Genotypic exploration of the fibrinogen phenotype in a black South African population

HT Cronjé
23520825

Dissertation submitted in fulfilment of the requirements for the degree *Master of Science* in Nutrition at the Potchefstroom Campus of the North-West University

Supervisor: Dr C Nienaber-Rousseau
Co-supervisor: Prof. M Pieters
Co-supervisor Dr Z de Lange

November 2016

ACKNOWLEDGEMENTS

Firstly, I would like to thank the entire research team; it has been an honour to be your postgraduate student, and I am so thankful for all the opportunities you have given me. To all the co-authors, Lizelle, Tinashe, Tertia and Fiona, thank you for your individual contributions to this dissertation, it was a privilege to work with each one of you.

Furthermore, I would like to express my sincerest gratitude to the following individuals:

My parents, you are my greatest blessing. Thank you for always believing in me and for granting me the opportunity to further my studies. Your unconditional love and support from the smallest to the biggest things mean everything to me, and I love you dearly. **Cornelie**, thank you for your passion, and the drive and enthusiasm you teach and encourage with. You are an inspiration, the calm in the storm, and I appreciate you more than words can say. **Prof. Marlien**, thank you for your patience and the valuable lessons and trade secrets you have shared with me. I truly had the privilege to learn from the best. You are a remarkable researcher and mentor; thank you for your leadership. **Zelda**, you are a constant source of guidance and reassurance. I have learnt a lot from your work ethic and the precision and effort you do every task, big or small with. **Lizelle**, thank you for all the encouragement you gave throughout the journey and for always making time to try and explain the (sometimes) unexplainable. I appreciate everything you have taken the time to teach me, and the sincerity with which you did it. **Ellenor**, thank you for patiently teaching me about all things ELISA and life. This year would not have been the same without you and you have shaped a great part of the years ahead.

With words borrowed from the National Institutes of Health's director Francis Collins who led the international human genome project: "The God of the Bible is also the God of the genome. He can be worshiped in the cathedral or in the laboratory. His creation is majestic, awesome, intricate and beautiful." What a privilege to study the genome with the God who created the genome. **For from Him and to Him are all things. In Him we live and move and have our being.** To Him be the glory forever! Amen.

ABSTRACT

INTRODUCTION AND AIM

Increased total and γ' fibrinogen concentrations are associated with cardiovascular disease (CVD) risk in part through their effects on fibrin clot properties. Interleukin-6 (IL-6) promotes the expression of fibrinogen and is also independently associated with CVD. The fibrinogen phenotype is heritable, although discrepancies exist between the outcome of heritability studies and the ability of genome-wide association studies to identify contributing polymorphisms. This dissertation aimed to address the 'missing heritability' of fibrinogen by investigating an African population known to have higher fibrinogen and IL-6 concentrations, together with greater genetic variability and lower linkage disequilibrium (LD) than previously reported in Europeans. Three approaches were followed; firstly, polymorphisms and haplotypes within the fibrinogen gene cluster were investigated with the aim of identifying functional variants that had not been identified due to high LD in the fibrinogen gene cluster in Europeans; secondly, as genetic variation within the fibrinogen genes alters the magnitude of the IL-6-induced expression of fibrinogen, the interactive effect of fibrinogen polymorphisms and IL-6 was investigated; and lastly, the pleiotropic and polygenic co-regulation of fibrinogen was investigated in a candidate gene analysis of multiple targeted genes, largely outside the fibrinogen genes. These approaches were investigated in terms of total and γ' fibrinogen, as well as their functional effects through turbidity-derived indicators of clot formation, structure and lysis.

METHODS

Eighty-one single nucleotide polymorphisms (SNPs) were investigated in 2010 apparently healthy Tswana individuals. Genotyping was performed through restriction fragment length polymorphism techniques, TaqMan-based assays, the beadXpress® platform and competitive allele-specific polymerase chain reaction methods. Fourteen SNPs were located in the fibrinogen gene cluster and 67 spanned the *APOB*, *APOE*, *CBS*, *CRP*, *F13A1*, *LDL-R*, *MTHFR*, *MTR*, *PCSK-9* and *SERPINE-1* genes. Total and γ' fibrinogen concentrations were quantified *via* the modified Clauss and enzyme-linked immunosorbent assay methods, respectively. IL-6 was quantified by means of automatic electrochemiluminescence and fibrin clot properties were determined through turbidimetric analyses. Independent and IL-6-interactive associations of the 14 fibrinogen SNPs with the outcome phenotypes were determined. In addition, 78 candidate SNPs were investigated in terms of their individual and accumulative associations (through genetic risk score analyses) with the outcome phenotypes together with possible co-regulatory processes as a result of the gain and loss of transcription factor binding sites (TFBS).

RESULTS

Lower minor allele frequencies and higher recombination rates for the investigated SNPs were observed in our study population compared to what has previously been reported for Europeans. None of the common European fibrinogen haplotypes was present. Seven of the fibrinogen SNPs were significantly associated with one or more of the outcome phenotypes. The fibrinogen SNPs contributed 0.5% of the variance in total fibrinogen. Fibrinogen significantly associated with IL-6, and thereby mediated associations with clot formation and structure. Several significant interactions between the fibrinogen SNPs and IL-6 were observed relating to total and γ' fibrinogen and fibre diameter. These interactions were additive in their association with total fibrinogen concentrations. *FGB*-rs7439160, -1420G/A and -148C/T were acknowledged as SNPs possibly functional in the PURE population. The candidate gene analysis revealed SNPs within and outside of the fibrinogen gene cluster to associate with fibrinogen and clot-related phenotypes, including SNPs located in the *F13A1*, *LDL-R*, *PCSK-9*, *CBS* and *CRP* genes, through the regulatory effects induced by the gain and loss of 75 TFBS. An accumulative genetic risk was observed through genetic risk scores that were significantly associated with all phenotypes apart from fibre diameter.

CONCLUSION

This dissertation highlights the distinctive African genome and stresses the importance of conducting genetic research among Africans. Original contributions to the literature include the investigation of three novel SNPs, a report of a lack of haplotypes in the fibrinogen genes and an additive effect of risk alleles within the fibrinogen gene cluster of Africans, as well as evidence of the involvement of *PCSK-9* SNPs in the heritability of fibrinogen concentration. Furthermore, evidence is given that polygenic transcriptional co-regulation of these SNPs through their effects on TFBS forms the basis of their associations with the respective phenotypes. The current study contributes to the investigation of fibrinogen's missing heritability by widening the scope of involved genes from what has been discovered thus far, as well as providing a new avenue for the exploration of transcriptional co-regulation of SNPs outside of, and additive gene-environment contributions within the fibrinogen gene cluster.

KEYWORDS FGA, FGB, FGG, thrombosis, genetics of haemostasis, interleukin-6, inflammation, fibrinogen gamma prime, gene-environment interactions, turbidity

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ..... **I**

ABSTRACT ..... **II**

LIST OF TABLES ..... **VIII**

LIST OF FIGURES ..... **X**

LIST OF SYMBOLS AND ABBREVIATIONS ..... **XI**

INTRODUCTION **2**

1.1 BACKGROUND **2**

1.2 AIMS AND OBJECTIVES **6**

1.3 STRUCTURE OF THE DISSERTATION **6**

1.4 RESEARCH TEAM **7**

LITERATURE REVIEW **11**

2.1 INTRODUCTION **11**

2.2 FIBRINOGEN: BIOCHEMISTRY, PHYSIOLOGY AND PATHOPHYSIOLOGY **12**

2.2.1 Biochemical composition of the fibrinogen protein 12

2.2.2 The (patho)physiology of fibrinogen 13

2.2.2.1 The (patho)physiology of fibrinogen: haemostasis and thrombogenesis 14

2.2.2.2 The (patho)physiology of fibrinogen: inflammation and atherogenesis 17

2.2.2.3 Fibrinogen and cardiovascular disease risk 18

2.3 GENETICS OF FIBRINOGEN **19**

2.3.1 The molecular composition of fibrinogen 19

2.3.2 Transcriptional regulation of fibrinogen 21

2.3.3 Common genetic variation within the fibrinogen gene cluster 24

2.3.4 Common fibrinogen splice variants 28

2.3.5 The heritability of fibrinogen 31

2.4 ADDRESSING THE MISSING HERITABILITY OF FIBRINOGEN 34

2.4.1 Genetic linkage, linkage disequilibrium and haplotypes 34

2.4.2 Gene-environment interactions 37

2.4.3 Pleiotropic and polygenic co-regulation 39

2.5 CONCLUSION 43

INDEPENDENT AND IL-6-INTERACTIVE ASSOCIATIONS OF SELECTED FIBRINOGEN POLYMORPHISMS IN PREDICTING FIBRINOGEN AND CLOT-RELATED PHENOTYPES 44

