pearson correlation coefficients for these paired physiologically related lipid parameters were high whether we analyzed all GAIT subjects (and did not account for nonindependence between observations) or only the (independent) founders ($\rho \geq 0.95$; $P < .001$). Finally, the "effect" of D1241E on FVIII:C level was found to be similar whether we substituted TG levels for VLDL levels or TC levels for LDL levels as covariates in our analyses (not shown). Since removing both LDL levels and TC levels had a negligible impact upon the estimated effect of D1241E on FVIII (14.0 IU dL^{-1}; $P = .019$), these variables do not appear to confound the relationship under study[70] and therefore were not included in our final analysis (Table 3). In contrast, removal of both VLDL and TG levels resulted in a noticeable change (not shown). While the choice between using TG or VLDL levels was arbitrary and the results were comparable, we present the findings obtained with VLDL level as a covariate (Table 3), since fewer subjects had missing data for this measurement.

Discussion

We performed this study to identify genetic determinants contributing to the broad normal range for FVIII:C level,[5,6] a quantitative trait that influences thrombosis risk,[35] both venous[1-3] and arterial.[2,4] Despite involving a substantial heritable component,[31] functional allelic variants affecting the interindividual variability in this trait have been identified in only 1 gene.[22] While *F8* represents an obvious candidate, the negative findings from previous studies using linkage analysis[36,37,39,71] have led some investigators to conclude that the encoding structural locus is not a FVIII QTL.[71] Linkage studies are in general, however, less powerful than investigations based on association analysis, which usually directly examine potential functional variants, or polymorphisms in high LD with nongenotyped functional sites, especially for X-linked loci like *F8* that presently can be evaluated only by single-marker linkage analysis. Although *F8* has been the focus of numerous candidate gene studies, no polymorphisms were found within the functional regions examined in these investigations.[25,41-43,45] Nevertheless, *F8* should not be excluded as a possible FVIII QTL because less than 25% of all known functional regions, and therefore less than 2.5% of the entire structural locus, were scanned for variants in these studies. Despite the recent discovery of approximately 100 *F8* polymorphisms by SeattleSNPs[53] in a high-throughput resequencing study that examined 71 X-chromosomes from 47 unrelated individuals representing 2 racial groups,[53] these potential FVIII determinants could not be evaluated since no phenotypic data were available.

As our first step to investigate *F8,* we resequenced all known functional regions in a collection of 222 potentially distinct alleles from 137 unrelated nonhemophilic individuals representing 7 racial groups. We identified 47 variants distributed throughout *F8,* including 45 SNPs, a 1-bp INDEL polymorphism, and the mild hemophilic missense mutation W225C, found in a single asymptomatic heterozygous female (Table 2). Despite their location in 1 of the most extensively investigated human loci, 18 variants were previously unknown. Because our study, together with the SeattleSNPs scan,[53] identified 119 different polymorphisms (Table 2), *F8* may be as variable as other human genes. Furthermore, because the 4 SNPs 61620G>A, 91317A>G, 92714C>G, and 162161A>G encode the nonsynonymous substitutions R484H, R776G, D1241E, and M2238V, respectively (Figure 1), wild-type FVIII is a variable protein in nonhemophilic populations and not monomorphic as long thought.[6] Although SeattleSNPs identified 98 polymorphisms,[53] only 27 were located in the regions examined in this study; the larger number of total variants found is likely due to the fact that approximately 40 kb of deep intronic sequence was investigated.

The distribution of FVIII:C levels in GAIT displayed the typical broad variability observed for this trait.[5,6,41] Because the degree of LD across *F8* was found to be weak overall (Figure 2), we separately evaluated the 12 functional-region SNPs variable in white individuals (Table 2) for potential contributions to FVIII variability by measured-genotype association analysis. Preliminary analyses (not shown), in which we accounted for the potential confounders age, sex, ABO genotype, and smoking and OC status as covariates, demonstrated that FVIII:C level was significantly associated ($P = .007$) with only 92714C>G (D1241E).[39] Scanavini et al[40] confirmed this association in a case-control study of women with idiopathic thrombophilia. Specifically, they reported a higher mean FVIII:C level in female subjects with the C/C genotype (D/D phenotype) compared with those with either C/G or G/G genotypes (D/E or E/E phenotypes). Because no other *F8* polymorphisms were (i) found in our preliminary analyses to be significantly associated with FVIII levels or the alleles of D1241E,[40] or (ii) investigated by Scanavini et al,[40] the structurally distinct proteins encoded by this ns-SNP are likely to be functionally distinct and contribute to FVIII:C variability. However, when reanalyzed only among GAIT founders (Figure 2), 92714C>G demonstrated a high LD ($r^2 = 0.84$) with 56010G>A, an SNP that had incomplete genotypic data in the preliminary analysis.[39] Since 56010G>A is located within a potentially functional segment of the 3′ SJ of intron 7 (ie, position −27) and was found to have a significant association ($P < .05$) with FVIII:C level when reanalyzed using complete data (not shown), we cannot, based on data from this study alone, unequivocally establish which of these variants, if either (see below), is the true QTN.

To evaluate further the relationship of these 2 SNPs with FVIII, we controlled for all potentially confounding variables available in GAIT. Support for an association of age with FVIII:C level is among the strongest.[7-10,12,72] Results from our analyses of FVIII:C level versus age are consistent with this whether assessed alone (not shown) or with covariates (Table 3). In a large cross-sectional study with a narrower and younger age range than GAIT subjects, Green et al[17] did not observe an association between age and FVIII levels. Because Miller et al[73] also found no association in a study of women of similar age, we used splines at 15 and 50 years to allow a less rigorous imposition on the data across the broad age range in GAIT (2-87 years). This approach allows flexibility in the effects of age on FVIII:C levels in the young and the old instead of modeling, for instance, a quadratic effect across both, in which the age data for older subjects influences the estimated response for the young and vice versa. The relationship between sex and FVIII:C level remains uncertain, with some studies demonstrating significant associations[7,9] and others not[8,18]; we did not find evidence of a genotype-by-sex interaction. The impact of OC status on FVIII:C level has not been established.[8,10,74,75] Although we found an apparently large effect for OC, comparable to that observed for ABO genotype (Table 3), oral contraceptives were used by only a small number of female GAIT subjects. Two large cross-sectional studies found modestly lower FVIII:C levels in smokers,[7,15] whereas other investigations observed no association with this variable.[8,17] We found that smokers have a lower mean FVIII:C level but could not characterize this relationship further because detailed smoking information is lacking in GAIT. Since FVIII is an acute-phase reactant, its plasma concentration may be transiently affected by various conditions, including inflammation.[69] Although CRP level was not measured, we incorporated the surrogate marker fibrinogen level as a covariate in our analyses (Table 3), similarly to Kreuz et al.[68]

There is strong support for an association of ABO blood type with both FVIII:C level[17-21] and VWF:Ag level.[24-26,28,29] Using linkage analysis, Souto et al[22] demonstrated that the functional ABO polymorphisms responsible for the antigens of this blood group are likely to directly influence the levels of both hemostasis traits. Until recently,[39,40] these polymorphisms were the only known QTNs for FVIII. Although not completely understood, the effect of ABO genotype or phenotype on these 2 proteins may involve the same mechanism. Indeed, in a bivariate analysis of FVIII:C and VWF:Ag levels, we found potential overlap in the sets of genes associated with variability in their plasma levels. While less studied, associations between FIX and FVIII have also been reported.[24] Since neither FIX nor FVIII functions outside of intrinsic Xase, regulatory mechanisms might exist that coordinate their levels to prevent a potential cellular resource misallocation. Consistent with this, our findings from a bivariate analysis demonstrated a potential overlap in the set of genes that affect FIX and FVIII levels. Including factors that share pleiotropic influences with the trait of interest as covariates decreases the power to detect gene variants that influence both phenotypes jointly but may increase power to detect those with unique influences on the focal trait. If *F8* variants are in part responsible for the observed pleiotropy between FVIII and FIX or VWF, then model 1, excluding variables that are genetically correlated with FVIII, provides the most power. In contrast, if unidentified loci contribute to their shared heritability, then model 4 is preferable because it incorporates the known correlates of FVIII that may confound analysis. Without knowing the genetic architecture underlying pleiotropy in these traits, we focus on the results from model 1—the most conservative from the perspective of identifying functional gene variants, the goal of the present study—which demonstrated an additive increase in mean FVIII:C level of 14.3 IU dL^{-1} per D-allele (Table 3). Scanavini et al[40] recently reported similar findings in that subjects with the C/C genotype (D/D phenotype) had a 19 IU dL^{-1} higher unadjusted mean FVIII:C level than the group of subjects with either a C/G or G/G genotype (D/E or E/E phenotype).

Despite strong allelic association between 56010G>A and 92714C>G, we predict that 92714C>G represents the true QTN because it encodes structurally distinct wild-type FVIII molecules

(D1241E) that represent the only 2 forms of the protein expressed by white individuals, the population sampled in GAIT. Furthermore, while 56010G>A is near the 3′ end of intron 7, it is not located in essential cis-elements of the 3′ SJ. However, genetic studies in other populations with different patterns of allelic associations and/or in vitro functional assays are necessary to exclude the possibility that 56010G>A, or a nongenotyped functional polymorphism also in high LD with 92714C>G, is the FVIII QTN. Since the B domain undergoes numerous posttranslational modifications that limit FVIII expression,[38,76,77] and yet represents a portion of the protein known to be dispensable for its procoagulant activity once in the circulation,[78-81] we hypothesize that 92714C>G (D1241E) modulates FVIII:C levels by differentially affecting the secretion of this molecule.[82,83] Because D1241E could also influence FVIII clearance and/or activation, elucidating the molecular basis for its effect on FVIII:C level will require in vitro investigations since genetic studies alone cannot differentiate between these biologically plausible alternative mechanisms.

Regardless of whether 92714C>G is the actual QTN or in strong LD with it, we estimate that *F8* accounts for approximately 10% of the total FVIII variability in GAIT (not shown). This is consistent with the lack of evidence for linkage of FVIII:C levels to X-linked microsatellites within or near *F8,* which were genotyped in prior genome-wide screens,[36,37] including 1 performed in GAIT,[39] as we would not expect to detect QTLs with effect sizes in this range. As both 92714C>G and 56010G>A were polymorphic in all racial groups studied except SEA (Table 2), this *F8*-based determinant may influence FVIII variability in these populations. Because GAIT is composed entirely of SW subjects, however, appropriately designed studies in other racial groups are necessary to determine whether our findings apply to nonwhite populations. Such studies are also required to define further the genetic architecture underlying human FVIII variability since 30 of the functional-region *F8* polymorphisms identified, including the 3 additional ns-SNPs R484H, R776G, and M2238V, were not evaluated as potential FVIII determinants because they were not variable in white individuals (Table 2). In summary, it is important to emphasize that we resequenced less than 15% of *F8* and not all GAIT founders were enrolled for study. Therefore, we did not examine entirely every potentially distinct *F8* allele segregated in GAIT. Thus, the possibility that this locus contains additional FVIII determinants, either as common polymorphisms[84] or rare variants,[85] since both have been shown to contribute substantially to complex trait variability at the population level, cannot be excluded.

Acknowledgments

We are grateful to all individuals who participated in this study and the physicians who enrolled them. We would also like to thank Cynthia Channell, Ming Shen, and Thuy Tran for excellent technical assistance; Art Thompson, Earl Davie, and Bruce Evatt for helpful discussions and/or reviewing the manuscript; and both Alexander (Sandy) Duncan and Pete Lollar for early support and mentoring.

This work was supported in part by the following grants: NIH-MH-59490 (J.B.); NIH-HL-70751 (L.A.); NIH-HL-71130 and NIH-HL-68016 (T.E.H.), Fondo Investigación Sanitaria (FIS)–02/0375, SAF2002-03449 (FEDER, Spanish Ministry of Science and Technology), FIS-RECAVA C03/01 (Fundació "La Caixa" and Fundació d'Investigació Sant Pau). J.M.S. and A.B. were supported by FIS-99/3048 and FIS-01/A046 (Fondo Investigación Sanitaria; Spanish Ministry of Health).

Authorship

Contribution: J.C.S., J.B., J.F., J.M.S., L.A., and T.E.H. contributed the study concept and design; J.C.S., J.F., and S.T.W. enrolled the study subjects; K.R.V., D.K.M., J.C.S., J.F., J.M.S., and T.E.H. acquired the data; K.R.V., D.K.M., D.M.W., A.B., K.F., J.P., T.S., J.B., S.P., W.D.F., L.A., and T.E.H. provided data analysis and interpretation; K.R.V., J.M.S., W.D.F., L.A., and T.E.H. drafted the manuscript; and K.R.V., D.K.M., M.K., W.D.F., L.A., and T.E.H. provided critical manuscript revision for important intellectual content.

Conflict-of-interest disclosure: We certify that we have no affiliation with or financial involvement in any organization or entity with a direct financial interest in the subject matter or materials discussed in the manuscript. All financial support for this research project is identified in "Acknowledgments."

K.R.V. and D.K.M. contributed equally to this work.

Correspondence: Tom E. Howard, Southwest Foundation for Biomedical Research, Department of Genetics, Rm 12.306, Slick-Urschel Bldg, San Antonio, TX 78245-0549; e-mail: thoward@sfbrgenetics.org.

References

1. Koster T, Blann AD, Briet E, Vandenbroucke JP, Rosendaal FR. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep-vein thrombosis. Lancet. 1995;345:152-155.
2. Bank I, Libourel EJ, Middeldorp S, et al. Elevated levels of FVIII:C within families are associated with an increased risk for venous and arterial thrombosis. J Thromb Haemost. 2005;3:79-84.
3. Kyrle PA, Minar E, Hirschl M, et al. High plasma levels of factor VIII and the risk of recurrent venous thromboembolism. N Engl J Med. 2000;343:457-462.
4. Cortellaro M, Boschetti C, Cofrancesco E, et al. The PLAT Study: hemostatic function in relation to atherothrombotic ischemic events in vascular disease patients. Principal results. PLAT Study Group. Progetto Lombardo Atero-Trombosi (PLAT) Study Group. Arterioscler Thromb. 1992;12:1063-1070.
5. Wells PS, Langlois NJ, Webster MA, Jaffey J, Anderson JA. Elevated factor VIII is a risk factor for idiopathic venous thromboembolism in Canada: is it necessary to define a new upper reference range for factor VIII? Thromb Haemost. 2005;93:842-846.
6. Thompson AR. Structure and function of the factor VIII gene and protein. Semin Thromb Hemost. 2003;29:11-22.
7. Conlan MG, Folsom AR, Finch A, et al. Associations of factor VIII and von Willebrand factor with age, race, sex, and risk factors for atherosclerosis: The Atherosclerosis Risk in Communities (ARIC) Study. Thromb Haemost. 1993;70:380-385.
8. Balleisen L, Bailey J, Epping PH, Schulte H, van de Loo J. Epidemiological study on factor VII, factor VIII and fibrinogen in an industrial population, I: baseline data on the relation to age, gender, body-weight, smoking, alcohol, pill-using, and menopause. Thromb Haemost. 1985;54:475-479.
9. Tracy RP, Bovill EG, Fried LP, et al. The distribution of coagulation factors VII and VIII and fibrinogen in adults over 65 years: results from the Cardiovascular Health Study. Ann Epidemiol. 1992;2:509-519.
10. Lowe GD, Rumley A, Woodward M, et al. Epidemiology of coagulation factors, inhibitors and activation markers: the Third Glasgow MONICA Survey, I: illustrative reference ranges by age, sex and hormone use. Br J Haematol. 1997;97:775-784.
11. Lindeberg S, Berntorp E, Nilsson-Ehle P, Terent A, Vessby B. Age relations of cardiovascular risk factors in a traditional Melanesian society: the Kitava Study. Am J Clin Nutr. 1997;66:845-852.
12. Yarnell JW, Sweetnam PM, Rumley A, Lowe GD. Lifestyle factors and coagulation activation markers: the Caerphilly Study. Blood Coagul Fibrinolysis. 2001;12:721-728.
13. Schlit AF, Grandjean P, Donnez J, Lavenne E.

Large increase in plasmatic 11-dehydro-TxB2 levels due to oral contraceptives. Contraception. 1995;51:53-58.

14. Chang SJ, Koh SB, Cha BS, Park JK. Job characteristics and blood coagulation factors in Korean male workers. J Occup Environ Med. 2002; 44:997-1002.
15. Geffken DF, Cushman M, Burke GL, Polak JF, Sakkinen PA, Tracy RP. Association between physical activity and markers of inflammation in a healthy elderly population. Am J Epidemiol. 2001; 153:242-250.
16. Cushman M, Yanez D, Psaty BM, et al. Association of fibrinogen and coagulation factors VII and VIII with cardiovascular risk factors in the elderly: the Cardiovascular Health Study. Cardiovascular Health Study Investigators. Am J Epidemiol. 1996;143:665-676.
17. Green D, Ruth KJ, Folsom AR, Liu K. Hemostatic factors in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Arterioscler Thromb. 1994;14:686-693.
18. Jeremic M, Weisert O, Gedde-Dahl TW. Factor VIII (AHG) levels in 1016 regular blood donors: the effects of age, sex, and ABO blood groups. Scand J Clin Lab Invest. 1976;36:461-466.
19. McCallum CJ, Peake IR, Newcombe RG, Bloom AL. Factor VIII levels and blood group antigens. Thromb Haemost. 1983;50:757.
20. Mohanty D, Ghosh K, Marwaha N, Kaur S, Chauhan AP, Das KC. Major blood group antigens: a determinant of factor VIII levels in blood? Thromb Haemost. 1984;51:414.
21. Orstavik KH, Magnus P, Reisner H, Berg K, Graham JB, Nance W. Factor VIII and factor IX in a twin population: evidence for a major effect of ABO locus on factor VIII level. Am J Hum Genet. 1985;37:89-101.
22. Souto JC, Almasy L, Muniz-Diaz E, et al. Functional effects of the ABO locus polymorphism on plasma levels of von Willebrand factor, factor VIII, and activated partial thromboplastin time. Arterioscler Thromb Vasc Biol. 2000;20:2024-2028.
23. Balleisen L, Assmann G, Bailey J, Epping PH, Schulte H, van de Loo J. Epidemiological study on factor VII, factor VIII and fibrinogen in an industrial population, II: baseline data on the relation to blood pressure, blood glucose, uric acid, and lipid fractions. Thromb Haemost. 1985;54: 721-723.
24. Cadroy Y, Daviaud P, Saivin S, Sie P, Boneu B. Distribution of 16 hemostatic laboratory variables assayed in 100 blood donors. Nouv Rev Fr Hematol. 1990;32:259-264.
25. Morange PE, Tregouet DA, Frere C, et al. Biological and genetic factors influencing plasma factor VIII levels in a healthy family population: results from the Stanislas cohort. Br J Haematol. 2005; 128:91-99.
26. Miller CH, Haff E, Platt SJ, et al. Measurement of von Willebrand factor activity: relative effects of ABO blood type and race. J Thromb Haemost. 2003;1:2191-2197.
27. Weiss HJ, Hoyer IW. Von Willebrand factor: dissociation from antihemophilic factor procoagulant activity. Science. 1973;182:1149-1151.
28. Weiss HJ, Sussman, II, Hoyer LW. Stabilization of factor VIII in plasma by the von Willebrand factor: studies on posttransfusion and dissociated factor VIII and in patients with von Willebrand's disease. J Clin Invest. 1977;60:390-404.
29. Tuddenham EG, Lane RS, Rotblat F, et al. Response to infusions of polyelectrolyte fractionated human factor VIII concentrate in human haemophilia A and von Willebrand's disease. Br J Haematol. 1982;52:259-267.
30. van Hylckama Vlieg A, van der Linden IK, Bertina RM, Rosendaal FR. High levels of factor IX increase the risk of venous thrombosis. Blood. 2000;95:3678-3682.
31. Souto JC, Almasy L, Borrell M, et al. Genetic determinants of hemostasis phenotypes in Spanish families. Circulation. 2000;101:1546-1551.
32. de Lange M, Snieder H, Ariens RA, Spector TD, Grant PJ. The genetics of haemostasis: a twin study. Lancet. 2001;357:101-105.
33. Kamphuisen PW, Houwing-Duistermaat JJ, van Houwelingen HC, Eikenboom JC, Bertina RM, Rosendaal FR. Familial clustering of factor VIII and von Willebrand factor levels. Thromb Haemost. 1998;79:323-327.
34. Vossen CY, Hasstedt SJ, Rosendaal FR, et al. Heritability of plasma concentrations of clotting factors and measures of a prethrombotic state in a protein C-deficient family. J Thromb Haemost. 2004;2:242-247.
35. Souto JC, Almasy L, Borrell M, et al. Genetic susceptibility to thrombosis and its relationship to physiological risk factors: the GAIT study. Genetic Analysis of Idiopathic Thrombophilia. Am J Hum Genet. 2000;67:1452-1459.
36. Soria JM, Almasy L, Souto JC, et al. A new locus on chromosome 18 that influences normal variation in activated protein C resistance phenotype and factor VIII activity and its relation to thrombosis susceptibility. Blood. 2003;101:163-167.
37. Berger M, Mattheisen M, Kulle B, et al. High factor VIII levels in venous thromboembolism show linkage to imprinted loci on chromosomes 5 and 11. Blood. 2005;105:638-644.
38. Kemball-Cook G, Tuddenham EG, Wacey AI. The factor VIII structure and mutation resource site: HAMSTeRS version 4. Nucleic Acids Res. 1998; 26:216-219.
39. Machiah DK, Viel K, Almasy L, et al. A common SNP in the factor VIII (f-VIII) gene encodes a conservative aspartate to glutamate substitution (Asp1241Glu) in the B-domain that influences f-VIII activity levels [abstract]. Blood. 2003;102: 55a. Abstract 181.
40. Scanavini D, Legnani C, Lunghi B, Mingozzi F, Palareti G, Bernardi F. The factor VIII D1241E polymorphism is associated with decreased factor VIII activity and not with activated protein C resistance levels. Thromb Haemost. 2005;93: 453-456.
41. Mansvelt EP, Laffan M, McVey JH, Tuddenham EG. Analysis of the F8 gene in individuals with high plasma factor VIII:C levels and associated venous thrombosis. Thromb Haemost. 1998;80: 561-565.
42. Brummer J, Groth J, Flayeh R, Wagener C, Jung R. Absence of mutations at the APC interacting sites of factor VIII in Caucasians. Thromb Haemost. 2002;87:170.
43. Hooper WC, Dilley A, Austin H, et al. Absence of mutations at APC cleavage sites Arg306 in factor V and Arg336, Arg562 in factor VIII in African-Americans. Thromb Haemost. 1998;79:236.
44. Kamphuisen PW, Eikenboom JC, Rosendaal FR, et al. High factor VIII antigen levels increase the risk of venous thrombosis but are not associated with polymorphisms in the von Willebrand factor and factor VIII gene. Br J Haematol. 2001;115: 156-158.
45. Roelse JC, Koopman MM, Buller HR, et al. Absence of mutations at the activated protein C cleavage sites of factor VIII in 125 patients with venous thrombosis. Br J Haematol. 1996;92:740-743.
46. Gitschier J, Wood WI, Goralka TM, et al. Characterization of the human factor VIII gene. Nature. 1984;312:326-330.
47. Figueiredo MS, Brownlee GG. cis-acting elements and transcription factors involved in the promoter activity of the human factor VIII gene. J Biol Chem. 1995;270:11828-11838.
48. McGlynn LK, Mueller CR, Begbie M, Notley CR, Lillicrap D. Role of the liver-enriched transcription factor hepatocyte nuclear factor 1 in transcriptional regulation of the factor V111 gene. Mol Cell Biol. 1996;16:1936-1945.
49. Bertina RM, Koeleman BP, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. Nature. 1994; 369:64-67.
50. Williamson D, Brown K, Luddington R, Baglin C, Baglin T. Factor V Cambridge: a new mutation (Arg306—Thr) associated with resistance to activated protein C. Blood. 1998;91:1140-1144.
51. Poort SR, Rosendaal FR, Reitsma PH, Bertina RM. A common genetic variation in the 3'-untranslated region of the prothrombin gene is associated with elevated plasma prothrombin levels and an increase in venous thrombosis. Blood. 1996;88:3698-3703.
52. Howard TE, Marusa M, Channell C, Duncan A. A patient homozygous for a mutation in the prothrombin gene 3'-untranslated region associated with massive thrombosis. Blood Coagul Fibrinolysis. 1997;8:316-319.
53. National Heart, Lung, and Blood Institute Program for Genomic Applications. SeattleSNPs. http://pga.gs.washington.edu. Accessed November 2004.
54. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred, I: accuracy assessment. Genome Res. 1998;8:175-185.
55. Ewing B, Green P. Base-calling of automated sequencer traces using phred, II: error probabilities. Genome Res. 1998;8:186-194.
56. Zerylnick C, Torroni A, Sherman SL, Warren ST. Normal variation at the myotonic dystrophy locus in global human populations. Am J Hum Genet. 1995;56:123-130.
57. Karolchik D, Baertsch R, Diekhans M, et al. The UCSC Genome Browser Database. Nucl Acids Res. 2003;31:51-54.
58. Wacey AI, Kemball-Cook G. The Haemophilia A Mutation, Structure, Test and Resource Site. http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm. Accessed November 2004.
59. National Center for Biotechnology Information, National Library of Medicine. The Single Nucleotide Polymorphism Database (dbSNP) of Nucleotide Sequence Variation. http://www.ncbi.nlm.nih.gov/SNP/. Accessed November 2004.
60. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet. 1998;62:1198-1211.
61. Zhao JH. 2LD, GENECOUNTING and HAP: computer programs for linkage disequilibrium analysis. Bioinformatics. 2004;20:1325-1326.
62. Almasy L, Dyer TD, Blangero J. Bivariate quantitative trait linkage analysis: pleiotropy versus coincident linkages. Genet Epidemiol. 1997;14:953-958.
63. Blangero J, Goring HH, Kent JW Jr, et al. Quantitative trait nucleotide analysis using Bayesian model selection. Hum Biol. 2005;77:541-559.
64. Ruppert D, Wand MP, Carroll RJ. Semiparametric Regression. Cambridge, NY: Cambridge University Press; 2003.
65. Greenland S. Dose-response and trend analysis in epidemiology: alternatives to categorical analysis. Epidemiology. 1995;6:356-365.
66. Motulsky HJ, Ransnas LA. Fitting curves to data using nonlinear regression: a practical and nonmathematical review. FASEB J. 1987;1:365-374.
67. Folsom AR, Pankow JS, Tracy RP, et al. Association of C-reactive protein with markers of prevalent atherosclerotic disease. Am J Cardiol. 2001; 88:112-117.
68. Kreuz W, Stoll M, Junker R, et al. Familial elevated factor VIII in children with symptomatic venous thrombosis and post-thrombotic syndrome: results of a multicenter study. Arterioscler Thromb Vasc Biol. 2006;26:1901-1906.

69. Reitsma PH, Branger J, Van Den Blink B, Weijer S, Van Der Poll T, Meijers JC. Procoagulant protein levels are differentially increased during human endotoxemia. J Thromb Haemost. 2003;1: 1019-1023.
70. Rothman KJ, Greenland S. Modern Epidemiology. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1998.
71. De Visser MC, Sandkuijl LA, Lensen RP, Vos HL, Rosendaal FR, Bertina RM. Linkage analysis of factor VIII and von Willebrand factor loci as quantitative trait loci. J Thromb Haemost. 2003;1: 1771-1776.
72. Lindeberg S, Berntorp E, Carlsson R, Eliasson M, Marckmann P. Haemostatic variables in Pacific Islanders apparently free from stroke and ischaemic heart disease: the Kitava Study. Thromb Haemost. 1997;77:94-98.
73. Miller CH, Dilley AB, Drews C, Richardson L, Evatt B. Changes in von Willebrand factor and factor VIII levels during the menstrual cycle. Thromb Haemost. 2002;87:1082-1083.
74. Meade TW, Brozovic M, Chakrabarth R, Howarth DJ, North WR, Stirling Y. An epidemiological study of the haemostatic and other effects of oral contraceptives. Br J Haematol. 1976;34:353-364.
75. Kadir RA, Economides DL, Sabin CA, Owens D, Lee CA. Variations in coagulation factors in women: effects of age, ethnicity, menstrual cycle and combined oral contraceptive. Thromb Haemost. 1999;82:1456-1461.
76. Toole JJ, Knopf JL, Wozney JM, et al. Molecular cloning of a cDNA encoding human antihaemophilic factor. Nature. 1984;312:342-347.
77. Vehar GA, Keyt B, Eaton D, et al. Structure of human factor VIII. Nature. 1984;312:337-342.
78. Langner KD, Bird RE, McCandliss R, et al. Synthesis of biologically active deletion mutants of human factor VIII:C. Behring Inst Mitt. 1988:16-25.
79. Eaton DL, Wood WI, Eaton D, et al. Construction and characterization of an active factor VIII variant lacking the central one-third of the molecule. Biochemistry. 1986;25:8343-8347.
80. Toole JJ, Pittman DD, Orr EC, Murtha P, Wasley LC, Kaufman RJ. A large region (approximately equal to 95 kDa) of human factor VIII is dispensable for in vitro procoagulant activity. Proc Natl Acad Sci U S A. 1986;83:5939-5942.
81. Burke RL, Pachl C, Quiroga M, et al. The functional domains of coagulation factor VIII:C. J Biol Chem. 1986;261:12574-12578.
82. Pittman DD, Tomkinson KN, Kaufman RJ. Post-translational requirements for functional factor V and factor VIII secretion in mammalian cells. J Biol Chem. 1994;269:17329-17337.
83. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 form a cargo receptor complex and interact with coagulation factor VIII in the early secretory pathway. J Biol Chem. 2005;280:25881-25886.
84. Soria JM, Almasy L, Souto JC, et al. Linkage analysis demonstrates that the prothrombin G20210A mutation jointly influences plasma prothrombin levels and risk of thrombosis. Blood. 2000;95:2780-2785.
85. Soria JM, Almasy L, Souto JC, Sabater-Lleal M, Fontcuberta J, Blangero J. The F7 gene and clotting factor VII levels: dissection of a human quantitative trait locus. Hum Biol. 2005;77:561-575.

Chapter 3

The C allele of the Factor VIII (*F8*) 092714 single nucleotide polymorphism is additively associated with increased Factor VIII activity levels in white women, but only weakly associated in black women.

3-I. Introduction

Factor VIII (FVIII) is a 2332 amino acid glycoprotein that serves, in activated form, as the co-factor to activated Factor IX (FIX), a vitamin K-dependent serine protease. Activated FVIII (FVIIIa) enhances the relative catalytic efficacy of FIXa by several orders of magnitude(3, 5, 9). Insufficient levels or absence of FVIII activity (FVIII:C), seen in the bleeding disorder hemophilia A, demonstrates that FVIII is indispensable to normal hemostasis. High FVIII:C levels, at the other end of the bleeding spectrum, are associated with the pathogenesis of both venous(131, 134, 139) and arterial thrombosis(139, 140).

Recent studies in non-hemophilic populations have found an association between FVIII:C levels and an aspartate (D) to glutamate (E) substitution at residue 1241 of the FVIII protein(141-143). Encoded, respectively, by the C and G alleles of nucleotide 092714 (rs1800291) of *F8*, the FVIII structural gene, this residue lies in the B domain. Importantly, heterotrimeric FVIIIa does not contain the B-domain, which is freed by the activation cleavages, and thus the B-domain is not directly material to FVIII activity. Indeed, Pittman et al.(144) reported that "B-domainless" recombinant FVIII had specific activity that was indistinguishable from wildtype FVIII. Further adding to the enigma of this association, D1241E, one of two non-synonymous SNPs of *F8* originally reported by Gitschier et al. in 1984(22), has also been implicated in hemophilia(145).