GUIDE FOR AUTHORS: BRITISH JOURNAL OF HAEMATOLOGY 45

ARTICLE 53

SUMMARY 54

INTRODUCTION 55

MATERIALS AND METHODS 57

Study population and ethical considerations 57

Blood collection and storage 57

Biochemical analyses 57

DNA isolation, SNP selection and genotyping 58

Statistical analyses 59

RESULTS 61

Description of individual polymorphisms 61

LD and haplotype construction 61

Associations of individual polymorphisms with fibrinogen-related phenotypes 62

The association of IL-6 with fibrinogen-related phenotypes 64

Genotype-IL-6 interactions in terms of phenotype predictions 66

DISCUSSION 70

ACKNOWLEDGEMENTS 74

AUTHOR CONTRIBUTIONS 74

SUPPORTING INFORMATION 75

REFERENCES 75

CANDIDATE GENE ANALYSIS OF THE FIBRINOGEN PHENOTYPE REVEALS THE IMPORTANCE OF POLYGENIC CO-REGULATION 89

GUIDE FOR AUTHORS: MATRIX BIOLOGY 90

LETTER OF ACCEPTANCE FOR PUBLICATION IN MATRIX BIOLOGY 104

ARTICLE 105

ABSTRACT 106

INTRODUCTION 107

RESULTS AND DISCUSSION 109

Individual association of SNPs with fibrinogen variables 111

Polygenic association of SNPs with fibrinogen variables 114

Polygenic co-regulation 115

METHODS 117

Study population 117

Blood sampling and biochemical analyses 117

Selection of genetic variants 118

Genotyping 118

Statistical analysis 119

Quality control and statistical analysis of genetic data 119

In silico network analyses 120

AUTHOR CONTRIBUTION 120

CONFLICT OF INTEREST 121

ACKNOWLEDGEMENTS 121

REFERENCES 122

CONCLUSIONS AND RECOMMENDATIONS 140

5.1 INTRODUCTION 140

5.2 GENETIC COMPOSITION OF THE PURE POPULATION 141

5.3 FUNCTIONALITY OF FIBRINOGEN SNPS AND HAPLOTYPES IN TERMS OF THEIR ASSOCIATION WITH FIBRINOGEN-RELATED PHENOTYPES 141

5.4 PLEIOTROPIC AND POLYGENIC REGULATION OF THE FIBRINOGEN PHENOTYPES BY SNPS BEYOND THOSE IN THE FIBRINOGEN GENES 143

5.5 CONCLUSION 144

BIBLIOGRAPHY 145

ANNEXURE A 177

LIST OF TABLES

Table 1.1 Individual contributions of the research team members to this dissertation 8

Table 2.1 The association of selected fibrinogen polymorphisms with total fibrinogen concentrations and (patho)physiological outcomes..... 25

Table 2.2 Studies performed to estimate the heritability of fibrinogen 32

Table 2.3 Summary of genome-wide association studies with fibrinogen as phenotype 40

Table 3.1 Association of selected upstream *FGB* polymorphisms with fibrinogen-related phenotypes..... 63

Table 3.2 Outcome phenotypes descriptive statistics and association with IL-6- quartiles 65

Table 3.3 Genotype-IL-6 interactions modulating total and γ' fibrinogen concentrations..... 67

Table 3.4 Genotype-IL-6 interactions modulating clot properties..... 68

Table 3.5 Primer and synthetic control sequences used for KASP analyses..... 85

Table 3.6 Associations of individual SNPs with fibrinogen γ′, γ′ ratio and total fibrinogen as published by Kotzé et al, (2015)..... 86

Table 3.7 Associations of individual SNPs with clot-related phenotypes as published by Kotzé et al, (2015)..... 87

Table 3.8 Genotype-IL-6 interactions modulating total and γ' fibrinogen concentrations upon removal of individuals with IL-6 > 100pg/mL..... 88

Table 4.1 Descriptive characteristics of the study participants 109

Table 4.2 Phenotypes correlating significantly with total and γ' fibrinogen concentration and clot properties 110

Table 4.3 SNPs significantly associated with the fibrinogen phenotypes..... 112

Table 4.4 Contribution of genetic risk scores to phenotypes 115

Table 4.5 Supplementary Table 1: SNPs, their minor allele frequencies and the methods used for genotyping 129

Table 4.6 Supplementary Table 2: Single nucleotide polymorphisms and the gain and loss of transcription factor binding sites 133

LIST OF FIGURES

Figure 2.1 Biochemical composition of the fibrinogen protein 13

Figure 2.2 Blood coagulation cascade 15

Figure 2.3 Composition of the fibrinogen gene cluster 20

Figure 2.4 Transcriptional regulation of the fibrinogen gene cluster 23

Figure 2.5 High, low and very low molecular weight fibrinogen molecules 29

Figure 2.6 Alternative splicing mechanism of the fibrinogen γ chain 31

Figure 2.7 Human karyotype indicating loci significantly associated with the fibrinogen phenotype 41

Figure 3.1 Fourteen polymorphisms spanning the fibrinogen gene cluster 61

Figure 3.2 Pairwise LD structure of 14 fibrinogen SNPs, illustrated by D' values on an r^2 colour scheme 62

Figure 3.3 CLT across IL-6 quartiles before (A) and after (B) adjustment for total fibrinogen, γ' fibrinogen and PAI-1_{act} 66

Figure 3.4 Association of total fibrinogen concentrations with circulating interleukin-6 by risk score groups 69

Figure 4.1 The in silico portrayal of the 13 genes included in a CVD network 111

LIST OF SYMBOLS AND ABBREVIATIONS

A	Adenine
α	Alpha
Aα	A alpha
αC	Alpha C
ACTN1	Actinin alpha 1
AFR	African
ANCOVA	Analysis of co-variance
ANOVA	Analysis of variance
Apo-B	Apolipoprotein-B
APOB	Apolipoprotein-B (gene)
Apo-E	Apolipoprotein-E
APOE	Apolipoprotein-E (gene)
APR	Acute phase response
ASI	East Asian
au	Absorbance units
β	Beta
BMI	Body mass index
bp	Base pairs
c.	Coding
C	Cytosine
CBS	Cystathionine beta synthase
CD300LF	CD300 molecule like family member F
C/EPB	CCAAT enhancer-binding protein
CHD9	Chromodomain helicase deoxyribonucleic acid binding protein 9
CHD	Coronary heart disease
CI	Confidence interval
CLT	Clot lysis time

cm	Centimetres
CPS1	Carbamoyl-phosphate synthase 1
CPT1B	Carnitine palmitoyltransferase 1b
CRP	C-reactive protein
CVD	Cardiovascular disease
D	Distal regions
D’	Standardised disequilibrium
DIP2B	Disco interacting protein 2 homolog B
DNA	Deoxyribonucleic acid
DVT	Deep vein thrombosis
EA	European American
F	Factor
FARP2	FERM, ARH/RhoGEF and pleckstrin domain protein 2
F13A1	Factor XIII (gene)
FGA	Fibrinogen alpha chain gene
FGB	Fibrinogen beta chain gene
FGG	Fibrinogen gamma chain gene
FPA	Fibrinopeptide A
FPB	Fibrinopeptide B
FXIII	Factor XIII
γ	Gamma
γA	Gamma A
γ’	Gamma prime
g.	Genomic
g	Gram
g/L	Gram per litre
G	Guanine
GRE	Glucocorticoid responsive enhancer element