Since the *in-vitro* determination of activity involves a cell-free environment, hypothetical mechanisms potentially involving the detection of liberated B-domain, such as feedback involving secretion of stored FVIII or up-regulation of receptor-mediated

clearance, are not possible explanations for a functional role of the D1241E variant. Ruling out an effect of the rate of activation, one remaining plausible mechanism is an effect on the amount of circulating FVIII. Elucidation of the clearance pathways of FVIII is currently an active field in the study of the life cycle of this protein(146), but the known involvement of the B-domain in secretion(147) lends credence to a potential functional role of this variant. Another possibility, since *in-vitro* studies are lacking, is that this SNP may simply be in allelic association with the true functional variant in the populations studied to date, but may not be in other populations.

Given these uncertainties, further assessment of this association is important, especially in a different racial population. Black Americans are one such population, with the advantage that even a study of moderate size should have adequate numbers of subjects with each genotype of interest here (148, 149). Therefore, the goal of this investigation was to assess the association between FVIII:C levels and the C(092714)G SNP of *F8* in black and white women.

3-II. **MATERIAL AND METHODS**

3-II-A. Study Population

Dilley et al. (150) conducted a case-control study of unexplained uterine bleeding using a case definition comprised of diagnoses relating to menorrhagia (ICD-9 codes: 625.9, 626, 626.2, 626.6, and 626.9). The target was reproductive-age women between 18 and 45 years old who were members of a single group practice in Atlanta, Georgia. The investigators searched medical records between the two-year period from January 01,

1995 to January 01, 1997 and identified 580 women who had physician diagnoses that included at least one of these ICD-9 codes. They also identified and randomly selected an age-stratified sample of 500 women who did not have any of the qualifying ICD-9 codes in their medical records. By mail, they invited these 500 (control) women and all of the women meeting the case definition to participate in the study. One-hundred twenty-one women with menorrhagia and 123 control women enrolled. The exposure of interest was von Willebrand Factor (VWF) deficiency(151), however, the investigators measured additional coagulation factors to rule them out as potential causes of bleeding. Dilley et al. (150) provide full details regarding the design, execution, and results of this case-control study.

3-II-B. Genotyping

We used an Applied Biosystems TaqMan® SNP Genotyping Assay to genotype nucleotide 092714 of *F8* (rs1800291). We performed the genotyping on an Applied Biosystems 7900HT Fast Real Time PCR instrument and set the thermal cycling conditions in accordance with the pre-calculated TaqMan® assay conditions. The analytic controls consisted of a known sample for each genotype from a previous re-sequencing based study and a negative, no template, control. We included up to four replicates of each control per run and used the Applied Biosystems SDS 7900HT software (v2.2.2) to discriminate the alleles. A subsequent run to resolve missing or ambiguous determinations included a convenience set of samples for replication.

3-II-C. Statistical Analysis

We determined the frequency of the C allele and calculated the expected HWE frequency of the CC, CG, and GG genotypes, which are p^2, $2p(1-p)$, and $(1-p)^2$, respectively. Here p equals the frequency of the C allele. We then employed a chi-squared test to determine if HWE was violated, comparing the observed and expected genotype frequencies.

We compared the mean FVIII:C level by genotype. Within each race group, we compared the means of the women homozygous for major and minor alleles using two-sample t-tests, since means of these groups are expected to differ the most. We also performed a simple linear regression in which the C092714G genotype was scored as 1 for the CC genotype, as 0 for the CG genotype, and as -1 for the GG genotype. This coding scheme imposes an additive effect that not only orders the mean FVIII:C levels, but requires that the mean for the CG group lie midway between the means for the two homozygous groups, i.e. that the difference between the CG and GG means is equal to the difference between the CC and CG means. To adjust for potentially important covariates, we subsequently also used multiple linear regression.

Based on a priori considerations and availability of information, we were able to assess possible confounding by the following covariates: C092714G genotype score, VWF:Ag (IU/dL), FIX:C level (IU/dL), race, age, BMI, oral contraceptive use, exercise units, the maximum number of alcoholic drinks the respondents reported consuming during one episode in the last 12 months, number of cigarette smoked per day, reported history of high cholesterol, reported history of diabetes, ABO phenotype, and case-control status. We limited our focus to those subjects who reported either black or white race, and represented race with an indicator (dichotomous) variable. Exercise was coded

as the number of 20 minute units of exercise in the last month, thus one hour of exercise contributes 3 units, regardless of whether it occur during one day or over three days. The number of cigarettes smoked per day reflected reported current average consumption. Both high cholesterol and diabetes were dichotomous variables indicating whether the respondents reported a history of those conditions. The ABO phenotype was a series of indicator variables (A, B, AB) with the O phenotype as the referent group.

Preliminary analyses suggested that race might be an effect measure modifier, so we conducted regression analyses in each racial group separately. To limit the number of comparisons, we chose an *a priori* model (**Table 3-3)** to investigate the adjusted association between FVIII:C levels and the C092714G alleles.

We also performed several sensitivity analyses. In the first, we repeated the analyses using the *a priori* model in the combined black and white population. We then considered an adjustment to correct the unequal selection probabilities of cases and controls by applying differential weighting to the cases and controls. We chose to weight using the inverse probability of selection(152) based on three approximations of an estimated prevalence of menorrhagia of 13.3%(153). Finally, we also performed a control-only analysis. In the second, we expanded the model to control for additional covariates: the indicator variables for ABO phenotype, reported average number of cigarettes smoked per day, history of high cholesterol, the maximum number of drinks consumed in a 24 hour period, and the number of 20 minute periods of exercise in the last month. In the third, we investigated the extent to which influential observations, based on a deleted studentized residual (DSR) of 3.6 or greater, might impact the results by excluding them. Finally, we also examined three additional variations of the *a priori*

model: exclusion from the model of FIX:C levels, of VWF:Ag levels, and of FIX:C and VWF:Ag levels jointly. The appendix provides a discussion of the rationale, results and interpretations for these additional models. We used SAS® version 9.1 and perl to analyze the data. We performed all statistical procedures using SAS®.

3-III. RESULTS

The frequency (95% CI) of the G allele in black participants was 64.3% (55.7%, 72.9%) with 17, 51, and 51 having the CC, CG, and GG genotypes, respectively. The frequency (95% CI) of the G allele in white participants was 21.3% (11.6%, 31.1%) with 42, 23, and 3 having the CC, CG, and GG genotypes, respectively. Neither group departed significantly from the expected HWE distributions.

3-III-A. Main Findings

The mean (SD) FVIII:C level was 122 (44) IU/dL and the range was 51-382 IU/dL. As shown in **Table 3-1**, a monotonic relationship between the number of C alleles and FVIII:C levels is present in whites. In this racial group, the mean FVIII:C level among C/C homozygotes is greater than that among heterozygotes ($p = 0.046$); the number of white women with the G/G genotype is small ($n = 3$). Although the mean FVIII:C was not monotonic in blacks, two black women had FVIII:C levels that were more than 4 SD from the mean: 314 and 382 IU/dL. Both women had C/G genotypes; after excluding them the resulting mean (SD) for this group was 122 (37) IU/dL. A test for trend, i.e. a simple linear regression that models the genotype as an additive effect,

Table 3-1. The mean FVIII:C levels by race and genotype at the C097214G allele of the *F8* gene, which encodes the D1241E allele in the B domain, among the women of the CDC study of menorrhagia.

Race	092714 Genotype	N	FVIII:C level (IU/dL)				
			Mean	Std Dev	Median	Minimum	Maximum
Black	GG	51	123	36	122	51	190
	CG*	49	122	40	116	59	225
	CC	17	128	45	125	66	246
White	GG	3	99	12	95	89	113
	CG	23	103	28	94	59	158
	CC	42	120	36	115	64	212

(*) Two women who had FVIII:C levels of 314 and 382 IU/dL were excluded from this group. The mean (SD) prior to the exclusion was 131(59) IU/dL.

showed a statistically significant association in white women (slope = 14.7, p = 0.036), but little or no association in black women (slope = 1.76, p = 0.73).

3-III-B. Sensitivity Analyses

As described in the appendix in more detail, expanding the model to adjust for additional potential confounders did not meaningfully affect the results, suggesting that inclusion of additional covariates did not induce numeric instability in the estimates and that residual confounding by the additional factors was not important. Further, excluding one or two subjects, for whom the model may not have adequately adjusted or who had extreme values, whether due to measurement error or not, actually strengthened the associations. Not adjusting for VWF:Ag and/or FIX:C levels also strengthened the associations. For instance, when we did not adjust for VWF:Ag, the additive association for black and white women was 6.3 (-2.2, 14.8) and 15.3 (3.6, 27.1), respectively.

Results were similar, although confidence intervals somewhat wider, when we restricted to controls (Appendix).

3-III-C. Additional Adjustments

Most of the covariates assessed as potential confounders (**Table 3-2**) did not appear to be strongly associated with genotype. A possible exception, however, is VWF levels which increase with increasing numbers of C alleles, particularly in white women

Table 3-2. The characteristics of the 185 women in the CDC case-control study of menorrhagia genotyped at the *F8* 092714 nucleotide, which encodes amino acid 1241 in the B domain of FVIII.

	Black			**White**		
	G/G	**C/G**	**C/C**	**G/G**	**C/G**	**C/C**
N	51	49	17	3	23	42
VWF:Ag (IU/dL)	129.9 (41.5)	131.5 (50.1)	132.4 (50.4)	97.7 (15.2)	108.0 (32.1)	119.1 (42.0)
FIX:C (IU/dL)	153.0 (30.0)	139.6 (27.8)	148.6 (36.0)	176.0 (43.1)	132.0 (28.5)	145.5 (31.0)
Age (years)	34.3 (7.2)	36.2 (7.9)	32.7 (7.5)	36.0 (8.7)	33.4 (7.7)	36.8 (8.2)
BMI (kg/m^2)	29.4 (7.5)	29.2 (5.7)	30.9 (9.6)	23.9 (3.4)	24.6 (5.6)	25.9 (6.9)
Cigarettes/day	1.1 (5.8)	1.6 (4.3)	0.0 (0.0)	3.3 (5.8)	3.3 (9.5)	3.7 (7.4)
MaxDrinks*	1.2 (1.7)	1.6 (1.7)	1.1 (1.5)	3.3 (1.5)	3.0 (3.2)	2.5 (2.7)
Exercise units$^+$	15.1 (19.3)	16.8 (21.1)	12.4 (17.3)	30.0 (11.8)	27.0 (43.7)	23.3 (28.7)
Oral contraceptive use	11 (21.6%)	9 (18.4%)	5 (29.4%)	1 (33.3%)	5 (21.7%)	9 (21.4%)
History of diabetes	3 (5.9%)	0 (0.0%)	2 (11.8%)	0 (0.0%)	2 (8.7%)	3 (7.1%)
History of high cholesterol	9 (17.6%)	10 (20.4%)	4 (23.5%)	0 (0.0%)	1 (4.3%)	6 (14.3%)
ABO phenotype						
A	7 (13.7%)	11 (22.4%)	3 (17.6%)	0 (0.0%)	13 (56.5%)	15 (35.7%)
AB	5 (9.8%)	2 (4.1%)	1 (5.9%)	0 (0.0%)	1 (4.3%)	2 (4.8%)
B	14 (27.5%)	11 (22.4%)	3 (17.6%)	2 (66.7%)	2 (8.7%)	10 (23.8%)
O	25 (49.0%)	25 (51.0%)	10 (58.8%)	1 (33.3%)	7 (30.4%)	15 (35.7%)

*The maximum number of alcoholic drinks the respondents reported consuming during one episode in the last 12 months. $^+$The number of 20 minute units of exercise in the last month.

In multiple linear regression models to control for potential confounding by these factors, we found consistent results (Table 3-3). Among the white women, the mean FVIII:C level increased additively by 9.3 (0.3, 18.2) IU/dL with each additional C allele. Among

Table 3-3. Results of the multiple linear regressions of FVIII:C levels in the separate black-only and white-only populations.

Parameter	Black-only n = 117	White-only n = 68
Intercept	18.13 (-19.7, 55.91)	55.86 (21.80, 89.92)
C092714G score[a]	2.43 (-4.29, 9.15)	9.28 (0.33, 18.22)
VWF:Ag (IU/dL)	0.56 (0.46, 0.67)	0.62 (0.47, 0.76)
FIX:C level (IU/dL)	0.21 (0.02, 0.40)	0.19 (-0.03, 0.41)
Age (years)	0.36 (-0.30, 1.02)	-0.43 (-1.10, 0.24)
BMI (kg/m^2)	-0.36 (-1.15, 0.43)	-1.06 (-2.20, 0.08)
Oral contraceptive use	-6.76 (-18.8, 5.25)	-19.6 (-33.6, -5.73)
History of diabetes	5.09 (-18.8, 28.96)	17.89 (-2.64, 38.42)

(a) CC = 1, CG = 0, GG = -1

black women, the estimated increase in FVIII:C levels was smaller (2.4) and confidence intervals included the possibility of no association (-4.3, 9.2) IU/dL. FVIII:C levels increased linearly with VWF:Ag level with similar slopes (95% CI) of 0.6 (0.5, 0.7) IU/dL and 0.6 (0.5, 0.8) IU/dL among black and white women, respectively.

3-IV. DISCUSSION

This study supports an additive association between the C092714G SNP of *F8* and FVIII:C levels among white women. Of the three previous investigations of this locus, one had sufficient numbers of G/G females, but no additive pattern was discernable among the women of that study, even though the difference between the hemizygous males was significant and supportive of the current results. In this study, the frequency of the G allele among black women was 64.3%, ensuring adequate numbers of women with each of the three genotypes. Despite this, evidence of an additive effect, though consistent with an increased mean FVIII:C level associated with the C allele, was weak or absent, such that the confidence intervals included the possibility of no association. Replication, preferably with a larger sample size, is clearly required to document and support this association among black women.

Several factors may explain why the effect was not also definitively seen in the subset of black women of this study. First, the results could be spurious, resulting from differential measurement error or sampling variability. Second, the variability of FVIII:C levels in black populations was higher, meaning the "noise" may decrease the ability to detect a true effect. Third, this locus may not be functional, but its allelic association with the true functional locus may be stronger in the white than the black population.

Regardless of separate replications of the association between this locus and FVIII:C levels, consideration of the plausibility of the results is the true functional locus may be stronger in the white than the black population warranted, since the three known studies to date have measured activity but not antigen levels of FVIII, and the B domain is absent from the activated protein. Whereas the allele may affect the rate of activation

in the *in vitro* system in which the activity level was measured, it is not likely that the allele could affect the rate of thrombin cleavages to the extent necessary to explain the observed differences. Further, any potential feedback mechanism that might react to the presence of freed B domain is not possible within the *in vitro* system since most, if not all, of the cells that store and possibly release FVIII in response to detecting the freed B domain should be absent, thus ruling out an indirect affect on activity via an release of the protein during a clotting event. The B domain, however, may be involved in the clearance of intact, inactivated FVIII from the circulating blood. Any study of the D1241E allele in regard to this mechanism must include not only the genotype, but also determination of the extent of proteolysis of the B domain during transit through the secretory pathway is particularly relevant to confirm that this residue is in fact present.

If the allele is functional, then it might affect post-translational processing, secretion, storage, or the strength of non-covalent bonding of FVIII with VWF. The B domain is important to the secretion process(23, 144) and interacts with chaperon molecules within the endoplasmic reticulum-Golgi apparatus(154). The greatly reduced, but significant residual FV and FVIII activity, found in patients with the autosomal recessive disease, combined FV/FVIII deficiency (CF5F8D), suggests that efficiency of secretion can affect FVIII:C levels without altering the function of the protein(24). Thus, *in vitro* studies may be required to establish whether the E allele creates a bottleneck during post-translational modification or during interaction with chaperon molecules that might affect folding or secretion rates. In the meantime, a next logical step would include measurement of FVIII:Ag levels in addition to FVIII:C levels in observational studies of this allele.

This study had several limitations. Sample size was modest and, since we genotyped only one SNP, we could not address allelic associations (haplotypes). Bias could have affected results, although we were able to investigate several potential sources in our sensitivity analyses (Appendix). For instance, since menorrhagia is a bleeding disorder and FVIII:C level affects bleeding, selecting cases with a higher probability may have created a selection bias. In general, the sensitivity analyses suggested that these results are probably not importantly affected by this and the other sources of bias addressed, and may even be strengthened by some of the corrections. Important strengths of the study included measurement of VWF and FIX levels, inclusions two racial groups from the same group practice, and availability of important covariate data, such as use of oral contraceptives and history of diabetes.

Although not the focus of this study, we did find that mean VWF:Ag levels increased approximately additively with each additional C allele of the 092714 SNP of *F8* among white women (Table 3-2). Since the VWF gene lies on chromosome 12, it is not possible for allelic association to account for this. Suspicion, therefore, first falls on chance.

Regardless of the source, the association of VWF:Ag levels with the 092714 SNP potentially impacts the conclusions of this analysis. VWF levels are the most important known factor causally influencing basal FVIII levels, aside from deleterious mutations in either structural gene. FVIII:C levels increase linearly in association with VWF:Ag levels; the estimate in this study was approximately 0.6 IU/dL. If the distribution of VWF:Ag observed in white women (Table 3-2) was random, i.e. not caused by the 092714 genotype, then failure to control for VWF:Ag level could spuriously strengthen

the additive effect that this SNP had on FVIII:C levels. However, by unknown regulation posited in the appendix, FVIII:C levels may provide feedback, so that an increase in the FVIII:C level leads to a concordant increase in the VWF:Ag level to conserve resources required to synthesis and store these proteins. Under such mechanisms, the group of women with the higher mean FVIII:C levels would have a higher mean VWF:Ag level and the relationship observed in Table 3-2 would be expected, but perhaps not so pronounced. Although speculative, especially since VWF molecules are polymers of varying sizes such that individual subunits are in stoichiometric excess of FVIII molecules, these data may provide empirical suggestion for such an unknown mechanism. Our results, however, were largely robust to control: regardless of whether or not we controlled for VWF:Ag level, mean FVIII:C levels increased additively with each C allele.

In summary, these results lend credence to the functional involvement of the 1241 residue of the FVIII protein in variability of the FVIII:C levels. In addition, these analyses offer an intriguing perspective on the potential coordinate regulation of the levels of FVIII, FIX, and VWF. While replication of this finding in other black populations is necessary, future investigations using other racial groups should be worthwhile.

3-V. **APPENDIX** Sensitivity Analyses

3-V-A. Possible Selection bias

Cases and controls were selected in approximately equal numbers, so the selection probability for cases was higher than that of controls assuming the prevalence of

menorrhagia is close to published estimates, 13.3%(153). We therefore, weighted subjects in proportion to the inverse of the selection probability(152). We chose to calculate the weights for three prevalences: 5%, 10%, and 15%. The weight for cases was calculated as W = (Prevalence * $N_{control}$)/(N_{case}*(1-Prevalence)) and that for controls was 1. As shown in **Table 3-A-1**, the results are robust to the choice of prevalence, i.e. weights, suggesting that cases and controls are similar with respect to this association.

Table 3-A-1. The effects of adjustments for the potential bias resulting from unequal selection probabilities on the association between FVIII:C levels and the *F8* C(097214)G SNP.

Selection criteria	Case Weight (Prevalence)*	Black-only‡	White-only‡	Combined‡
Cases and controls	1.00 (0.50)	2.43 (-4.29, 9.15)	9.28 (0.33, 18.22)	3.98 (-1.43, 9.40)
Cases and controls	0.18 (0.15)	1.56 (-4.87, 7.99)	8.37 (-0.33, 17.08)	2.68 (-2.55, 7.90)
Cases and controls	0.11 (0.10)	1.45 (-4.94, 7.84)	8.30 (-0.34, 16.95)	2.51 (-2.68, 7.71)
Cases and controls	0.05 (0.05)	1.34 (-5.01, 7.69)	8.25 (-0.34, 16.84)	2.36 (-2.82, 7.53)
Controls only	0.00 (N/A)	1.23 (-7.30, 9.77)	8.21 (-3.05, 19.47)	2.21 (-4.70, 9.11)

(*)C092714G score (CC = 1, CG = 0, GG = -1); (+) Weight = (Prevalence * $N_{control}$)/(N_{case}*(1-Prevalence)); (‡) Adjusted for VWF:Ag (IU/dL), FIX:C level (IU/dL), Race(0/1, Combined only), Age (years), BMI (kg/m2), Oral contraceptive use (0/1), History of diabetes (0/1)

3-V-B. Control of confounding

We adjusted for additional covariates, in addition to those in Table 3-3, by first additionally controlling for the number of cigarettes smoked per day, history of high cholesterol, and the three indicator variables for ABO phenotype. As this did not meaningfully alter the results, we further adjusted for the maximum number of drinks consumed in a 24 hour period and the number of 20 minute periods of exercise in the last month. The estimated slopes (95% CI) were: 2.3 (-4.4, 8.9) and 10.2 (1.4, 18.9) for black

and white women, respectively, suggesting that results retained a similar level of stability with the larger number of covariates and achieved little meaningful, additional control of confounding.

3-V-C. Influential observations

We considered observations with deleted studentized residuals (DSR) equal to or greater than 3.6 to be potential influential points. Since the DSR has a t-distribution(155), we chose 3.6 over the conventional 2.0 level to account for the multiple testing inherent in examining the DSRs without specifying the subject of interest *a priori*. A simulation study suggested that this level was sufficiently stringent for the range of samples sizes in this investigation. We re-ran the regression analyses in which these points were excluded to investigate their impact on the results. In each regression of Table 3-3, we found one potential influential observation with DSRs of 4.3 and 4.0 in the black and white women, respectively. After exclusion, the associations strengthened: 4.1 (-2.2, 10.4) and 10.3 (2.3, 18.3) for the black and white women, respectively.

3-V-D. Confounders or potentially co-regulated factors

Since FVIII has a greatly reduced half-life when not in complex with VWF and its sole known function is as the co-factor to FIX, coordinate regulation of the levels of these three proteins may occur. For instance, thrombin will cleave free FVIII that is not being carried by VWF. The body may respond to these FVIII cleavage products by up-regulating VWF, either by producing new protein or by releasing stored protein, to "rescue" the FVIII being secreted. Analogously, the body may detect the remnants of

activated FVIII (FVIIIa) and up-regulate both FVIII and FIX production. In this case, detection of the freed B-domain of appropriately activated FVIII could indicate that coagulation reactions are consuming FVIII and, thus, the likely need for both FVIII and FIX. Additionally, detection of the dissociated A2 domain could indicate the dissociation of FVIIIa, potentially from the membrane bound FVIIIa/FIXa (tenase) complex. From the perspective of these analyses, we are unable to distinguish between basal and perturbed levels of the VWF and FIX proteins in the current study. If the subjects have basal levels, or we have adequately controlled for factors causing perturbation, then we may obscure the effect on the SNP by including either VWF or FIX in the model.

These considerations are relevant to our analyses as follows. If FVIII levels affect those of VWF and FIX through co-regulation or other mechanism, then simple analytic adjustment for them might well be inappropriate and could artifactually weaken the association. On the other hand, if that does not occur, then confounding may be present and adjustment appropriate and preferred.

As shown in **Table** 3-**A-2**, when we did not adjust for VWF:Ag the estimated, additive association between the SNP and FVIII:C levels was stronger. We cannot

Table 3-A-2. The effects of excluding FIX:C and VWF:Ag levels on the estimate for the effect of the *F8* 097214 locus.

	Full model*	**No FIX:C**	**No VWF:Ag**	**No FIX:C or VWF:Ag**
Combined	4.11 (-1.28, 9.51)	3.13 (-2.30, 8.57)	8.15 (1.20, 15.09)	6.60 (-0.54, 13.74)
Black women	2.45 (-4.24, 9.13)	0.98 (-5.73, 7.68)	6.27 (-2.23, 14.77)	4.02 (-4.68, 12.71)
White women	9.84 (0.82, 18.85)	9.42 (0.25, 18.59)	15.32 (3.59, 27.05)	14.89 (2.91, 26.88)

(*)C092714G score (CC = 1, CG = 0, GG = -1),VWF:Ag (IU/dL),FIX:C level (IU/dL),Race(0/1, Combined only),Cigarettes per day,Age (years),BMI (kg/m2),Oral contraceptive use (0/1), History of diabetes (0/1), History of high cholesterol (0/1), A phenotype (0/1),B phenotype (0/1),AB phenotype (0/1)

determine, with the data available to this investigation, whether this reflects coordination of the levels of these proteins or acute phase perturbation of the basal levels, but it offers an interesting perspective, nonetheless. Possibly, VWF level is partly a confounder, but may also partly be a consequence of the FVIII level; if so, a more complicated approach might be required; ultimate resolution of this issue may require results from in vitro studies as well.

Chapter 4

The association between Depression and plasma Factor VIII activity (FVIII:C) levels in reproductive aged women

4-I. Introduction

Depression is an important public health problem, with a lifetime prevalence of major depression in the general, non-institutionalized population estimated to be between 13%-16%(112-115), and an estimated 12-month prevalence for the period just prior to interview between 5.3%-6.6%(113, 114). Depression is associated with several chronic diseases. In particular, it has been associated with coronary artery disease(121, 156), myocardial infarction(157, 158), cerebrovascular disease(159), and stroke(160-163) (collectively, cardiovascular disease and accidents: CVDA).

Depression has also been associated with increased levels of Factor VIII activity (FVIII:C) in several epidemiologic studies(119, 120, 122) . For example, Kop et al.(119) employed a modified 10-item Center for Epidemiological Study Depression (CES-D) instrument(116) to assess depression in the absence of a physician diagnosis, and found that the depression score was significantly associated with FVIII:C levels.

Depression may cause increased Factor VIII activity, which can increase in response to mental activity(73), stress(74, 164), inflammation(84, 85), and exercise(69-71) underscoring that FVIII is an acute phase reactant. The latter three exposures bear resemblance to the clinical criteria for major depressive disorder (MDD)(111).

Since depression is a relatively common condition and increased Factor VIII levels can increase the risk of serious conditions such as CVDA(83, 98, 139, 165), further study of the association between depression and Factor VIIII is needed. The goal of this investigation is to determine whether FVIII:C levels are predicted by depression as measured by CES-D scores with the full 20 question instrument among black and white women. The use of data from a previous case-control study allows us to take advantage

of measured Factor VIIII levels and detailed covariate information in a predominantly black American population.

4-II. Materials and Methods

4-II-A. Study Population

Dilley et al. (150) conducted a case-control study of reproductive-age women between 18 and 45 years old. The case definition, comprised of diagnoses relating to menorrhagia (ICD-9 codes: 625.9, 626, 626.2, 626.6, and 626.9), was dysfunctional uterine bleeding (DUB). Searching medical records of members of a single health maintenance organization (HMO) in Atlanta, Georgia, the investigators identified women who had at least one qualifying diagnosis during a visit to a practice physician in either 1995 or 1996. They randomly selected an age-stratified sample of 500 females who were HMO members in 1995 or 1996 and did not have a qualifying visit during this time period, regardless of whether they had visited a practice physician during this two year period or not. Women were invited to participate by mail. Those women who returned a post-card indicating their interest in participating in the study and signed an informed consent were interviewed in person and provided a venous blood sample during a single visit to the Centers for Disease Control and Prevention Hemostasis laboratory. The study was separately approved by the internal review boards of Emory University, the Centers for Disease Control and Prevention, and the HMO.

FVIII:C was measured using an automated analyzer using mechanical end-point determination (STA, Diagnostica Stago) by a one-stage assay (Diagnostica Stago, Parsippany, NJ). Enzyme-linked immunosorbent assay (ELISA) with a polyclonal antiserum (Asserachrom von Willebrand Factor, Diagnostica Stago) was used to obtain von Willebrand Factor antigen (VWF:Ag) measurements. FIX:C activity was measured on an automated analyzer (STA, Diagnostica Stago) using deficient plasma (Diagnostica Stago). The initial focus of the original study was to investigate the association between VWF deficiency and DUB(151). Dilley et al. (150) provide full details regarding the design, execution, and results of this case-control study.

4-II-B. CES-D instrument

The Center for Epidemiologic Studies – Depresssion Scale (CES-D)(116) is a standard, well-validated instrument used to screen for depression in the general population. Twenty questions concerning depressive symptoms appeared as a group in the questionnaire. Each question has four possible answers: (Strongly Agree, Agree, Disagree, Strongly Disagree). Three of the questions are phrased positively to interrupt a possible flow of set responses. Each answer receives a score of 0-3, with the positively phrased questions being scored in reverse. The final CES-D score, with penalties for missing answers, ranges from 0-60, which we dichotomized using a cut point of 16(166). Wong(117) and McDowell(118), among others, provide salient discussions of the CES-D scale, including its use, validity, and reliability.

4-II-C. Statistical Analysis.

We limited analyses to women who reported being either black or white and had no obvious coagulation deficiency that might be expected to affect FVIII:C level. Eight women who had VWD deficiency(151) were excluded, but no other obvious coagulopathies relevant to FVIII were found. We transformed the FVIII:C levels to improve normality by taking the natural logarithm of the values and used the transformed value as the outcome variable in all analyses. Using a two-sample t-test, we compared FVIII:C levels by race. To take advantage of a large population of black Americans, we were interested in exploring race-specific associations between FVIII:C levels and depression and thus performed separate analyses for each race. To assess the association between FVIII:C levels and depression after adjusting for potentially important covariates, we employed multiple linear regression. Since we were unsure of the functional relationship between the score and FVIII:C levels, we examined two alternative representations of the score: the rank of the ordered scores allowing for ties and a dichotomization with a cut-off of 16 or greater indicating "depression".

Based on current knowledge of factors thought to influence FVIII:C levels and availability of data, we were able to adjust for potential confounding by the following covariates selected *a priori*: VWF:Ag (IU/dL), FIX:C level (IU/dL), age (years), BMI (kg/m^2), oral contraceptive use (yes/no), reported history of diabetes (yes/no), reported history of high cholesterol (yes/no), the current number of cigarettes smoked per day, exercise units in the last month (with one unit defined as twenty minutes of exercise in the last month), the maximum number of drinks consumed in a single 24 hour period, and the ABO blood phenotype. The ABO phenotype was represented with two dummy

variables to indicate the four phenotypes A, B, AB, and O (the referent group). We chose the *a priori* model shown in **Table 4-3** to investigate the adjusted association between FVIII:C and the CES-D score (or its alternative representations). For ease of interpretation, we present the results of the regression analyses as the estimated percent difference in the geometric mean FVIII level among the exposed compared to the unexposed. This percent difference is estimated by 100%*(exp(β)-1), when a log-transformation has been used(167). For a continuous exposure, such as the depression score, the expression 100%*(exp(β)-1) is interpretable as the difference, expressed as a percent, in geometric mean FVIII:C level among those with a given depression score compared with the corresponding mean among those with a score that is one unit lower(167).

We also performed a series of sensitivity analyses. The first consisted of differentially weighting women with and without DUB to adjust for the unequal probabilities of selection. Using the estimated prevalence of menorrhagia of 13.3%(153), we calculated the weights based on the inverse probability of selection(152). We also performed separate analyses among women with DUB and without DUB. Finally, we repeated these analyses using the combined black and white populations. In the second set of sensitivity analyses, we examined three additional variations of the *a priori* model: exclusion from the model of FIX:C levels, of VWF:Ag levels, and of FIX:C and VWF:Ag levels jointly. The appendix provides a discussion of the rationale, results and interpretations for these sensitivity analyses.