GRS	Genetic risk score
GWAS	Genome-wide association studies
$HbA1_C$	Glycated haemoglobin
Hcy	Homocysteine
HDLBP	High-density lipoprotein binding protein
HDL-c	High-density lipoprotein cholesterol
HepG2	Hepatocellular carcinoma cell lines
HGFAC	Hepatocyte growth factor activator
Hip C	Hip circumference
HIS	Hispanic
HIV	Human immunodeficiency virus
HMW	High molecular weight
HNF1	Hepatocyte nuclear factor 1
HNF4α	Hepatocyte nuclear factor 4 alpha
HWE	Hardy-Weinberg equilibrium
ICAM-1	Intracellular adhesion molecule-1
IHD	Ischemic heart disease
IL-1β	Interleukin-1β
IL1R1	Interleukin 1 receptor, type I
IL1RN	Interleukin 1 receptor antagonist
IL-6	Interleukin-6
IL6R	Interleukin 6 receptor
IL-6RE(s)	Interleukin-6 responsive element(s)
IRF1	Interferon regulatory factor 1
JMJD1C	Jumonji domain containing 1C
KASP	Competitive allele-specific polymerase chain reaction
kb	Kilo base pairs
kDA	Kilo Dalton
kg/m^2	kilogram per square metre of body weight

L	Litre
LD	Linkage disequilibrium
LDL-c	Low-density lipoprotein cholesterol
LDL-R	Low density lipoprotein receptor
LEPR	Leptin receptor
LMW	Low molecular weight
µl	Microliter
Mac-1	Macrophage-1 antigen
MAF(s)	Minor allele frequency(ies)
mg/L	Milligram per litre
MI	Myocardial infarction
mmHg	Millimetre of mercury
mmol/L	Millimoles per litre
mRNA	Messenger ribonucleic acid
MS4A6A	Membrane Spanning 4-Domains A6A
MTHFR	Methylenetetrahydrofolate reductase
MTR	Methionine synthase
MW	Molecular weight
N	Sample/population size
NAT9	N-Acetyltransferase 9
ng/mL	Nanogram per millilitre
NIDDM	Non-insulin dependent Diabetes Mellitus
NFKβ	Nuclear factor kappa β
NLRP3	Nucleotide-binding domain and leucine-rich repeat family containing pyrin domain containing 3
nm	Nanometre
NWU	North-West University
p	Indicator of statistical significance
PAI-1	Plasminogen activator inhibitor type 1

$PAI\text{-}1_{act}$	Plasminogen activator inhibitor type 1 activity
PA index	Physical activity index
PCCB	Propionyl-coa carboxylase beta subunit
PCR	Polymerase chain reaction
PCSK-9	Proprotein convertase subtilisin/kexin type 9
PDLIM4	PDZ and LIM domain protein 4
PLEC1	Plectin 1
PSMG1	Proteasome assembly chaperone 1
pg/mL	Picogram per millilitre
PURE	Prospective Urban and Rural Epidemiology
QC	Quality control
r	Correlation coefficient
r^2	Correlation coefficient squared
RFLP	Restriction fragment length polymorphism
rs	Reference sequence
RT-PCR	Real time polymerase chain reaction
SBP	Systolic blood pressure
SD	Standard deviation
SERPINE-1	Plasminogen activator inhibitor type 1 (gene)
SHANK3	SH3 and multiple ankyrin repeat domains 3
SH2B3	SH2B adaptor protein 3
SLC22A4	Solute carrier family 22 member 4
SLC22A5	Solute carrier family 22 member 5
SLC9A3R1	Solute carrier family 9 member a3 regulator 1
SMAC	Sequential Multiple Analyser Computer
SNP(s)	Single nucleotide polymorphism(s)
SOCS-3	Suppressor of cytokine signalling-3
SPPL2A	Signal peptide peptidase like 2A
SPSS	Statistical package for the social sciences

STAT-3	Signal transducer and activator of transcription-3
T	Thymine
TC	Total cholesterol
TF	Tissue factor (Chapter 3)
TF	Transcription factor (Chapter 4)
TFBS	Transcription factor binding site
TG	Triglycerides
TLR	Toll-like receptor
TNF-α	Tumor necrosis factor-α
TOMM7	Translocase of outer mitochondrial membrane 7
tPA	Tissue plasminogen activator
TSS	Transcription start site
µmol	Micromole
U/mL	Units per millilitre
uPA	Urokinase plasminogen activator
USF	Upstream stimulatory factor
UTR	Untranslated region
VLMW	Very low molecular weight

CHAPTER 1

INTRODUCTION

INTRODUCTION

1.1 BACKGROUND

Central to blood coagulation, fibrinogen is the precursor of fibrin, the main constituent of a blood clot (Blombäck, 1996; Herrick *et al.*, 1999). Fibrinogen is a hexameric protein, composed of two sets of non-identical, alpha (α), beta (β) and gamma (γ) polypeptide chains (Mosesson *et al.*, 2001). These chains are independently coded for by the α (*FGA*), β (*FGB*) and γ (*FGG*) chain genes, on the q-arm of chromosome four (Henry *et al.*, 1984). The γ chain is subject to alternative splicing and polyadenylation and two common variants, fibrinogen gamma A (γA) and gamma prime (γ') exist. Fibrinogen γ' contributes 8 to 15% of total plasma fibrinogen concentrations (Wolfenstein-Todel & Mosesson, 1980; Chung & Davie, 1984), and has a higher molecular weight due to alternative splicing in the carboxyl-terminal region of the γ polypeptide chain. The alternative splicing results in the translation of a 20-amino acid sequence of intron 9, that replaces the four-amino acid sequence of exon 10 (Wolfenstein-Todel & Mosesson, 1980; Chung & Davie, 1984).

Fibrinogen occurs in the circulation at concentrations of 1.5 to 4.5 g/L, and is involved in the inflammatory and blood coagulation processes, as an acute phase, and haemostatic protein, respectively (Clark *et al.*, 1982; Donaldson *et al.*, 1989; Sahni *et al.*, 1998; Sahni & Francis, 2000; Kamath & Lip, 2003). An increased fibrinogen concentration is an independent risk factor for cardiovascular pathology through its atherogenic and thrombogenic properties (Koenig, 2003; Ariëns, 2013). Fibrinogen contributes to atherogenesis by accumulating in atherosclerotic plaque, increasing both plaque growth and instability while promoting an endothelial environment that enhances atherogenic abilities (Feinbloom & Bauer, 2005; Ariëns, 2013). Increased fibrinogen concentrations also enhance fibrin clot stability by increasing clot size, density and stability (thrombogenesis), thereby suppressing the fibrinolytic process, allowing formed thrombi to remain intact in the vasculature for a longer time (Weisel & Nagaswami, 1992; Machlus *et al.*, 2011). Numerous studies, as reviewed by Undas and Ariëns (2011), have reported on the association between various clot properties (including effects on clot lysis), and cardiovascular disease (CVD) outcomes such as atherosclerosis, deep vein thrombosis, myocardial infarction, stroke and coronary heart disease. A dense clot structure consisting of compact, highly branched, thin fibres is more resistant to clot lysis and, therefore, has a greater association with CVD (Ariëns, 2011; Machlus *et al.*, 2011; Undas & Ariëns, 2011). Irrespective of the intermediate functional phenotypes, both total and γ' fibrinogen are independently associated with increased risk of CVD (Stec *et al.*, 2000; Lovely *et al.*, 2002; Danesh *et al.*,

2005; Mannila *et al.*, 2007; Cheung *et al.*, 2008; Cheung *et al.*, 2009; Tousoulis *et al.*, 2011; Macrae *et al.*, 2016), although causality has not been confirmed (Keavney *et al.*, 2006; Meade *et al.*, 2006; Ken-Dror *et al.*, 2012; Sabater-Lleal *et al.*, 2013).

The prevalence of CVD in South Africa is steadily increasing (Statistics South Africa, 2016), and it is currently one of the leading causes of mortality globally (Mozaffarian *et al.*, 2016). A recent longitudinal investigation of CVD risk progression in South Africa reported a worsening CVD risk over time, although the distribution of risk factors varied between ethnic groups. The authors observed a predisposition to hypercoagulability in black study participants, with greater increases in total fibrinogen concentrations in the follow-up period compared to their white counterparts (Hamer *et al.*, 2015). In agreement with the Hamer study, black ethnicities have been shown to have higher fibrinogen concentrations than whites in South Africa (Greyling *et al.*, 2007; Pieters & Vorster, 2008; Lammertyn *et al.*, 2015) and internationally (Folsom *et al.*, 1991; Albert *et al.*, 2009; Wassel *et al.*, 2011). This is already prevalent at adolescent age (Nienaber *et al.*, 2008).