4-III. Results

Data for a total of 217 white and black women were available for this for analysis. One hundred thirty-nine (64%) of these women reported their race as black. The range of the FVIII:C levels was 3.93 to 5.95, with a mean (SD) of 4.74 (0.34). Seventy-three (34%) subjects had a CES-D score greater than or equal to 16, and were categorized as depressed. The mean (SD) CES-D score was 13.7 (11.1), with a range from 0 to 60.

The geometric mean FVIII:C level, shown in **Table 4-1** by race and depression, is higher in depressed black women than in non-depressed black woman ($p = 0.11$). However, depressed white women did not have a higher mean than non-depressed white women ($p = 0.67$).

Table 4-1. The mean natural log transformed FVIII:C level by race and depression status among the women of the CDC study of menorrhagia.

Race	Depressed[a]	N	Ln(FVIII:C)				
			Mean	Std Dev	Median	Minimum	Maximum
Black	No	88	4.74	0.34	4.75	3.93	5.75
	Yes	51	4.84	0.38	4.88	4.06	5.95
White	No	56	4.68	0.32	4.65	4.03	5.36
	Yes	22	4.65	0.28	4.60	4.08	5.17

(a) No = CES-D score < 16, Yes = CES-D score ≥ 16;

The distributions of several factors that may moderately to strongly influence FVIII:C levels nominally differ among the race-depression strata (**Table 4-2)**. Importantly, average VWF:Ag and FIX:C levels were higher in depressed black women

Table 4-2. The characteristics of the women in the CDC case-control study of menorrhagia.

Factor	Black, Not Depressed*	Black, Depressed*	White, Not Depressed*	White, Depressed*
N	88	51	56	22
CES-D Score[1]	7.4(4.8)	26.8(8.2)	7.1(4.5)	25.1(11.0)
VWF:Ag (IU/dL) [1]	129.5(47.9)	135.5(46.5)	115.8(39.8)	106.9(38.4)
FIX:C (IU/dL) [1]	141.5(27.7)	154.1(33.7)	139.9(32.7)	151.9(28.0)
Age (years) [1]	35.5(7.4)	35.0(7.6)	36.2(7.9)	34.2(8.7)
BMI (kg/m^2) [1]	28.4(6.1)	30.4(8.3)	25.0(6.6)	27.1(6.7)
Cigarettes/day[1]	1.3(5.0)	1.0(4.0)	3.1(7.9)	3.6(6.8)
MaxDrinks[a, 1]	1.6(1.8)	1.0(1.3)	2.5(2.3)	3.1(3.7)
Exercise units[b, 1]	17.6(21.8)	12.5(20.3)	26.7(36.0)	16.9(19.0)
Oral contraceptive use[2]	19(21.6%)	9(17.6%)	13(23.2%)	6(27.3%)
History of diabetes[2]	2(2.3%)	3(5.9%)	2(3.6%)	3(13.6%)
History of high cholesterol[2]	17(19.3%)	9(17.6%)	8(14.3%)	2(9.1%)
DUB Case[c,2]	42(47.7%)	22(43.1%)	25(44.6%)	9(40.9%)
ABO phenotype				
A[2]	15(17.0%)	12(23.5%)	23(41.1%)	10(45.5%)
AB[2]	3(3.4%)	5(9.8%)	2(3.6%)	1(4.5%)
B[2]	24(27.3%)	8(15.7%)	11(19.6%)	3(13.6%)
O[2]	46(52.3%)	26(51.0%)	20(35.7%)	8(36.4%)

(*****) Center for Epidemiologic Studies Depression (CES-D) score ≥ 16; (a) the maximum number of alcoholic drinks the respondents reported consuming during one episode in the last 12 months; (b) the number of 20 minute units of exercise in the last month. (c) Dysfunctional Uterine Bleeding (DUB) case definition met: medical record contained at least one of the following ICD-9 codes during the period of the study: 625.9, 626, 626.2, 626.6, and 626.9. (1) The summary measure is mean (standard deviation). (2) The summary measure is frequency (percent).

than in non-depressed black women. Similarly FIX:C was higher, on average, in depressed white women than in non-depressed white women. In contrast, VWF:Ag was lower, on average, in depressed white women compared to non-depressed white women. In the appendix, we address the effects of incorporation or exclusion of both FIX:C and VWF:Ag levels in the model.

After adjusting for covariates, the mean FVIII:C level in depressed black women was slightly higher than in non-depressed black women, but the confidence interval included the possibility of no difference-the geometric mean level was 7.1% higher in depressed than non-depressed black women (95% CI: -1.6%, 16.7%), as shown in **Table 4-3**. Among white women, depression was weakly associated, if at all, with FVIII:C level, -0.7% (95% CI: -10.7%, 10.4%). When the CES-D score was treated as continuous, the FVIII:C levels tended to increase with CES-D score in black women: 0.3% increase per unit increase in CES-D score (95% CI -0.1%, 0.6%), but not in white women: 0.1% decrease (-0.5%, 0.4%), respectively. The results were similar when we used the rank of CES-D score, rather than the CES-D score itself (data not shown). In general, omitting VWF:Ag level from the model strengthened the association between FVIII:C level and depression for both black and white women. However, omitting FIX:C level only strengthened the association for black women. These results are presented in more detail in the appendix.

In sensitivity analyses after accounting for over-sampling of DUB cases by weighting their contribution or excluding them completely, the association between FVIII:C and depression was somewhat stronger and more positive in black women,

Table 4-3. Percent Difference in Geometric Mean FVIII:C levels, from multivariate analyses in the black- and white-only populations.

Parameter	Dichotomized CES-D score		CES-D score	
	Blacks (n = 137)	**Whites (n = 78)**	**Blacks (n = 137)**	**Whites (n = 78)**
	%	%	%	%
Depression[a]	7.13 (-1.57, 16.59)	-0.72 (-10.72, 10.40)		
CES-D score[b]			0.28 (-0.09, 0.64)	-0.06 (-0.52, 0.40)
VWF:Ag (IU/dL)	0.48 (0.38, 0.58)	0.47 (0.33, 0.61)	0.48 (0.38, 0.58)	0.47 (0.33, 0.61)
FIX:C level (IU/dL)	0.17 (0.02, 0.32)	0.13 (-0.07, 0.33)	0.16 (0.01, 0.31)	0.13 (-0.07, 0.34)
Cigarettes per day	0.18 (-0.68, 1.05)	0.06 (-0.58, 0.70)	0.13 (-0.73, 1.00)	0.07 (-0.57, 0.71)
Maxdrink[c]	1.82 (-0.71, 4.42)	-1.67 (-3.44, 0.13)	1.65 (-0.86, 4.21)	-1.69 (-3.45, 0.10)
Age (years)	0.50 (-0.07, 1.06)	-0.52 (-1.14, 0.11)	0.53 (-0.04, 1.10)	-0.53 (-1.15, 0.10)
BMI (kg/m^2)	-0.45 (-1.08, 0.18)	-0.19 (-1.15, 0.78)	-0.42 (-1.05, 0.22)	-0.19 (-1.15, 0.78)
Oral contraceptive use[d]	-5.54 (-14.68, 4.58)	-11.57 (-21.83, 0.04)	-5.26 (-14.47, 4.94)	-11.75 (-22.08, -0.05)
Exercise units[e]	-0.02 (-0.21, 0.17)	0.12 (-0.03, 0.27)	-0.03 (-0.21, 0.16)	0.12 (-0.03, 0.27)
History of diabetes[d]	2.68 (-17.92, 28.45)	20.41 (-1.62, 47.37)	2.32 (-18.24, 28.03)	20.99 (-1.53, 48.65)
History of high cholesterol[d]	2.95 (-7.23, 14.24)	-3.82 (-16.66, 10.99)	2.69 (-7.48, 13.98)	-3.84 (-16.65, 10.93)
A phenotype	11.80 (0.54, 24.32)	2.09 (-8.46, 13.86)	12.33 (1.02, 24.91)	2.01 (-8.52, 13.76)
B phenotype	6.74 (-5.00, 19.93)	2.05 (-11.16, 17.22)	7.13 (-4.76, 20.49)	2.12 (-11.10, 17.31)
AB phenotype	1.64 (-15.31, 21.97)	6.64 (-16.57, 36.29)	2.73 (-14.35, 23.21)	6.87 (-16.43, 36.68)

(a) CES-D score ≥ 16 = 1 versus CES-D score < 16 = 0; (b) Center for Epidemiologic Studies Depression score; (c) the maximum number of alcoholic drinks the respondents reported consuming during one episode in the last 12 months; (d) the number of 20 minute units of exercise in the last month; (e) Yes = 1, No = 0.

but more negative in white women (**Table 4-A-1** of the appendix). However, the numbers of subjects was small in these analyses, particularly for whites. We provide more details of these sensitivity analyses, and also discuss the rationale and impact of controlling or not controlling for possible confounding by VWF:Ag and FIX:C levels, in the appendix.

4-VI. Discussion

A consistent, positive association between FVIII:C levels and depression was not observed in this study. The association between FVIII:C levels and depression differed between white and black women. Among black women, FVIII:C levels were slightly higher among women with higher depression scores than among non-depressed women. However, this association was not statistically significant. Further, among white women, depression was not associated with increased FVIII:C levels.

Since women with DUB were over-represented in this study, we attempted to account for potential selection bias by studying controls only and, separately, by weighting subjects in inverse proportion to their estimated selection probabilities(152). The observed patterns did not substantially change with weighting for DUB-status or analyses stratified by race. Depressed black women consistently had higher FVIII:C levels than non-depressed black women; while this association was not seen in white women. However, since the number of white women was small, the power to identify an increase in FVIII:C levels among women with depression was limited since only 13 depressed white without DUB were studied.

Overall, the sample sizes for this study were relatively small. A further weakness was that we were unable to address whether depression had chronic or acute effects, which is especially pertinent as FVIII is an acute phase reactant. Therefore, measuring depression, along with exertion and inflammation, in the period immediately prior to phlebotomy would have been ideal. Important strengths of the study included measurement of VWF and FIX levels, inclusion of two racial groups from the same

group practice, and availability of important covariate data, such as use of oral contraceptives and history of diabetes.

The association of FVIII:C levels with depression in either black or white women was much stronger when we did not control for VWF:Ag level. This association was also much stronger when we did not control for FIX:C level among black, but not white women; however, the major impact was whether or not we adjusted for VWF:Ag level.

Several factors are important in the decision to adjust for VWF and FIX level. One is whether coordinate regulation of the levels of these proteins exist and another is the nature of the association and regulation. For example, if depression increases the expression or the secretion (release) of FVIII which, in turn, leads to increased FIX synthesis, controlling for FIX:C level might obscure true differences. The most appropriate analysis, may require a greater understanding of these relationships than is currently available in the literature or may require background information not available in the data at hand(168).

This study supports associations of FVIII:C levels with number of previously reported factors. As has been report previously, diabetes is associated with increased FVIII:C levels(80), an association supported in the current study, especially among white women. Similarly, the O blood group phenotype was associated with lower and VWF with higher FVIII:C levels in this study, consistent with previous findings(80, 83, 96, 124). Factor VIII:C levels have typically been positively associated with age (80, 88, 100), and this study supports this finding in black, although not white, women but the age range was narrow. On the other hand, BMI has also typically been reported to be associated with higher FVIII:C levels, (80, 83, 98). In contrast, we found a weakly,

negative association. That our findings support a number of previously-reported relationships suggests that major biases may be absent and that the study should have had validity to assess the main association of interest.

The main exposure in this investigation, depression, is a rather nebulous term, as is the clinical definition of a major depressive episode (MDE), for which a physiological response is ill defined(111). At best, the CES-D instrument provides an inexact approximation(116, 117), and deducing how certain components or the composite score may be associated with FVIII:C will require additional investigations. For instance, two questions that appear on the CES-D instrument are "I could not get 'going'" and "My sleep was restless". Further, two of the nine possible qualifying symptoms for MDE are (3) significant weight loss and (8) diminished ability to think or concentrate, any of which may plausibly be associated with FVIII:C level. Therefore, validation of the definition of depression in association with FVIII:C levels would be important for future examinations of this question.

Complicating interpretation of these analyses, are the possibilities that DUB, such as menorrhagia, may be a risk factor for depression(169) given their impacts on health, quality of life, and amplification of perceived discomfort(170). The converse may also be true. Therefore, even if a consistent pattern was observed, it would be difficult to interpret these results causally. Further, an increased mean FVIII among depressed black women but a decreased mean FVIII among depressed white women is difficult to understand biologically. Therefore replication of these findings must be made before any definitive conclusions about the relationship between depression and FVIII:C levels can be made.

Since physical and mental activity increase FVIII:C levels, a mental illness, such as depression, might also plausibly affect FVIII:C levels. Since high FVIII:C levels are associated with thrombosis the former association would be worrisome. The impact and prevalence of depression and CVDA, and the association between them, warrant additional studies, especially to establish whether FVIII:C level is a causal factor for depression in addition to its etiologic role in CVDA. The results of this study, a cross-sectional design that effectively precludes inferences on direction of causality, are inconclusive and varied across racial groups. Nevertheless, they support a positive association between depression and FVIII levels among black women. Before beginning the much more resource intensive investigation into temporality and causal mechanisms, it is important to establish the reported association and to have greater confidence that the current findings did not result from uncontrolled confounding or chance. Therefore, we encourage additional studies of this important issue.

4-V. Appendix

4-V-A. Possible Selection bias

Women with and without DUB were selected in approximately equal numbers, implying that the selection fraction for women with DUB was higher than that of women without DUB assuming the prevalence of menorrhagia is close to published estimates, 13.3%(153). Therefore, we weighted subjects in proportion to the inverse of the selection probability(152). We chose to calculate the weights for three prevalences: 5%, 10%, and

15%. The weight for women with DUB was calculated as W = (Prevalence * $N_{non\text{-}DUB}$)/(N_{DUB}*(1-Prevalence)) and that for women without DUB was 1. As shown in **Table 4-A-1**, the association between FVIII:C levels and depression appears to be different among women with and without DUB. In black women, the associations of FVIII:C with depression appear to strengthen with decreasing weights for the women with DUB. The association is positive and strongest among black women without DUB, although the confidence interval included the possibility of no association. The results for white women, with a smaller sample size, differ and suggest little association, or possibly even a negative association.

Table 4-A-1. The effects of adjustments for the potential bias resulting from unequal selection probabilities on the association between the FVIII:C level and depression.

Selection criteria	Weight for DUB (Prevalence)[+]	Black-only[‡]	White-only[‡]
Depression[*]			
With and without DUB	1.00 (0.50)	7.13% (-1.57%, 16.59%)	-0.72% (-10.72%, 10.40%)
With and without DUB	0.18 (0.15)	9.88% (0.89%, 19.67%)	-4.16% (-13.50%, 6.20%)
With and without DUB	0.11 (0.10)	10.51% (1.44%, 20.40%)	-4.66% (-13.91%, 5.58%)
With and without DUB	0.05 (0.05)	11.26% (2.08%, 21.26%)	-5.21% (-14.35%, 4.92%)
Without DUB		12.16% (-0.31%, 26.19%)	-5.80% (-17.64%, 7.74%)
With DUB		4.86% (-8.68%, 20.42%)	9.81% (-7.67%, 30.60%)
CES-D Score			
With and without DUB	1.00 (0.50)	0.28% (-0.09%, 0.64%)	-0.06% (-0.52%, 0.40%)
With and without DUB	0.18 (0.15)	0.35% (-0.01%, 0.70%)	-0.02% (-0.43%, 0.39%)
With and without DUB	0.11 (0.10)	0.37% (0.02%, 0.72%)	-0.02% (-0.42%, 0.39%)
With and without DUB	0.05 (0.05)	0.40% (0.04%, 0.75%)	-0.01% (-0.41%, 0.39%)
Without DUB		0.43% (-0.06%, 0.92%)	-0.01% (-0.53%, 0.52%)
With DUB		0.39% (-0.25%, 1.04%)	-0.21% (-1.12%, 0.70%)

(*****) Center for Epidemiologic Studies Depression (CES-D) score ≥ 16; (+) Weight = (Prevalence * $N_{non\text{-}DUB}$)/(N_{DUB}*(1-Prevalence)); (‡) Adjusted for VWF:Ag (IU/dL), FIX:C level (IU/dL), Race(0/1, Combined only), Age (years), BMI (kg/m2), Oral contraceptive use (0/1), History of diabetes (0/1).

4-V-B. Impact of Adjustment for Factor IX and for VWF:Ag

The sole known function of FVIII, which has a relative short half-life that is greatly reduced when it is not in complex with VWF, is to be the co-factor for FIX. Therefore, imbalances of FVIII with either VWF or FIX are a potential waste of resources. We speculate that there may be coordinate regulation for the levels of FVIII and VWF and FVIII and FIX. For instance, if FVIII is secreted but VWF is not present in sufficient quantities, FVIII is quickly cleaved and cleared. Potentially, the body might detect such an occurrence and up-regulate either VWF production or secretion. If depression, or some other factor, increases FVIII synthesis, by some mechanism not addressed, then VWF levels would also increase under this scenario. If so, then controlling for VWF:Ag levels may obscure detection of the association between depression on FVIII. Similarly, FVIII serves no known function unless it can be

Table 4-A-2. The effects of excluding FIX:C and VWF:Ag levels on the association between depression and FVIII:C levels.

	Full model*	**No FIX:C**	**No VWF:AG**	**No FIX:C or VWF:Ag**
	Percent difference (%)	Percent difference (%)	Percent difference (%)	Percent difference (%)
Black-only				
Depression	7.13 (-1.57, 16.59)	8.50 (-0.38, 18.16)	10.24 (-0.99, 22.75)	13.11 (1.33, 26.26)
CES-D score	0.28 (-0.09, 0.64)	0.35 (-0.02, 0.71)	0.40 (-0.06, 0.87)	0.54 (0.07, 1.01)
White-only				
Depression	-0.72 (-10.72, 10.40)	0.21 (-9.89, 11.44)	-7.71 (-19.01, 5.17)	-6.45 (-18.08, 6.84)
CES-D score	-0.06 (-0.52, 0.40)	-0.02 (-0.47, 0.44)	-0.42 (-0.97, 0.14)	-0.35 (-0.92, 0.22)

(*) See Table 4-3 for details concerning the model.

activated and couple with activated FIX. Therefore, it is conceivable that increased FVIII:C levels could result in an increase in FIX levels. **Table 4-A-2** presents the results of excluding FIX:C, VWF:Ag, and both jointly from multiple linear regressions. We have chosen the model that includes ABO blood type since it affects both FVIII and VWF.

Interestingly, dropping VWF:Ag level from the regression model strengthened the association of FVIII:C with depression. Dropping FIX:C level from the model slightly strengthened the association in black women but weakened it in white women. The most appropriate modeling approach may not always be clear, without additional knowledge(168). For instance VWF is partly a confounder, but may also partly be a consequence of FVIII, possibly requiring an alternative methodology or study design.

Chapter 5

Conclusion

5-I. Goals, Brief Overview of Background, and Importance

The goals of this dissertation were to study determinants of Factor VIII activity levels. Specifically, they were to survey and describe the DNA variants in the classically considered functional regions of *F8*; to determine associations of these variants with FVIII:C levels in populations without obvious FVIII-related coagulopathies; and to investigate the association between FVIII and depression, a common mental illness. The first goal is a precursor to investigation of FVIII that is a hallmark of the genomic era. The second goal follows logically from the first, with the newly identified variants constituting an augmentation to the more classical exposures or factors currently studied. The third goal represents a return to more conventional investigations that will eventually be assisted by the incorporation of knowledge from the first two goals.

5-I-A. FVIII is central to proper hemostasis

While all of the proteins in the coagulation system are essential, the bleeding and thrombotic events associated with extremes in FVIII:C level underscore its importance as the co-factor to FIX in the early amplification stage of the cascade reactions. Specifically, a deficiency of FVIII:C results in hemophilia A, a bleeding disorder. At the other extreme, high levels are associated with thromboses that may have severe consequences, particularly if they occur in the heart or brain. Even

peripheral thromboses, such as a deep vein thrombosis (DVT), carry the risk of life threatening pulmonary embolism (PE).

5-I-B. Epidemiologic and Public Health Importance

The individual and public health impacts of abnormal Factor VIII levels are great. Hemophilia A is rare, but requires expensive treatment, potentially costing more than $100,000 per year per patient(171) and places burdens on the individual and the health system.

Even in the absence of hemophilia, FVIII levels vary widely. For example, many people have levels above 150 IU/dL, a level widely cited as potentially deleterious(131). The risks associated with less extreme FVIII:C levels are uncertain, so identifying the environmental, endogenous, and genetic factors that are associated with variations in FVIII:C levels in this range will help elucidate the mechanisms involved in its life-cycle, thus assisting in the prevention or treatment of the exophenotypes to which FVIII contributes.

The reported association of FVIII:C levels with depression(119, 120) is also of concern. The causal nature of the association is still uncertain, but nevertheless important, as depression is a relatively common condition with important personal and public health consequences. Thus, understanding the genetic and environmental factors that affect FVIII:C levels is not just important to biology and physiology, but may underlie successful prevention and treatment of diseases with an immense public health impact.

The three papers of this dissertation are designed to provide new, important information and to add to the scientific literature concerning the determinants of FVIII:C levels. The first study sought to identify new polymorphisms in the *F8* gene and identify associations between these variants and FVIII:C levels, the second to confirm the association results found in the first, and the third to investigate a previously reported, but unconfirmed, association between depression and FVIII:C levels.

5-II. Review of Major Findings

5-II-A. Genetic Variation in *F8*

The first study, in which we sought to identify variants in 222 X chromosomes, was among the largest to date. We found only 47 variants, 18 of which were previously unreported. Importantly, of the eleven SNPs amongst the codons, five result in missense mutations: W0255C, R0484H, R0776G, D1241E, and M2238V. Barring any splice-affecting mutation, the remaining variants will not affect the protein structure, but rather may affect the level of FVIII through mechanisms such as, but not limited to, transcription regulation, RNA stability, and translation efficiency. For instance, we reported five promoter SNPs, including an insertion-deletion mutation, which could potentially affect the rate of transcription.

5-II-B. Association between *F8* variants and FVIII:C level

The measured-genotyped association analyses conducted as part of the first study offer more power than the corresponding linkage analyses. As a first step, we investigated each SNP independently and identified two SNPs that were strongly associated with FVIII:C levels: G056010A, an intronic SNP within 30 bases of the exon junction, and C092714G, which encodes D1241E. Notably, in the GAIT population that we studied, these two SNPs were in high allelic association ($r^2 = 0.84$, unadjusted). The second study, in which we attempt to replicate the results of the first study using separate study populations, extended the first by including another racial group. Within white subjects in this study, the association was consistent with the first study. Within the blacks, however, the association was not strong and included the possibility of no association.

5-II-C. Association between depression and FVIII:C level

In the third study of this dissertation, in which we attempted to replicate a previously reported association between depression and FVIII:C levels, the associations that we found were not consistent among the two racial groups under study. Among the black women, the results suggested the possibility that depression might be associated with higher FVIII:C levels, whereas the results among white women did not. Although we did not clearly replicate the association of depression with FVIII:C levels, we were able to replicate several other reported associations

including the associations with VWF:Ag, FIX:C, oral contraceptives, ABO blood group, and diabetes. The association with diabetes, however, was much weaker among black women and included the possibility of a null association in both racial groups. This study complements the first two by investigating a possible non-genetic determinant of FVIII:C levels.

5-II-D. New perspective on established covariates

Several factors have been established as important predictors of FVIII:C levels. Currently, among populations without coagulopathies, VWF:Ag is the strongest, most consistently replicated predictor. A correlation with FIX:C level has also been reported, but the magnitude is considerably less. This dissertation offers an important perspective on these two factors. Specifically, we may be the first to raise the possibility that these factors may not be just confounders, but also intermediates in a complex, yet unknown and unestablished, system of coordinate regulation. For example, feedback regulatory pathways could form a complex web in which measured VWF:Ag levels were affected by the C092714G genotype. Therefore, the considerations raised here suggest that adjusting for VFW:Ag or FIX:C may obscure the detection of a true association between the exposure of interest and FVIII:C levels.

5-III. Cohesion of results

5-III-A. Pre-requisite variant discovery

The first, discovery portion of this dissertation was essential to identify the *F8* variants that were present in the enrolled families. Identifying and genotyping these variants is a pre-requisite for subsequent analyses that determine whether they might be associated with FVIII:C levels.

5-III-B. Replication is necessary

Replication of the key finding of an association between a variant (the SNPs G056010A and C092714G) and FVIII:C level is important, in part because the association might be artifactual or a chance finding. For example, an allelic association between one of these SNPs and an ungenotyped true functional variant could lead to a spurious association. Replication with a second population of unrelated individuals lessens, but does not preclude, the chance that allelic association is a source of confounding. Thus replication is important, especially in family-based studies like the first study in this dissertation, because unidentified confounders may be more likely shared among family members.

The second study in this dissertation provided this replication of the association between C092714G and FVIII:C levels that we reported in the first paper. Within the Caucasians of the second paper the results were consistent with those from the first

paper, providing encouraging evidence that the D allele is associated with an additive increase in FVIII:C levels. Among the black women of the second paper, though, the results were not as strong.

5-III-C. Adjustment for VWF:Ag and FIX:C level

Although not an initial focus of the studies here, we found that whether or not we adjusted in the analyses for VWF:Ag and FIX:C levels had profound effects on the association between the SNPs studied and FVIII:C. In all three studies of the dissertation, adjustment for VWF:Ag weakened the apparent associations with FVIII:C level, potentially obscuring the effect that the exposure might have had, suggesting that our speculations about coordinate regulation may have some credibility. Omission of FIX:C levels from the analyses resulted in more mild differences between the studies that were not easily summarized.

5-III-D. Implication of replication

The concordant results of the first and second papers of this dissertation lend strength to the main conclusion concerning the association of C092714G genotype with FVIII:C, and very likely FVIII:Ag . Furthermore, the finding in all three studies that adjustment for VWF:Ag had important impacts on the associations with two different exposures, a SNP on the one hand and depression on the other, also

underscores the importance of further understanding the causal pathways. Finally, understanding the role of depression may also be important.

5-IV. Agreement with current knowledge

5-IV-A. The association between C092714G and FVIII:C levels was novel

Our report of an association between C092714G and FVIII:C was novel, even though it is one of the oldest known *F8* missense SNPs. Since our report, a subsequent case-control study using an Italian population (GAIT was among Catalans) supported our finding(141). Nossent et al.(142), however, only supported this finding among the male sub-populations of their investigation and the haplotype analyses suggested that C092714G may not be functional.

5-IV-B. Depression and FVIII:C levels

Our finding of a positive association between depression and FVIII:C levels in the third study provides some weak support of previous reports (119, 120) . Although, results among white women in the third study were inconsistent, the confidence intervals were wide reflecting the small study size. In both the second and third studies, the associations between FVIII:C levels and most of the covariates

considered were consistent with current knowledge, providing support that important systematic biases were absent.

5-V. Important Strengths

The studies in this dissertation had a number of important strengths, that reflect the contributions they make to the scientific literature concerning determinants of FVIII:C levels. Important strengths include the relatively large number of X chromosomes used in the variation scan that was part of the first study. In contrast, the scan by SeattleSNPs used only 47 males and females(149) whereas our scan included 222 X chromosomes from unrelated people. Some later studies may have used larger numbers. Another strength was availability of fairly detailed environmental and endogenous exposures, permitting greater adjustment for covariates than otherwise would have been possible.

The second study built on the first, and attempted to replicate the findings of the first in not only a separate study population, but also in a separate racial group. The second and third studies were also strengthened by the availability of a relatively large number of covariates. Thus, potential confounding by these covariates could be addressed.

Finally, the way in which the three studies complement one another is also a strength. First, the subjects used in first study differ from those in the second two studies. Next, the second study attempted to provide independent replication of the

first in a different racial group and with a different study design. Lastly, the third study addressed a non-genetic determinant of FVIII:C levels whereas the first two studies addressed genetic determinants.

5-VI. Important Weaknesses

Future studies may benefit by consideration of three weaknesses that were shared by the studies in this dissertation. First, all of the studies would be benefited by additional information concerning the outcome. In particular, FVIII:Ag level was not measured and might have been useful. Nevertheless, we found important associations between SNPs and FVIII:C in the first two studies, despite not having measured FVIII:Ag levels.

Second, all of the studies might have been strengthened had more detailed information been available for some of the covariates. For example in the first study, data concerning smoking might have been more informative if it had been recorded as average number of cigarettes per day in the first study as opposed to yes or no.

Finally, the numbers of subjects in each study were modest. The original GAIT project was intended only as a pilot study, and thus the sample size was smaller than desired by design. The investigations in the second and third studies involved previously identified subjects, so that the sample size was based on the original investigation. Nonetheless, our results were informative, and at least supportive, although not definitive, of an association between the C092714G SNP and FVIII:C.

5-VII. Conclusions

5-VII-A. FVIII has multiple forms

The five missense mutations identified in the first study among populations without FVIII-coagulopathy, were found among a rather modest number of chromosomes. Therefore, the diversity of FVIII should no longer be ignored. Specifically, the proteins should be characterized in any investigation and caution should be used to avoid misclassification by grouping the FVIII proteins under study in too few categories.

5-VII-B. C092714G is associated with FVIII:C levels

The association between C092714G of *F8* with FVIII:C levels reported in the first study appears not to be spurious, in part because we were able to partially confirm it in the second study. The confirmation, however, was not definitive and the association should be evaluated in several more studies, particularly since 1241 lies within the B-domain, which floats free of the protein upon activation.

5-VII-C. The association between FVIII:C and depression was weakly supported

Depression may be associated with FVIII:C levels, at least among black women, consistent with previously reported findings. Although results were inconsistent among white women, this was the smallest group and confidence intervals included the possibility of a positive association. These results and their potential importance warrant additional evaluation of this association.

5-VIII. Future directions

5-VIII-A. Measurement of FVIII:Ag levels

Determining whether FVIII:Ag levels are associated with C092714G is one of the most interesting advances possible for the comparably modest cost. Ideally, FVIII:C levels would also be measured in the same study. Because we were able to only partially confirm this association, additional studies are warranted.

5-VIII-B. Characterizing the proteins with which FVIII interacts

At the moment, all variations of a given protein that interacts with FVIII, such as VWF and FIX, are classified as one group. Characterizing these proteins will allow

us to study whether distinct forms of these other proteins may be associated with variations in FVIII:C or FVIII:Ag level through currently unexplored mechanisms like differential binding or dissociation.

5-VIII-C. Explore associations with haplotypes

Since variants occur in specific combinations in genes, investigating these haplotypes (combinations of syntenic variants) is a necessary next step in observational studies. With encouraging results from investigations, including Bayesian quantitative trait nucleotide analyses(172), we may next explore *in vitro* functional investigations of a single change within a haplotype or of haplotypes that may not occur in nature but may nonetheless be informative or a useful therapeutic avenue.

5-VIII-D. The role of Depression

The association of depression with FVIII:C levels is intriguing but unclear. A better definition of depression itself might help to reduce misclassification and strengthen confidence in results. Identification of physiologic responses that discriminate between non-depressed and depressed patients or between levels of severity, and also that might affect FVIII:C levels, would be a useful advancement in this field.

If the reported association between depression and FVIII:C levels, by others and by us among blacks, is not spurious, it will be important to prospectively study cohorts of subjects to determine the temporality of the association which can be ambiguous in a cross-sectional study like ours. For instance, a cohort of non-depressed subjects classified by FVIII:C levels might be followed with modest effort to determine whether high FVIII:C levels are a risk factor for depression.