Family and twin studies investigating the genetics of fibrinogen in multiple ethnic groups have reported the heritability of fibrinogen concentrations to be between 20 and 51% (Hamsten *et al.*, 1987; Reed *et al.*, 1994; Friedlander *et al.*, 1995; Livshits *et al.*, 1996; de Lange *et al.*, 2001). A meta-analysis conducted by the Fibrinogen Studies Collaboration reported the cumulative ability of modifiable lifestyle characteristics known to influence fibrinogen concentrations, including physical activity, body mass index, smoking and alcohol consumption habits, to explain about half of fibrinogen's variation (Kaptoge *et al.*, 2007). The authors added that aiming to alter elevated fibrinogen concentrations with lifestyle modification only might be a feeble attempt (Kaptoge *et al.*, 2007).

Acknowledging the reports of great heritability in a pathologically relevant protein, researchers set out to identify the single nucleotide polymorphisms (SNPs) associated with the heritable component of the fibrinogen phenotype. With limited success, the totality of genome-wide association studies (GWAS) have only accounted for 3.7% of the variation in plasma fibrinogen by common SNPs to date (Sabater-Lleal *et al.*, 2013). The vast discrepancy between heritability and association studies has made a case for researchers to try to find the missing heritability of fibrinogen.

The present study aimed to address the missing heritability of fibrinogen in a study population suitable to overcome some of the barriers in the literature thus far, as will be discussed below. The study population was the South African arm of the Prospective Urban and Rural

Epidemiology (PURE) study, consisting of self-identified black Tswana-speaking individuals in the North West province. The high fibrinogen concentrations and great genetic variability in this population made it ideal for investigating the missing heritability, as both the phenotype and genotype are different from those of Europeans, among whom most of the research on fibrinogen's heritability has been conducted (Wassel *et al.*, 2011; Kotzé *et al.*, 2015). This study adopted three primary approaches to explore the missing heritability of the fibrinogen phenotype in an African population.

Firstly, polymorphisms within the fibrinogen gene cluster identified in the literature to have possible functional effects were investigated. These polymorphisms were investigated in terms of their individual association with the fibrinogen phenotypes, as well as in terms of the linkage disequilibrium (LD) and haplotypes between them. This is the first report of LD and haplotypes within the fibrinogen gene cluster in this African study population and is of international relevance, as it has previously not been possible to identify the true functional variants because of very high LD in the fibrinogen genes in Europeans (Green, 2001; Baumann & Henschen, 1994; Behague *et al.*, 1996; Mannila *et al.*, 2005; Verschuur *et al.*, 2005). Studies on the population-specific genetic variation have revealed Africans to have genetic diversity greater than any other population in the world (Chen *et al.*, 1995; Schuster *et al.*, 2010; Teo *et al.*, 2010). Low LD in the fibrinogen genes has been reported in the PURE study population specifically (Kotzé *et al.*, 2015). Research in African populations might, therefore, be able to identify which of the highly linked SNPs in Europeans are in fact causally associated with fibrinogen concentrations and functionality, and truly contribute to its heritability. Furthermore, the possibility of an additive effect when harbouring more than one truly independent risk allele was investigated for the first time, aiming to explain a greater part of the fibrinogen variance.

The second approach was the investigation of gene-environment interactions. The regulation of fibrinogen through genetic (up to 51%) and environmental influences (up to 64%) has been reported (Manolio *et al.*, 2009; Sabater-Lleal *et al.*, 2013). Environmental factors alter the magnitude of the effect of polymorphic variance on fibrinogen expression (Humphries *et al.*, 1997; Lim *et al.*, 2003; Baumert *et al.*, 2014). Investigating the interaction between these two great influencers could explain a larger part of the variance than their separate contribution could (Manolio *et al.*, 2009). The gene-environment interaction that was focused on was the interaction of fibrinogen polymorphisms with interleukin-6 (IL-6). The interaction between inflammatory markers, specifically IL-6 (as one of the main fibrinogen production stimuli), and the fibrinogen genes has been greatly explored in the global literature (Anderson *et al.*, 1993; Dalmon *et al.*, 1993; Verschuur *et al.*, 2005; Fish & Neerman-Arbez, 2012). In response to injury or physiological stress, IL-6 triggers the inflammatory response through the production of

acute phase proteins, such as fibrinogen, by hepatocytes (Heinrich *et al.*, 1990). The fibrinogen genes carry upstream IL-6 responsive elements that alter the level of fibrinogen expression *in vivo* (Fuller & Zhang, 2001; Fish & Neerman-Arbez, 2012). In addition to total fibrinogen regulation, IL-6 is also able to up-regulate the production of γ' fibrinogen independently (Alexander *et al.*, 2011; Rein-Smith *et al.*, 2013). Evidence for the differential expression of the fibrinogen genotypes and haplotypes during the acute phase influencing gene- and protein expression is clear (Verschuur *et al.*, 2005; Morozumi *et al.*, 2009), although the effect of IL-6 on fibrinogen genotype behaviour in Africans has not been determined

The third approach was to explore SNPs beyond those in the fibrinogen gene cluster we hypothesise to have pleiotropic associations with the fibrinogen phenotype, and therefore might be able to regulate fibrinogen expression polygenically. This approach is rooted in numerous GWAS identifying more associations with the fibrinogen phenotype outside of, than within the fibrinogen gene cluster (de Vries et al., 2016). Fibrinogen-related GWAS have only been performed in European and ad-mixed population groups, from whom the extrapolation of genetic data to the unique African genome is not possible (Tishkoff *et al.*, 2009; Teo *et al.*, 2010). In addition, the lack of a genome-wide chip for ethnic subgroups in Africa hinders the possibility of pursuing a genome-wide approach (Teo *et al.*, 2010). Consequently, a candidate gene approach was followed to overcome the barrier of the lack of a genome-wide method, and strengthened by the ability to identify candidate SNPs of plausible relevance. The approach involved the identification of SNPs that code for proteins significantly associated with the fibrinogen phenotype that might regulate fibrinogen concentrations on a molecular, rather than protein level. Biochemical markers are susceptible to a large amount of variance due to lifestyle, metabolism, environmental factors, season or physiological stress. Genes, on the other hand, are constant predictors of a baseline phenotype unchangeable by the above-mentioned factors (Keavney *et al.*, 2006). The genotype underlying several phenotypes known to be associated with fibrinogen may, therefore, provide greater mechanistic insight into the high fibrinogen concentrations observed in the South African and other black populations. These variants were also investigated in terms of their co-regulatory transcriptional properties, alongside the polygenic effect of carrying several candidate risk alleles concurrently through the investigation of transcription factor binding sites and genetic risk score models.

To this end, this dissertation presents a focused investigation on selected fibrinogen SNPs and their haplotypes in terms of their independent (first approach) and IL-6-interactive (second approach) effects on the fibrinogen phenotype (Chapter 3). In addition, a candidate gene analysis investigating the effects of polymorphisms from several genes related to proteins known to be associated with the fibrinogen phenotype is presented (third approach, Chapter 4).

The fibrinogen phenotype investigated in this study consists of three components. Both total and γ' fibrinogen concentrations were investigated, as γ' fibrinogen is independently physiologically relevant in terms of clot properties and CVD risk. Furthermore, as one of the main mechanistic pathways through which increased total and γ' fibrinogen contribute to CVD risk is through the effects thereof on plasma clot properties (Ariëns, 2013), measurable features of clot formation (lag time and slope), structure (maximum absorbance) and lysis (clot lysis time), obtained from turbidimetry, were included as phenotype outcomes to establish the functional effects of changes in total and γ' fibrinogen concentrations. Investigating the molecular mechanisms that underlie the susceptibility to hypercoagulability in black South Africans will be insightful in the development of individualised prevention frameworks that could lower the burden of CVD-related pathology in Africa.

1.2 AIMS AND OBJECTIVES

This study is a genotypic exploration of the fibrinogen phenotype in terms of protein concentration (both total and γ' fibrinogen) and functionality (as measured through turbidimetric analysis) investigated using cross-sectional data from the South African PURE study population.

The primary objectives of this study were:

1. To investigate the association of specific fibrinogen polymorphisms and their haplotypes with the fibrinogen phenotypes;

2. To determine the IL-6-interactive effect of polymorphisms and haplotypes within the fibrinogen gene cluster on total and γ' fibrinogen concentration and clot properties;

3. To identify SNPs beyond those in the fibrinogen genes that are associated with the fibrinogen phenotype based on the principles of pleiotropic and polygenic regulation.