5-VIII-E. Refine the definition of diabetes

In all three studies of this dissertation, diabetes was recorded as a dichotomous variable. Although, the association with FVIII:C levels was not consistent or strong, previous reports have suggested an association(80, 98) and the mechanism (insult to the endothelium) has strong biological plausibility. It seems worthwhile to measure other factors related to diabetes such as duration, blood glucose level, and insulin level since it is another moderately prevalent disease with immense public health impact.

References

1. Furie B, Furie BC. Molecular and cellular biology of blood coagulation. N Engl J Med 1992;326:800-6.
2. Krishnaswamy S. Exosite-driven substrate specificity and function in coagulation. Journal of Thrombosis and Haemostasis 2004;3:54-67.
3. Mann KG. Biochemistry and physiology of blood coagulation. Thromb Haemost 1999;82:165-74.
4. Colman RW HJ, Marder VJ, Clowes AW, and George JN. Hemostasis and Thrombosis: Basic Principles and Clinical Pratice. Fourth Edition. 2001.
5. Saenko EL, Ananyeva NM, Tuddenham EG, Kemball-Cook G. Factor VIII - novel insights into form and function. Br J Haematol 2002;119:323-31.
6. Fay PJ. Activation of factor VIII and mechanisms of cofactor action. Blood Rev 2004;18:1-15.
7. Panteleev MA, Saenko EL, Ananyeva NM, Ataullakhanov FI. Kinetics of Factor X activation by the membrane-bound complex of Factor IXa and Factor VIIIa. Biochem J 2004;381:779-94.
8. Lollar P. Structure and function of Factor VIII. Adv Exp Med Biol 1995;386:3-17.
9. High KA, Roberts HR. Molecular basis of thrombosis and hemostasis. New York: M. Dekker, 1995.
10. Bertina RM, Koeleman BP, Koster T, et al. Mutation in blood coagulation factor V associated with resistance to activated protein C. Nature 1994;369:64-7.
11. Ananyeva NM, Kouiavskaia DV, Shima M, Saenko EL. Catabolism of the coagulation factor VIII: can we prolong lifetime of f VIII in circulation? Trends Cardiovasc Med 2001;11:251-7.
12. McGlynn LK, Mueller CR, Begbie M, Notley CR, Lillicrap D. Role of the liver-enriched transcription factor hepatocyte nuclear factor 1 in transcriptional regulation of the factor V111 gene. Mol Cell Biol 1996;16:1936-45.
13. Weiss HJ, Sussman, II, Hoyer LW. Stabilization of factor VIII in plasma by the von Willebrand factor. Studies on posttransfusion and dissociated factor VIII and in patients with von Willebrand's disease. J Clin Invest 1977;60:390-404.
14. Tuddenham EG, Lane RS, Rotblat F, et al. Response to infusions of polyelectrolyte fractionated human factor VIII concentrate in human haemophilia A and von Willebrand's disease. Br J Haematol 1982;52:259-67.
15. Lenting PJ, van Mourik JA, Mertens K. The life cycle of coagulation factor VIII in view of its structure and function. Blood 1998;92:3983-96.
16. Schwarz HP, Lenting PJ, Binder B, et al. Involvement of low-density lipoprotein receptor-related protein (LRP) in the clearance of factor VIII in von Willebrand factor-deficient mice. Blood 2000;95:1703-8.
17. Saenko EL, Yakhyaev AV, Mikhailenko I, Strickland DK, Sarafanov AG. Role of the low density lipoprotein-related protein receptor in mediation of factor VIII catabolism. J Biol Chem 1999;274:37685-92.

18. Kaufman RJ, Pipe SW, Tagliavacca L, Swaroop M, Moussalli M. Biosynthesis, assembly and secretion of coagulation factor VIII. Blood Coagul Fibrinolysis 1997;8 Suppl 2:S3-14.
19. Bontempo FA, Lewis JH, Gorenc TJ, et al. Liver transplantation in hemophilia A. Blood 1987;69:1721-4.
20. Gordon FH, Mistry PK, Sabin CA, Lee CA. Outcome of orthotopic liver transplantation in patients with haemophilia. Gut 1998;42:744-9.
21. Hisatake GM, Chen TW, Renz JF, et al. Acquired hemophilia A after liver transplantation: a case report. Liver Transpl 2003;9:523-6.
22. Gitschier J, Wood WI, Goralka TM, et al. Characterization of the human factor VIII gene. Nature 1984;312:326-30.
23. Pittman DD, Tomkinson KN, Kaufman RJ. Post-translational requirements for functional factor V and factor VIII secretion in mammalian cells. J Biol Chem 1994;269:17329-37.
24. Nichols WC, Ginsburg D. From the ER to the golgi: insights from the study of combined factors V and VIII deficiency. Am J Hum Genet 1999;64:1493-8.
25. Nichols WC, Seligsohn U, Zivelin A, et al. Mutations in the ER-Golgi intermediate compartment protein ERGIC-53 cause combined deficiency of coagulation factors V and VIII. Cell 1998;93:61-70.
26. Thompson AR. Structure and function of the factor VIII gene and protein. Semin Thromb Hemost 2003;29:11-22.
27. Kaufman RJ. Post-translational modifications required for coagulation factor secretion and function. Thromb Haemost 1998;79:1068-79.
28. Bjorkman S, Berntorp E. Pharmacokinetics of coagulation factors: clinical relevance for patients with haemophilia. Clin Pharmacokinet 2001;40:815-32.
29. Sarafanov AG, Ananyeva NM, Shima M, Saenko EL. Cell surface heparan sulfate proteoglycans participate in factor VIII catabolism mediated by low density lipoprotein receptor-related protein. J Biol Chem 2001;276:11970-9.
30. Moestrup SK, Gliemann J, Pallesen G. Distribution of the alpha 2-macroglobulin receptor/low density lipoprotein receptor-related protein in human tissues. Cell Tissue Res 1992;269:375-82.
31. Bovenschen N, Herz J, Grimbergen JM, et al. Elevated plasma factor VIII in a mouse model of low-density lipoprotein receptor-related protein deficiency. Blood 2003;101:3933-9.
32. Bovenschen N, van Dijk KW, Havekes LM, Mertens K, van Vlijmen BJ. Clearance of coagulation factor VIII in very low-density lipoprotein receptor knockout mice. Br J Haematol 2004;126:722-5.
33. Mertens K, Bovenschen N, Havekes LM, van Vlijmen BJ. Role of Low Density Lipoprotein Receptor in the Clearance of Coagulation Factor VIII In Vivo. Blood 2004;104.
34. Jespersen J, Bertina RM, Haverkate F. Laboratory techniques in thrombosis : a manual. Dordrecht ; Boston: Kluwer Academic, 1999.
35. Thompson SG, Martin JC, Meade TW. Sources of variability in coagulation factor assays. Thromb Haemost 1987;58:1073-7.

36. Folsom AR, Wu KK, Conlan MG, et al. Distributions of hemostatic variables in blacks and whites: population reference values from the Atherosclerosis Risk in Communities (ARIC) Study. Ethn Dis 1992;2:35-46.
37. Chambless LE, McMahon R, Finch A, et al. ARIC hemostasis study--III. Quality control. Atherosclerosis Risk in Communities. Thromb Haemost 1993;70:588-94.
38. Toole JJ, Knopf JL, Wozney JM, et al. Molecular cloning of a cDNA encoding human antihaemophilic factor. Nature 1984;312:342-7.
39. Karolchik D, Baertsch R, Diekhans M, et al. The UCSC Genome Browser Database. Nucl. Acids Res. 2003;31:51-54.
40. Mansvelt EP, Laffan M, McVey JH, Tuddenham EG. Analysis of the F8 gene in individuals with high plasma factor VIII: C levels and associated venous thrombosis. Thromb Haemost 1998;80:561-5.
41. Figueiredo MS, Brownlee GG. cis-acting elements and transcription factors involved in the promoter activity of the human factor VIII gene. J Biol Chem 1995;270:11828-38.
42. Levinson B, Kenwrick S, Lakich D, Hammonds G, Jr., Gitschier J. A transcribed gene in an intron of the human factor VIII gene. Genomics 1990;7:1-11.
43. Levinson B, Kenwrick S, Gamel P, Fisher K, Gitschier J. Evidence for a third transcript from the human factor VIII gene. Genomics 1992;14:585-9.
44. Lakich D, Kazazian HH, Jr., Antonarakis SE, Gitschier J. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. Nat Genet 1993;5:236-41.
45. SeattleSNPs. NHLBI Program for Genomic Applications, UW-FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu) [November, 2004].
46. Higuchi M, Wong C, Kochhan L, et al. Characterization of mutations in the factor VIII gene by direct sequencing of amplified genomic DNA. Genomics 1990;6:65-71.
47. Morange PE, Tregouet DA, Frere C, et al. Biological and genetic factors influencing plasma factor VIII levels in a healthy family population: results from the Stanislas cohort. Br J Haematol 2005;128:91-99.
48. Brummer J, Groth J, Flayeh R, Wagener C, Jung R. Absence of mutations at the APC interacting sites of factor VIII in Caucasians. Thromb Haemost 2002;87:170.
49. Hooper WC, Dilley A, Austin H, et al. Absence of mutations at APC cleavage sites Arg306 in factor V and Arg336, Arg562 in factor VIII in African-Americans. Thromb Haemost 1998;79:236.
50. Roelse JC, Koopman MM, Buller HR, et al. Absence of mutations at the activated protein C cleavage sites of factor VIII in 125 patients with venous thrombosis. Br J Haematol 1996;92:740-3.
51. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 1986;51 Pt 1:263-73.
52. Watson JD. Recombinant DNA. New York: Scientific American Books: Distributed by W.H. Freeman, 1992.
53. Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998;8:186-94.

54. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res 1998;8:175-85.
55. Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. Am J Hum Genet 1998;62:1198-211.
56. Jacquard A. The genetic structure of populations. Berlin, New York: Springer-Verlag, 1974.
57. Dyke B. PEDSYS: A Pedigree Data Management System. PGL Tech Rep 2. Population Laboratory, Department of Genetics, Southwest Foundation for Biomedical Research, San Antonio, Tx. 1995.
58. Falconer DS. Introduction to quantitative genetics. New York: Ronald Press, 1960.
59. Biostatistical Genetics and Genetic Epidemiology: John Wiley & Sons, LTD, 2002.
60. Amos CI. Robust variance-components approach for assessing genetic linkage in pedigrees. Am J Hum Genet 1994;54:535-43.
61. Almasy L, Dyer TD, Blangero J. Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. Genet Epidemiol 1997;14:953-8.
62. Hopper JL, Mathews JD. Extensions to multivariate normal models for pedigree analysis. Ann Hum Genet 1982;46 (4):373-83.
63. Souto JC, Almasy L, Muniz-Diaz E, et al. Functional effects of the ABO locus polymorphism on plasma levels of von Willebrand factor, factor VIII, and activated partial thromboplastin time. Arterioscler Thromb Vasc Biol 2000;20:2024-8.
64. Balleisen L, Bailey J, Epping PH, Schulte H, van de Loo J. Epidemiological study on factor VII, factor VIII and fibrinogen in an industrial population: I. Baseline data on the relation to age, gender, body-weight, smoking, alcohol, pill-using, and menopause. Thromb Haemost 1985;54:475-9.
65. Haus E, Cusulos M, Sackett-Lundeen L, Swoyer J. Circadian variations in blood coagulation parameters, alpha-antitrypsin antigen and platelet aggregation and retention in clinically healthy subjects. Chronobiol Int 1990;7:203-16.
66. Iversen PO, Groot PD, Hjeltnes N, Andersen TO, Mowinckel MC, Sandset PM. Impaired circadian variations of haemostatic and fibrinolytic parameters in tetraplegia. Br J Haematol 2002;119:1011-6.
67. van Diest R, Hamulyak K, Kop WJ, van Zandvoort C, Appels A. Diurnal variations in coagulation and fibrinolysis in vital exhaustion. Psychosom Med 2002;64:787-92.
68. Salomaa V, Rasi V, Pekkanen J, et al. The effects of saturated fat and n-6 polyunsaturated fat on postprandial lipemia and hemostatic activity. Atherosclerosis 1993;103:1-11.
69. Bourey RE, Santoro SA. Interactions of Exercise, Coagulation, Platelets, and Fibrinolysis - a Brief Review. Medicine and Science in Sports and Exercise 1988;20:439-446.
70. Kopitsky RG, Switzer ME, Williams RS, McKee PA. The basis for the increase in factor VIII procoagulant activity during exercise. Thromb Haemost 1983;49:53-7.

71. van den Burg PJM, Hospers JEH, vanVliet M, Mosterd WL, Huisveld IA. Unbalanced haemostatic changes following strenuous physical exercise - A study in young sedentary males. European Heart Journal 1995;16:1995-2001.
72. Jilma B, Dirnberger E, Eichler HG, et al. Partial blockade of nitric oxide synthase blunts the exercise-induced increase of von Willebrand factor antigen and of factor VIII in man. Thrombosis and Haemostasis 1997;78:1268-1271.
73. Jern C, Eriksson E, Tengborn L, Risberg B, Wadenvik H, Jern S. Changes of plasma coagulation and fibrinolysis in response to mental stress. Thromb Haemost 1989;62:767-71.
74. Chang SJ, Koh SB, Cha BS, Park JK. Job characteristics and blood coagulation factors in Korean male workers. J Occup Environ Med 2002;44:997-1002.
75. Allman-Farinelli MA, Hall D, Kingham K, Pang D, Petocz P, Favaloro EJ. Comparison of the effects of two low fat diets with different alpha-linolenic:linoleic acid ratios on coagulation and fibrinolysis. Atherosclerosis 1999;142:159-68.
76. Meade TW, Brozovic M, Chakrabarth R, Howarth DJ, North WR, Stirling Y. An epidemiological study of the haemostatic and other effects of oral contraceptives. Br J Haematol 1976;34:353-64.
77. Kadir RA, Economides DL, Sabin CA, Owens D, Lee CA. Variations in coagulation factors in women: effects of age, ethnicity, menstrual cycle and combined oral contraceptive. Thromb Haemost 1999;82:1456-61.
78. Lowe GD, Rumley A, Woodward M, et al. Epidemiology of coagulation factors, inhibitors and activation markers: the Third Glasgow MONICA Survey. I. Illustrative reference ranges by age, sex and hormone use. Br J Haematol 1997;97:775-84.
79. Schobersberger W, Fries D, Mittermayr M, et al. Changes of biochemical markers and functional tests for clot formation during long-haul flights. Thromb Res 2002;108:19-24.
80. Conlan MG, Folsom AR, Finch A, et al. Associations of factor VIII and von Willebrand factor with age, race, sex, and risk factors for atherosclerosis. The Atherosclerosis Risk in Communities (ARIC) Study. Thromb Haemost 1993;70:380-5.
81. Geffken DF, Cushman M, Burke GL, Polak JF, Sakkinen PA, Tracy RP. Association between physical activity and markers of inflammation in a healthy elderly population. Am J Epidemiol 2001;153:242-50.
82. van der Salm P, Ubachs HM, van Wersch JW. Cord blood clotting factors in neonates of smoking and non-smoking mothers. Int J Clin Lab Res 1994;24:177-9.
83. Green D, Ruth KJ, Folsom AR, Liu K. Hemostatic factors in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Arterioscler Thromb 1994;14:686-93.
84. Korkmaz C, Ozdogan H, Kasapcopur O, Yazici H. Acute phase response in familial Mediterranean fever. Ann Rheum Dis 2002;61:79-81.
85. Reitsma PH, Branger J, Van Den Blink B, Weijer S, Van Der Poll T, Meijers JC. Procoagulant protein levels are differentially increased during human endotoxemia. J Thromb Haemost 2003;1:1019-23.