1.3 STRUCTURE OF THE DISSERTATION

This dissertation is presented in the form of five chapters, two of which are original articles. Chapters 1, 2 and 5 meet the technical, language and referencing requirements stipulated by the North-West University, whereas Chapters 3 and 4 were written in accordance with the respective journals' specifications. A literature review, Chapter 2, follows this introductory chapter. In the literature review the biochemical structure and physiological functions of fibrinogen, together with the pathophysiological consequences of increased fibrinogen, are discussed. In addition, the genetics of fibrinogen, including the molecular composition, transcriptional regulation, heterogeneity, polymorphic variance and heritability are reviewed;

elucidating the problem of fibrinogen's missing heritability. Lastly, three approaches that could address the missing heritability are discussed, including the investigation of fibrinogen polymorphisms and their haplotypes, gene-environment interactions, specifically with regard to IL-6, and the pleiotropic and polygenic co-regulation of the fibrinogen phenotype.

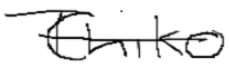
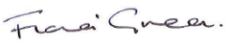
Chapter 3 is the first original research article to be submitted for publication in the British Journal of Haematology titled: "Independent and IL-6-interactive associations of selected fibrinogen polymorphisms in predicting fibrinogen and clot-related phenotypes". Thereafter, the second article, published in Matrix Biology, titled: "Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation" (Cronjé *et al., 2016,* e-pub ahead of print), forms Chapter 4. A concluding chapter, Chapter 5, captures the main findings and the implications of the results in the greater body of literature, as well as limitations of the current study and recommendations for future research. A bibliography, including references cited in the first, second and fifth chapters and addenda, concludes this dissertation.

1.4 RESEARCH TEAM

The primary research team consisted of Ms H. Toinét Cronje (M.Sc. candidate), Dr Cornelie Nienaber-Rousseau (supervisor), Prof. Marlien Pieters and Dr Zelda de Lange (co-supervisors). In addition, Mr Tinashe Chikowore, Dr Tertia van Zyl, Dr Lizelle Zandberg and Dr Fiona R. Green co-authored one or more of the manuscripts resulting from this dissertation. By signing below, each contributor accepted their indicated involvement as true, and approved the inclusion of the resultant manuscripts in this dissertation.

Table 1.1 Individual contributions of the research team members to this dissertation

Ms H. Toinét Cronjé (M.Sc. candidate and author of Chapters 3 and 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Wrote the study protocol and application for ethical approval; Performed the literature search, critically appraised the literature and compiled the literature review; Isolated the DNA (second round); Performed the KASP analyses; Prepared the files for statistical analyses in PLINK; Performed the statistical analyses using PLINK, Statistica®, and SPSS®; Interpreted the data reported in Chapters 3 and 4; Wrote and approved the final manuscripts; Critically planned and wrote this dissertation.	
Dr Cornelie Nienaber-Rousseau (Supervisor and co-author of Chapters 3 and 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Conceptualised the M.Sc. project with Prof. M. Pieters; Was co-responsible for the study protocol and application for ethical approval; Conceptualised Chapter 4; Isolated the DNA (first round) and performed the RFLP analyses; Assisted with statistical analyses using PLINK; Assisted with data interpretation and guided the writing process; Critically reviewed and approved the final manuscripts; Critically reviewed Chapters 1, 2 and 5, and the final dissertation.	
Prof. Marlien Pieters (Co-supervisor and co-author of Chapters 3 and 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Conceptualised the M.Sc. project with Dr C. Nienaber-Rousseau; Critically reviewed the study protocol; Conceptualised Chapter 3 with Dr Fiona R. Green; Supervised the laboratory analyses of all the haemostatic variables; Assisted in statistical analyses using SPSS® and Statistica®; Assisted with data interpretation and guided the writing process; Critically reviewed and approved the final manuscripts; Critically reviewed Chapters 1, 2 and 5, and the final dissertation.	
Dr Zelda de Lange (Co-supervisor and co-author of Chapter 3 and 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Critically reviewed the study protocol; Performed the global fibrinolytic assay; Critically reviewed the interpretation of results regarding clot properties; Critically reviewed and approved the final manuscripts; Critically reviewed Chapters 1, 2 and 5, and the final dissertation.	

Dr Lizelle Zandberg (Co-author of Chapters 3 and 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Performed gene functionality analyses; Interpreted the gene functionality results; Reviewed and approved the final manuscripts.	
Dr Tertia van Zyl (Co-author of Chapter 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Managed quality control of the BeadXpress® data; Critically reviewed the interpretation of results regarding lipid mediators; Reviewed and approved the final manuscript.	
Mr Tinashe Chikowore (Co-author of Chapter 4)	
Centre of Excellence for Nutrition at the North-West University (Potchefstroom Campus)	
Assisted with statistical analyses using PLINK and SPSS; Reviewed and approved the final manuscript.	
Dr Fiona R. Green (Co-author of Chapter 3)	
Division of Cardiovascular Sciences (University of Manchester)	
Conceptualised Chapter 3 with Prof. M. Pieters; Critically reviewed the interpretation of results; Critically reviewed and approved the final manuscript.	

CHAPTER 2

LITERATURE REVIEW

LITERATURE REVIEW

2.1 INTRODUCTION

Cardiovascular disease (CVD) is one of the leading causes of mortality worldwide (Mozaffarian *et al.*, 2016), and is at the moment, the second largest cause of death in South Africa, steadily increasing from 16.4% in 2012 to 17.3% in 2014 (Statistics South Africa, 2015). An increased fibrinogen concentration is an established risk factor for cardiovascular pathology, and is consistently associated with a greater risk of myocardial infarction (MI), atherosclerosis, deep vein thrombosis (DVT), stroke and coronary heart disease (CHD), owing to its involvement in thrombotic and inflammatory processes (Danesh *et al.*, 1998; Danesh *et al.*, 2005; Kaptoge *et al.*, 2007; Kaptoge *et al.*, 2012; Kaptoge *et al.*, 2013). It was recently reported that black South Africans had a greater longitudinal worsening of CVD risk compared to whites, particularly with regard to unfavourable haemostatic profiles (Hamer *et al.*, 2015). In addition, local and international epidemiological research has shown that black ethnicities have higher fibrinogen concentrations than their white counterparts (Folsom *et al.*, 1991; Cook *et al.*, 2001; Greyling *et al.*, 2007; Pieters & Vorster, 2008; Albert *et al.*, 2009; Lammertyn *et al.*, 2015).

Fibrinogen concentrations correlate strongly between family members, specifically twins, suggesting that the fibrinogen phenotype has a large heritable component. Heritability estimates ranged from 20 to 51% in various pedigree structures (Hamsten *et al.*, 1987; Friedlander *et al.*, 1995; de Lange *et al.*, 2001). Environmental factors such as body mass index (BMI), tobacco and alcohol use, lipid profiles and physical activity contribute to the remaining variance (Kaptoge *et al.*, 2007; Arbustini *et al.*, 2013). Although heritability estimates have been significant, genome-wide association studies (GWAS) have allocated a mere 3.7% of this heritability to common single nucleotide polymorphisms (SNPs) to date (Sabater-Lleal *et al.*, 2013; de Vries *et al.*, 2016). Furthermore, genetic association studies in Europeans have had some difficulty in trying to identify functional SNPs in the fibrinogen gene cluster, owing to great linkage disequilibrium (LD), particularly in the upstream β chain gene region where highly associated SNPs are in full LD (Green, 2001; Baumann & Henschen, 1994; Behague *et al.*, 1996; Mannila *et al.*, 2005; Verschuur *et al.*, 2005). The black South African population, in contrast, have great genetic variability, as observed in their complex LD and haplotype pattern (Teo *et al.*, 2010; Kotzé *et al.*, 2015), which can be utilised in an effort to identify possible functional SNPs.