86. Gibbs NM, Crawford GP, Michalopoulos N. Postoperative changes in coagulant and anticoagulant factors following abdominal aortic surgery. J Cardiothorac Vasc Anesth 1992;6:680-5.
87. Elkeles RS, Chakrabarti R, Vickers M, Stirling Y, Meade TW. Effect of treatment of hyperlipidaemia on haemostatic variables. Br Med J 1980;281:973-4.
88. Tracy RP, Bovill EG, Fried LP, et al. The distribution of coagulation factors VII and VIII and fibrinogen in adults over 65 years. Results from the Cardiovascular Health Study. Ann Epidemiol 1992;2:509-19.
89. Lindeberg S, Berntorp E, Carlsson R, Eliasson M, Marckmann P. Haemostatic variables in Pacific Islanders apparently free from stroke and ischaemic heart disease--the Kitava Study. Thromb Haemost 1997;77:94-8.
90. Miller CH, Haff E, Platt SJ, et al. Measurement of von Willebrand factor activity: relative effects of ABO blood type and race. J Thromb Haemost 2003;1:2191-7.
91. Yarnell JW, Sweetnam PM, Rumley A, Lowe GD. Lifestyle factors and coagulation activation markers: the Caerphilly Study. Blood Coagul Fibrinolysis 2001;12:721-8.
92. Sweeney JD, Hoernig LA. Age-dependent effect on the level of factor IX. Am J Clin Pathol 1993;99:687-8.
93. Briet E, Bertina RM, van Tilburg NH, Veltkamp JJ. Hemophilia B Leyden: a sex-linked hereditary disorder that improves after puberty. N Engl J Med 1982;306:788-90.
94. Kurachi S, Deyashiki Y, Takeshita J, Kurachi K. Genetic mechanisms of age regulation of human blood coagulation factor IX. Science 1999;285:739-43.
95. Lindeberg S, Berntorp E, Nilsson-Ehle P, Terent A, Vessby B. Age relations of cardiovascular risk factors in a traditional Melanesian society: the Kitava Study. Am J Clin Nutr 1997;66:845-52.
96. Jeremic M, Weisert O, Gedde-Dahl TW. Factor VIII (AHG) levels in 1016 regular blood donors. The effects of age, sex, and ABO blood groups. Scand J Clin Lab Invest 1976;36:461-6.
97. Iso H, Folsom AR, Wu KK, et al. Hemostatic variables in Japanese and Caucasian men. Plasma fibrinogen, factor VIIc, factor VIIIc, and von Willebrand factor and their relations to cardiovascular disease risk factors. Am J Epidemiol 1989;130:925-34.
98. Cushman M, Yanez D, Psaty BM, et al. Association of fibrinogen and coagulation factors VII and VIII with cardiovascular risk factors in the elderly: the Cardiovascular Health Study. Cardiovascular Health Study Investigators. Am J Epidemiol 1996;143:665-76.
99. Cigolini M, Targher G, Bergamo Andreis IA, Tonoli M, Agostino G, De Sandre G. Visceral fat accumulation and its relation to plasma hemostatic factors in healthy men. Arterioscler Thromb Vasc Biol 1996;16:368-74.
100. Balleisen L, Assmann G, Bailey J, Epping PH, Schulte H, van de Loo J. Epidemiological study on factor VII, factor VIII and fibrinogen in an industrial population--II. Baseline data on the relation to blood pressure, blood glucose, uric acid, and lipid fractions. Thromb Haemost 1985;54:721-3.
101. Blann AD. How a damaged blood vessel wall contibutes to thrombosis and hypertenasion. Pathophysiol Haemost Thromb 2003;33:445-8.

102. Bern MM, Cassani MP, Horton J, Rand L, Davis G. Changes of fibrinolysis and factor VIII coagulant, antigen, and ristocetin cofactor in diabetes mellitus and atherosclerosis. Thromb Res 1980;19:831-9.
103. Hughes A, McVerry BA, Wilkinson L, Goldstone AH, Lewis D, Bloom A. Diabetes, a hypercoagulable state? Hemostatic variables in newly diagnosed type 2 diabetic patients. Acta Haematol 1983;69:254-9.
104. Herlihy OM, Barrow BA, Grant PJ, Levy JC. Hyperglycaemic siblings of Type II (non-insulin-dependent) diabetic patients have increased PAI-1, central obesity and insulin resistance compared with their paired normoglycaemic sibling. Diabetologia 2002;45:635-41.
105. Khrenov A, Sarafanov A, Ananyeva N, et al. Molecular basis for different ability of low-density and high-density lipoproteins to support activity of the intrinsic Xase complex. Thromb Res 2002;105:87-93.
106. Sie P, Caron C, Azam J, et al. Reassessment of von Willebrand factor (VWF), VWF propeptide, factor VIII:C and plasminogen activator inhibitors 1 and 2 during normal pregnancy. Br J Haematol 2003;121:897-903.
107. Stirling Y, Woolf L, North WR, Seghatchian MJ, Meade TW. Haemostasis in normal pregnancy. Thromb Haemost 1984;52:176-82.
108. Sanchez-Luceros A, Meschengieser SS, Marchese C, et al. Factor VIII and von Willebrand factor changes during normal pregnancy and puerperium. Blood Coagul Fibrinolysis 2003;14:647-51.
109. Brenner B. Haemostatic changes in pregnancy. Thromb Res 2004;114:409-14.
110. Miller CH, Dilley AB, Drews C, Richardson L, Evatt B. Changes in von Willebrand factor and factor VIII levels during the menstrual cycle. Thromb Haemost 2002;87:1082-3.
111. American Psychiatric Association., American Psychiatric Association. Task Force on DSM-IV. Diagnostic and statistical manual of mental disorders : DSM-IV-TR. Washington, DC: American Psychiatric Association, 2000.
112. Doris A, Ebmeier K, Shajahan P. Depressive illness. Lancet 1999;354:1369-75.
113. Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 2005;62:1097-106.
114. Kessler RC, Berglund P, Demler O, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). Jama 2003;289:3095-105.
115. Riolo SA, Nguyen TA, Greden JF, King CA. Prevalence of depression by race/ethnicity: findings from the National Health and Nutrition Examination Survey III. Am J Public Health 2005;95:998-1000.
116. Radloff LS. The CES-D Scale: A self-report depression scale for research in the general population. Applied Psychological Measurement 1977;1:385-401.
117. Wong YL. Measurement properties of the Center for Epidemiologic Studies-Depression Scale in a homeless population. Psychol Assess 2000;12:69-76.
118. McDowell I. Measuring health : a guide to rating scales and questionnaires. New York: Oxford University Press, 2006.

119. Kop WJ, Gottdiener JS, Tangen CM, et al. Inflammation and coagulation factors in persons > 65 years of age with symptoms of depression but without evidence of myocardial ischemia. Am J Cardiol 2002;89:419-24.
120. Eskandari F, Mistry S, Martinez PE, et al. Younger, premenopausal women with major depressive disorder have more abdominal fat and increased serum levels of prothrombotic factors: implications for greater cardiovascular risk. Metabolism 2005;54:918-24.
121. Shimbo D, Davidson KW, Haas DC, Fuster V, Badimon JJ. Negative impact of depression on outcomes in patients with coronary artery disease: mechanisms, treatment considerations, and future directions. J Thromb Haemost 2005;3:897-908.
122. Doulalas AD, Rallidis LS, Gialernios T, et al. Association of depressive symptoms with coagulation factors in young healthy individuals. Atherosclerosis 2006;186:121-5.
123. Panagiotakos DB, Pitsavos C, Chrysohoou C, et al. Inflammation, coagulation, and depressive symptomatology in cardiovascular disease-free people; the ATTICA study. Eur Heart J 2004;25:492-9.
124. McCallum CJ, Peake IR, Newcombe RG, Bloom AL. Factor VIII levels and blood group antigens. Thromb Haemost 1983;50:757.
125. Mohanty D, Ghosh K, Marwaha N, Kaur S, Chauhan AP, Das KC. Major blood group antigens--a determinant of factor VIII levels in blood? Thromb Haemost 1984;51:414.
126. Orstavik KH, Magnus P, Reisner H, Berg K, Graham JB, Nance W. Factor VIII and factor IX in a twin population. Evidence for a major effect of ABO locus on factor VIII level. Am J Hum Genet 1985;37:89-101.
127. Cadroy Y, Daviaud P, Saivin S, Sie P, Boneu B. Distribution of 16 hemostatic laboratory variables assayed in 100 blood donors. Nouv Rev Fr Hematol 1990;32:259-64.
128. Vlot AJ, Koppelman SJ, van den Berg MH, Bouma BN, Sixma JJ. The affinity and stoichiometry of binding of human factor VIII to von Willebrand factor. Blood 1995;85:3150-7.
129. Fischer BE, Kramer G, Mitterer A, et al. Effect of multimerization of human and recombinant von Willebrand factor on platelet aggregation, binding to collagen and binding of coagulation factor VIII. Thromb Res 1996;84:55-66.
130. van Hylckama Vlieg A, van der Linden IK, Bertina RM, Rosendaal FR. High levels of factor IX increase the risk of venous thrombosis. Blood 2000;95:3678-82.
131. Koster T, Blann AD, Briet E, Vandenbroucke JP, Rosendaal FR. Role of clotting factor VIII in effect of von Willebrand factor on occurrence of deep-vein thrombosis. Lancet 1995;345:152-5.
132. van der Meer FJ, Koster T, Vandenbroucke JP, Briet E, Rosendaal FR. The Leiden Thrombophilia Study (LETS). Thromb Haemost 1997;78:631-5.
133. Patel RK, Ford E, Thumpston J, Arya R. Risk factors for venous thrombosis in the black population. Thromb Haemost 2003;90:835-8.
134. Kyrle PA, Minar E, Hirschl M, et al. High plasma levels of factor VIII and the risk of recurrent venous thromboembolism. N Engl J Med 2000;343:457-62.

135. Cristina L, Benilde C, Michela C, Mirella F, Giuliana G, Gualtiero P. High plasma levels of factor VIII and risk of recurrence of venous thromboembolism. Br J Haematol 2004;124:504-10.
136. Ananyeva NM, Kouiavskaia DV, Shima M, Saenko EL. Intrinsic pathway of blood coagulation contributes to thrombogenicity of atherosclerotic plaque. Blood 2002;99:4475-85.
137. Kamphuisen PW, Eikenboom JC, Vos HL, et al. Increased levels of factor VIII and fibrinogen in patients with venous thrombosis are not caused by acute phase reactions. Thromb Haemost 1999;81:680-3.
138. Souto JC, Almasy L, Borrell M, et al. Genetic determinants of hemostasis phenotypes in Spanish families. Circulation 2000;101:1546-51.
139. Bank I, Libourel EJ, Middeldorp S, et al. Elevated levels of FVIII:C within families are associated with an increased risk for venous and arterial thrombosis. J Thromb Haemost 2005;3:79-84.
140. Cortellaro M, Boschetti C, Cofrancesco E, et al. The PLAT Study: hemostatic function in relation to atherothrombotic ischemic events in vascular disease patients. Principal results. PLAT Study Group. Progetto Lombardo Atero-Trombosi (PLAT) Study Group. Arterioscler Thromb 1992;12:1063-70.
141. Scanavini D, Legnani C, Lunghi B, Mingozzi F, Palareti G, Bernardi F. The factor VIII D1241E polymorphism is associated with decreased factor VIII activity and not with activated protein C resistance levels. Thromb Haemost 2005;93:453-6.
142. Nossent AY, Eikenboom JC, Vos HL, et al. Haplotypes encoding the factor VIII 1241 Glu variation, factor VIII levels and the risk of venous thrombosis. Thromb Haemost 2006;95:942-8.
143. Viel KR, Machiah DK, Warren DM, et al. A sequence variation scan of the coagulation factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels. Blood 2007;109:3713-3724.
144. Pittman DD, Alderman EM, Tomkinson KN, Wang JH, Giles AR, Kaufman RJ. Biochemical, immunological, and in vivo functional characterization of B-domain-deleted factor VIII. Blood 1993;81:2925-35.
145. El-Maarri O, Singer H, Klein C, et al. Lack of F8 mRNA: a novel mechanism leading to hemophilia A. Blood 2006;107:2759-65.
146. Strickland DK, Medved L. LRP-mediated clearance of activated blood coagulation co-factors and proteases: clearance mechanism or regulation? Journal of Thrombosis and Haemostasis;0.
147. Zhang B, Kaufman RJ, Ginsburg D. LMAN1 and MCFD2 Form a Cargo Receptor Complex and Interact with Coagulation Factor VIII in the Early Secretory Pathway. J Biol Chem 2005;280:25881-6.
148. Viel K, Machiah DK, Warren DM, et al. A Sequence Variation Scan of the Coagulation Factor (F)VIII Structural Gene and Associations with Plasma FVIII Activity (FVIII:C) Levels. 2006.
149. SeattleSNPs. NHLBI Program for Genomic Applications, UW-FHCRC, Seattle, WA (URL: http://pga.gs.washington.edu) [May, 2006].

150. Dilley A, Drews C, Miller C, et al. von Willebrand disease and other inherited bleeding disorders in women with diagnosed menorrhagia. Obstet Gynecol 2001;97:630-6.
151. Miller CH, Dilley A, Richardson L, Hooper WC, Evatt BL. Population differences in von Willebrand factor levels affect the diagnosis of von Willebrand disease in African-American women. Am J Hematol 2001;67:125-9.
152. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615-25.
153. Cote I, Jacobs P, Cumming D. Work loss associated with increased menstrual loss in the United States. Obstet Gynecol 2002;100:683-7.
154. Colman RW. Hemostasis and thrombosis : basic principles and clinical practice. Philadelphia: Lippincott Williams & Wilkins, 2001.
155. Neter J, Kutner KH, Wasserman W, Nachtsheim CJ. Applied linear statistical models. Chicago: Irwin, 1996.
156. Rudisch B, Nemeroff CB. Epidemiology of comorbid coronary artery disease and depression. Biol Psychiatry 2003;54:227-40.
157. Creed F. The importance of depression following myocardial infarction. Heart 1999;82:406-8.
158. Ziegelstein RC. Depression after myocardial infarction. Cardiol Rev 2001;9:45-51.
159. Steffens DC, Krishnan KR, Crump C, Burke GL. Cerebrovascular disease and evolution of depressive symptoms in the cardiovascular health study. Stroke 2002;33:1636-44.
160. Krishnan KR. Depression as a contributing factor in cerebrovascular disease. Am Heart J 2000;140:70-6.
161. Ostir GV, Markides KS, Peek MK, Goodwin JS. The association between emotional well-being and the incidence of stroke in older adults. Psychosom Med 2001;63:210-5.
162. Robinson RG. Poststroke depression: prevalence, diagnosis, treatment, and disease progression. Biol Psychiatry 2003;54:376-87.
163. Williams LS. Depression and stroke: cause or consequence? Semin Neurol 2005;25:396-409.
164. Zgraggen L, Fischer JE, Mischler K, Preckel D, Kudielka BM, von Kanel R. Relationship between hemoconcentration and blood coagulation responses to acute mental stress. Thromb Res 2005;115:175-83.
165. Wang TD, Chen WJ, Su SS, et al. Increased levels of tissue plasminogen activator antigen and factor VIII activity in nonvalvular atrial fibrillation: relation to predictors of thromboembolism. J Cardiovasc Electrophysiol 2001;12:877-84.
166. McDowell I, Newell C. Measuring health : a guide to rating scales and questionnaires. New York: Oxford University Press, 1996.
167. Flanders WD, DerSimonian R, Freedman DS. Interpretation of linear regression models that include transformations or interaction terms. Ann Epidemiol 1992;2:735-44.
168. Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology 2001;12:313-20.

169. Harris T. Depression in women and its sequelae. J Psychosom Res 2003;54:103-12.
170. Hurskainen R, Aalto AM, Teperi J, et al. Psychosocial and other characteristics of women complaining of menorrhagia, with and without actual increased menstrual blood loss. Bjog 2001;108:281-5.
171. Globe DR, Curtis RG, Koerper MA. Utilization of care in haemophilia: a resource-based method for cost analysis from the Haemophilia Utilization Group Study (HUGS). Haemophilia 2004;10 Suppl 1:63-70.
172. Blangero J, Goring HH, Kent JW, Jr., et al. Quantitative trait nucleotide analysis using Bayesian model selection. Hum Biol 2005;77:541-59.