Acknowledging that twin studies have allocated up to 51% of fibrinogen concentrations to be heritable, and GWAS have only been able to allocate 3.7% thereof, a case is made to

investigate the missing heritability of the fibrinogen phenotype. This dissertation aims to do just that, using three primary approaches to explore the genotypic composition of the fibrinogen phenotype in Africans. Firstly, investigating polymorphisms and haplotypes within the fibrinogen genes themselves in a population known for genetic diversity and low LD could assist in the identification of functional variants. Secondly, gene-environment interactions, particularly regarding the effect of cytokines such as IL-6, have not been investigated in an African context where the prevalence of low-grade chronic inflammation is significant (Pieters *et al.*, 2011; Lammertyn *et al.*, 2015). These interactions could explain a larger part of the variance than the contribution of polymorphisms and environmental influencers separately. Lastly, an investigation of candidate genes and polymorphisms coding for phenotypes biologically relevant to fibrinogen, overcomes the methodological barriers in terms of a genome-wide investigation in Africans (Chapter 1), and utilises the genetic and phenotypic diversity of the African study population to identify novel associations. Furthermore, investigating these candidate SNPs in terms of their pleiotropic and polygenic co-regulatory properties could contribute to the broader understanding of the genetic regulation of the fibrinogen phenotype.

This literature review appraises the theoretical principles that underlie these three approaches by reviewing the basic principles of linkage, LD and haplotypes, gene-environment interactions and cytokine involvement, as well as pleiotropic and polygenic regulation of complex phenotypes. Prior to the overview of these three approaches, the fibrinogen phenotype will be reviewed in terms of its biochemical composition, (patho)physiological relevance, genetic regulation and heterogeneity, as well as heritability. The research question in terms of missing heritability will then be presented as an introduction to the three approaches.

2.2 FIBRINOGEN: BIOCHEMISTRY, PHYSIOLOGY AND PATHOPHYSIOLOGY

2.2.1 Biochemical composition of the fibrinogen protein

Fibrinogen is a 45 nm soluble glycoprotein with a molecular weight (MW) of 340 kilo Dalton (kDa), synthesised and assembled mainly by the hepatic parenchymal cells (Caspary & Kekwick, 1957). Secretion of fibrinogen by the intestinal and alveolar epithelial cells upon stimulation by inflammatory cytokines has also been described (Guadiz *et al.*, 1997; Lawrence & Simpson-Haidaris, 2004).

The hexameric fibrinogen protein is composed of two sets of non-identical disulphide-bridged alpha (α), beta (β) and gamma (γ) chains (Blombäck *et al.*, 1976; Henschen *et al.*, 1983; Zhang & Redman, 1992; Huang *et al.*, 1993; Herrick *et al.*, 1999). Fibrinogen has a trinodular structure that consists of firstly a central E-domain containing the amino-termini of the six polypeptide

chains and secondly, two distal D-regions (Blombäck, 1996) containing binding pockets crucial for fibrinogen polymerisation. These three structural components are joined by two α-helical coiled coils made up of three polypeptide chains (AαBβγ), which provide elasticity to the fibrinogen molecule (Henschen *et al.*, 1983). The amino-termini of the α and β chains represent the fibrinopeptides A (FPA) and B (FPB) and are situated in the E-region as the cleavage site for thrombin (Yang *et al.*, 2000). The carboxyl termini of the β and γ chains terminate at the respective D-regions, whereas the α chains' carboxyl termini extend to flexible αC domains of more than 350 residues protruding from the D-regions (Mosesson, 1998; Tsurupa *et al.*, 2009). Fibrinogen is depicted in Figure 2.1.

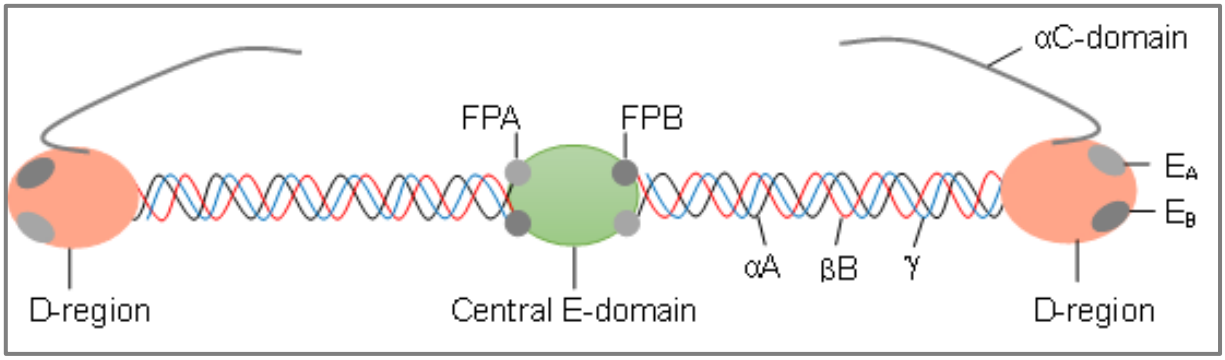


Figure 2.1 Biochemical composition of the fibrinogen protein

Adapted from Herrick et al. (1999) and Mosesson (2005)

Common variants of the fibrinogen protein exist, including high, low, and very low MW variants (3.40, 3.05 and 2.70 kDa respectively), depending on the presence of both, one or none of the α chain carboxyl termini (Holm & Godal, 1984; Holm *et al.*, 1985; Holm *et al.*, 1986). Variation in the β and γ chains has also been reported (Mosesson *et al.*, 1972; Brennan *et al.*, 2009). These heterogeneities will be discussed in Section 2.3.4.

2.2.2 The (patho)physiology of fibrinogen

Fibrinogen is present in the plasma at a basal concentration of 1.5 to 4.5 g/L, and varies as a result of numerous genetic and environmental influences (Kamath & Lip, 2003; Weinstock & Ntefidou, 2006). As an acute phase protein, fibrinogen concentrations increase significantly during times of physiological stress, and have a biological half-life of approximately 100 hours (Herrick *et al.*, 1999; Pulanić & Rudan, 2005).

Fibrinogen, both in shortage and excess, has pathophysiological consequences. In shortage, a lack of blood coagulation causes increased haemorrhaging and excessive blood loss, whereas excess fibrinogen concentrations result in hypercoagulability and delayed fibrinolysis (Ariëns,

2013). This literature review will focus on pathophysiology resulting from excess fibrinogen, as it is the most prominent occurrence in the South African population and a known contributor to CVD (Lovely *et al.*, 2002; Mannila *et al.*, 2007; Bertram *et al.*, 2013; Hamer *et al.*, 2015).

In this section the main physiological functions of fibrinogen, haemostasis and inflammation, as well as the pathophysiological consequences related to these functions in the presence of high fibrinogen concentrations, are discussed. The section concludes with an overview of the literature regarding the prospective association of increased fibrinogen concentration with CVD risk.

2.2.2.1 The (patho)physiology of fibrinogen: haemostasis and thrombogenesis

Haemostasis refers to the maintenance of day-to-day blood flow and the physiological ability to respond to excessive bleeding by sealing damaged vessel walls to cease haemorrhaging (Widmaier *et al.*, 2011). The coagulation cascade (Figure 2.2) is composed of two pathways, intrinsic and extrinsic, that differ owing to the cause of coagulation initiation (Davie *et al.*, 1991). Both pathways consist of a range of stepwise enzymatic conversions of inactive plasma proteins (zymogens) to, mostly, serine proteases (Davie & Ratnoff, 1964). The intrinsic pathway initiates when Factor (F)XII is activated through contact activation by collagen or connective tissue (Wilner *et al.*, 1968) to form FXIIa. The extrinsic pathway initiates when the membrane protein, tissue factor (TF), which has a strong affinity for FVII, forms an enzymatically active TF-FVIIa complex in the presence of calcium ions (Jesty & Nemerson, 1974). The two pathways converge at the formation of active serine protease FXa. FX is activated through stepwise reactions within the intrinsic pathway, resulting in FIXa, which, in the presence of FVIII, phospholipids and calcium, is able to initiate the cleavage of FX (Davie *et al.*, 1991). Extrinsically, the FX-to-FXa conversion occurs through either the direct 'attack' of the TF-FVIIa complex on the FX protein (Davie & Ratnoff, 1964; Nemerson, 1966; Jesty & Nemerson, 1974), or indirectly through the activation of FXI (Osterud & Rapaport, 1977; Marlar *et al.*, 1982). The presence of phospholipids and calcium ions assists in the formation of a FXa-FVa complex, also known as "prothrombinase" (Rosing *et al.*, 1980). This converts prothrombin to thrombin through hydrolysis. Factors Va and FVIIIa participate in the cascade as cofactors (Davie *et al.*, 1991).

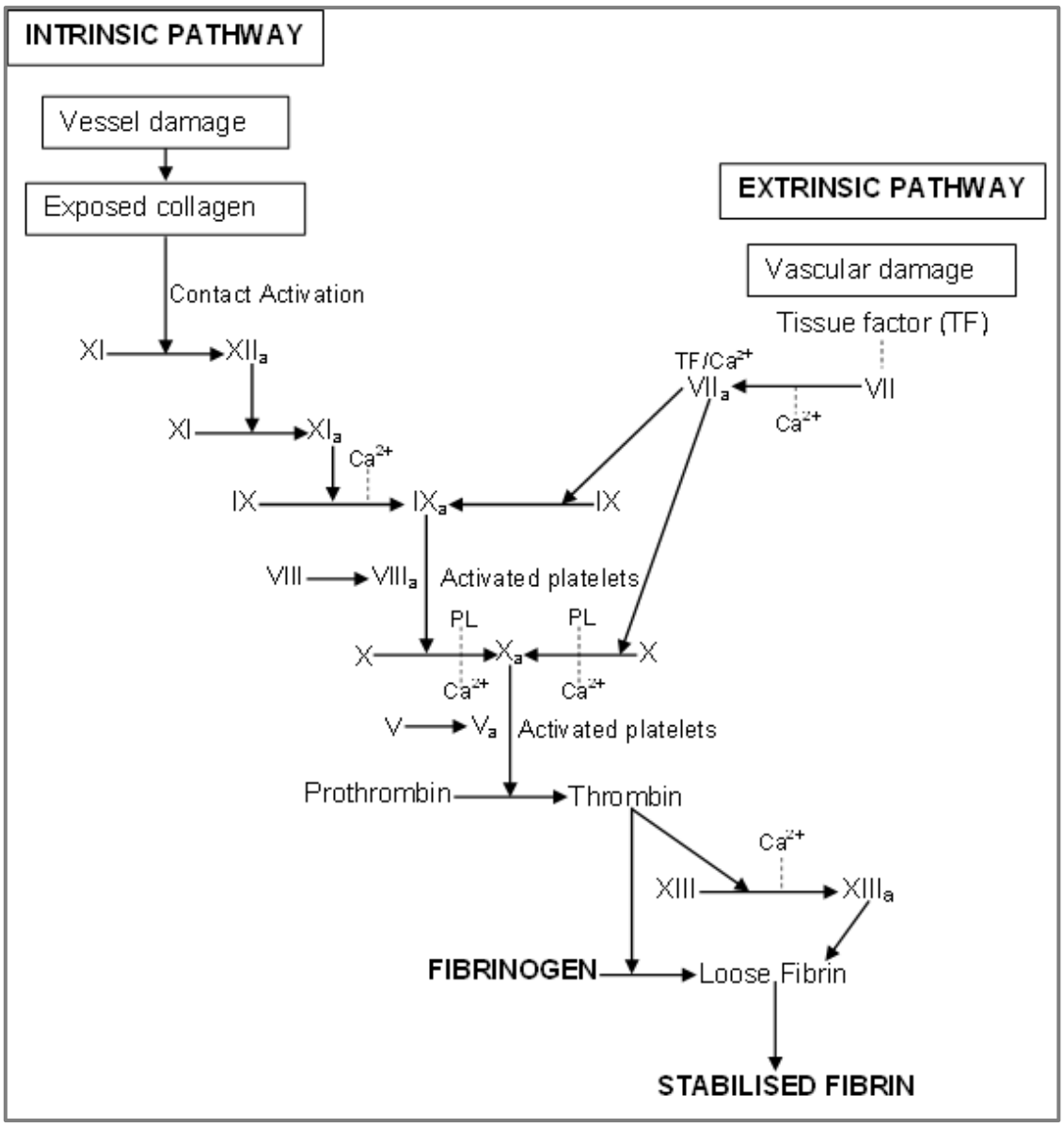


Figure 2.2 Blood coagulation cascade

Adapted from Davie et al. (1991)and Widmaier et al. (2011)

Upon the generation of thrombin, fibrinogen is converted to fibrin, the greatest constituent of a blood clot. This conversion takes place when the proteolytic cleavage of the peptide bonds in the amino-termini of both the α and β chains releases FPA and FPB (Scheraga & Laskowski, 1957). Cleavage of FPA exposes the E_A polymerisation site at the amino-terminal of the α chain (17^{th} to 20^{th} amino acid) and the β chain (between the 15^{th} and 42^{nd} amino acid), respectively (Yang *et al.*, 2000; Mosesson *et al.*, 2001). The consequent cleavage of FPB exposes an alternative polymerisation site at the E-terminal (E_B) (Sporn *et al.*, 1995). Complementary binding pockets in the γ chains situated in the D-domains (D_A and D_B) bind to the E_A and E_B polymerisation sites to stagger fibrin molecules in an end-to-middle attachment manner, so that

the central E-domain binds to the outer D-domain of the adjacent molecule (Smith, 1980; Mosesson *et al.*, 2001). The αC regions are also released by FPB cleavage, thereby allowing the lateral aggregation of the protofibrils (Tsurupa *et al.*, 2009; Riedel *et al.*, 2011).

During fibrin formation, thrombin also facilitates the conversion of FXIII to an active FXIIIa enzyme that enables the covalent cross-linking of the fibrin monomers by covalent bonds in the α and γ chains (Ariëns *et al.*, 2002). Cross-linking by FXIII and the incorporation of additional plasma proteins (fibrinectin and α-antiplasmin) assist in the stabilisation and strengthening of a secure fibrin clot (Davie *et al.*, 1991; Ariëns *et al.*, 2002).

In addition to the role of fibrinogen in clot formation, fibrinogen chronically regulates blood viscosity and promotes wound healing through vasoconstriction at injury sites, and enhanced wound stability and cell-to-cell interaction and adhesion (Herrick *et al.*, 1999; Drew *et al.*, 2001; Reinhart, 2003; Pulanić & Rudan, 2005). Furthermore, fibrinogen acts as an adhesive by enhancing platelet aggregation through its binding to the platelet glycoprotein IIIb/IIIa receptor, thereby producing a platelet-rich blood clot (Cahill *et al.*, 1992; Calvete, 1995; Lefkovits *et al.*, 1995; Koenig, 2003).

Although the formation of a blood clot is essential, all blood clots that have formed have to be lysed upon adequate tissue repair and blood loss control. The fibrinolytic system is initiated when plasminogen, activated by tissue plasminogen activator (tPA) and urokinase plasminogen activator (uPA), forms plasmin (Davie *et al.*, 1991; Collen, 1999). Plasmin is able to lyse the fibrin clot by cleaving individual fibrin fibres at C-terminal lysines and releasing soluble fibrin degradation products into the vasculature (Collen, 1999). Hyperfibrinogenaemia interferes with the ability of clots to lyse efficiently, and disturbs the balance between haemostatic and fibrinolytic function in favour of thrombogenesis, consequently leading to a hypercoagulable state (McDonagh & Lee, 1997; Machlus *et al.*, 2011). Hyperfibrinogenaemia is associated with a larger platelet and fibrin rich clot structure, with thin fibres packed densely in a rigid clot network structure with lower porosity (Fatah *et al.*, 1992; Scrutton *et al.*, 1994; Mills *et al.*, 2002; Collet *et al.*, 2006; Undas & Ariëns, 2011). These clots remain in the vasculature for a longer time owing to the inability of degradation enzymes to enter the clot, thereby impairing clot lysis (Machlus *et al.*, 2011). Hyperfibrinogenaemia also increases blood viscosity and aggregation of erythrocytes, further enhancing the risk of thrombosis (Koenig & Ernst, 1992; Lowe, 1992).

Apart from fibrinogen's role in blood coagulation and thrombosis, fibrinogen is also an acute phase protein, involved in the inflammatory process, and in excess, contributes to

atherogenesis and inflammatory disease. These (patho)physiological properties of fibrinogen are discussed in the following section.

2.2.2.2 The (patho)physiology of fibrinogen: inflammation and atherogenesis

Fibrinogen's involvement as both a haemostatic and an acute phase reactant in the coagulation and inflammatory systems is essential, as are both of these biological pathways (Pulanić & Rudan, 2005). The interplay and cross-talk between these two systems is complex and the initiation or up-regulation of one is mirrored in the other (Dahlbäck, 2012). Coagulation and inflammation are kept under strict homeostatic control by anti-coagulant and anti-inflammatory mechanisms and the chronic disturbance thereof has pathological consequences (Dahlbäck, 2012; Davalos & Akassoglou, 2012).

IL-6 and glucocorticoids primarily up-regulate the hepatic expression of fibrinogen during the acute phase initiated by pro-inflammatory cytokines such as interleukin-1β (IL-1β) and tumour necrosis factor-α (TNF-α) (Herrick *et al.*, 1999; Redman & Xia, 2000). The increase in the expression of fibrinogen is observed upon physiological stress such as smoking, strenuous exercise or surgery (Koster *et al.*, 1994; Folsom, 2001; Danesh *et al.*, 2005; Smith *et al.*, 2005; Folsom *et al.*, 2007; Kaptoge *et al.*, 2007; Spiel *et al.*, 2008; Davalos & Akassoglou, 2012). In addition to hepatic up-regulation, fibrinogen synthesis is also stimulated in the lung and endothelial epithelium (Guadiz *et al.*, 1997; Lawrence & Simpson-Haidaris, 2004).

The inflammatory response relies on the interaction of leukocytes and their surface receptors (integrins) with inflammatory proteins such as fibrinogen. The main integrins through which fibrinogen exerts its inflammatory effect are $CD11_B/CD_{18}$ (Macrophage-1 antigen, Mac-1) and $CD11_C/CD_{18}$ ($\alpha_X\beta_2$). Fibrinogen is able to bind to Mac-1 as well as the intercellular adhesion molecule-1 (ICAM-1), through which monocyte-endothelial cell interaction is enhanced owing to the bridging of monocytes (Mac-1) to endothelial cells (ICAM-1) (Languino *et al.*, 1995; Van de Stolpe *et al.*, 1996; Duperray *et al.*, 1997). Fibrinogen also has the ability to up-regulate the expression of ICAM-1, specifically through the release of FPB, resulting in greater monocyte adhesion and cellular proliferation (Gardiner & D'Souza, 1997; Harley *et al.*, 2000; Tsakadze *et al.*, 2002). The migration and binding of leukocytes to the endothelial tissue through ICAM-1 (Languino *et al.*, 1995) enhance subsequent chemotaxis, resulting in the influx of monocytes, lymphocytes, eosinophils, fibroblasts and endothelial cells to the site of injury (Forsyth *et al.*, 2001). Rubel *et al.* (2001) also observed that the binding of fibrinogen to integrins activated neutrophils and resulted in a delay in phagocytosis. Binding of fibrinogen to the endothelial cells through ICAM-1 also stimulates the release of vasoactive mediators, that can rapidly induce vasodilation or constriction (Hicks *et al.*, 1996; Herrick *et al.*, 1999).

Pathogenically, endothelial attachment modulates the permeability of the endothelial tissue that allows the deposition of fibrin(ogen) remnants in the sub-endothelium, thereby providing an adsorptive surface for LDL-c and apolipoprotein-A (Lou *et al.*, 1998; Retzinger *et al.*, 1998). Fibrin(ogen) deposits have been found in atherosclerotic plaque (Bini *et al.*, 1989; Reinhart, 2003; Borissoff *et al.*, 2011), where it contributes to risk of atherosclerosis (Naito *et al.*, 1990; Lou *et al.*, 1998). Small arterial lesions that occur within the plaque allows fibrin(ogen) deposits, leading to the incorporation of fibrin(ogen) in the growing plaque mass, thereby contributing to chemotaxis of smooth muscle cells and greater instability of atherosclerotic plaque (Collet *et al.*, 2000; Feinbloom & Bauer, 2005; Ariëns, 2013). These mechanisms support the causal role of fibrinogen in atherosclerosis, although the lack of an effective isolated fibrinogen-lowering pharmacological product has made the investigation and establishment of fibrinogen as either a marker or mediator of atherosclerosis problematic (Reinhart, 2003). Opposing the above-mentioned pathophysiological consequences, fibrinogen also possesses antioxidant properties that are able to minimise oxidative damage during inflammation, offering a protective mechanism during the inflammatory process (Kaplan *et al.*, 2001; Olinescu & Kummerow, 2001).

The involvement of hyperfibrinogenaemia in thrombo- and atherogenesis provides mechanisms that support the increased risk of CVD observed in individuals with high fibrinogen concentrations. A brief summary of the prospective data on fibrinogen and the broad spectrum of cardiovascular pathology will now be discussed.

2.2.2.3 Fibrinogen and cardiovascular disease risk

Increased fibrinogen concentrations are prospectively associated with an increased risk of atherosclerosis (Chambless *et al.*, 2002), MI (Ernst & Resch, 1993; Danesh *et al.*, 2005; Mannila *et al.*, 2005), stroke (Wilhelmsen *et al.*, 1984; Kannel *et al.*, 1987; Qizilbash *et al.*, 1991; Lee *et al.*, 1993; Maresca *et al.*, 1999), DVT (Koster *et al.*, 1994; Uitte de Willige *et al.*, 2005) and CHD (Meade *et al.*, 1986; Palmieri *et al.*, 2003; Danesh *et al.*, 2005). Apart from fibrinogen's association with these CVDs, fibrinogen is known to contribute to autoimmune and inflammatory diseases, including inflammatory bowel disease and cancer, through its pro-inflammatory properties (Davalos & Akassoglou, 2012). The causality of fibrinogen concentrations in these above-mentioned illnesses remains a highly debated topic (Reinhart, 2003; Sabater-Lleal *et al.*, 2013). Fibrinogen has been reported as a marker of the increased inflammatory state that contributes to CVD risk, more so than as an independent CVD risk contributor (Sabater-Lleal *et al.*, 2013). Several mechanistic pathways do however exist through which an increased fibrinogen concentration can contribute to CVD. This includes the

formation of pro-atherogenic clot structure, increased viscosity, increased platelet activation and the above-mentioned inflammatory pathways (Machlus *et al.*, 2011; Sabater-Lleal *et al.*, 2013).

The study of causality is challenging, as fibrinogen is greatly associated with numerous independent CVD risk factors and intertwined with the inflammatory process, making the investigation of fibrinogen as a truly independent contributor difficult (Keavney *et al.*, 2006). Furthermore, effective fibrinogen-lowering medication is not available, inhibiting randomised control trials to isolate the true contribution of fibrinogen to CVD (Lowe & Rumley, 2001). Regardless of the controversy, the Emerging Risk Factors Collaboration (2012) reported that the inclusion of fibrinogen concentrations in the conventional list of risk factors could result in the prevention of one cardiovascular event in every 400 to 500 screened individuals in ten years.

In the South African context, higher fibrinogen concentrations have been reported in black population groups (Greyling *et al.*, 2007; Pieters & Vorster, 2008; Lammertyn *et al.*, 2015), in agreement with global literature (Folsom *et al.*, 1991; Green *et al.*, 1994; Lutsey *et al.*, 2006; Kaptoge *et al.*, 2007; Albert *et al.*, 2009; Wassel *et al.*, 2011). Research into the risk of CVD and fibrinogen concentrations in South Africa is limited, with a weak, but significant, association reported by Pieters *et al.* (2011). A predisposition to hypercoagulability has been noted in black South Africans (Hamer *et al.*, 2015) compared to their white counterparts, which greatly enhances their prospective risk of CVD outcomes.

A large portion of increased fibrinogen concentration and the consequent increased risk of pathogenic clot kinetics and structure have been reported to be heritable (de Maat, 2001; Mills *et al.*, 2002). The following section will review the genetics of fibrinogen.

2.3 GENETICS OF FIBRINOGEN

Fibrinogen concentration is to a large extent genetically pre-determined. This section will focus on the molecular composition and expression of the three fibrinogen genes. The well-known splice variants and polymorphic variation within these genes will also be discussed, after which the reported heritability will be reviewed and the case of fibrinogen's missing heritability will be presented.

2.3.1 The molecular composition of fibrinogen

The α, β and γ chain sets of the fibrinogen molecule are coded for individually by the respective α, β and γ chain genes (*FGA*, *FGB* and *FGG*), situated in a 50 kb region on the q-arm of

chromosome 4 (Henry *et al.*, 1984). The β chain transcript opposes that of the α and γ chain orientation, and the α chain gene encodes the largest of the three transcripts. The composition of the fibrinogen gene cluster is briefly summarised in Figure 2.3.

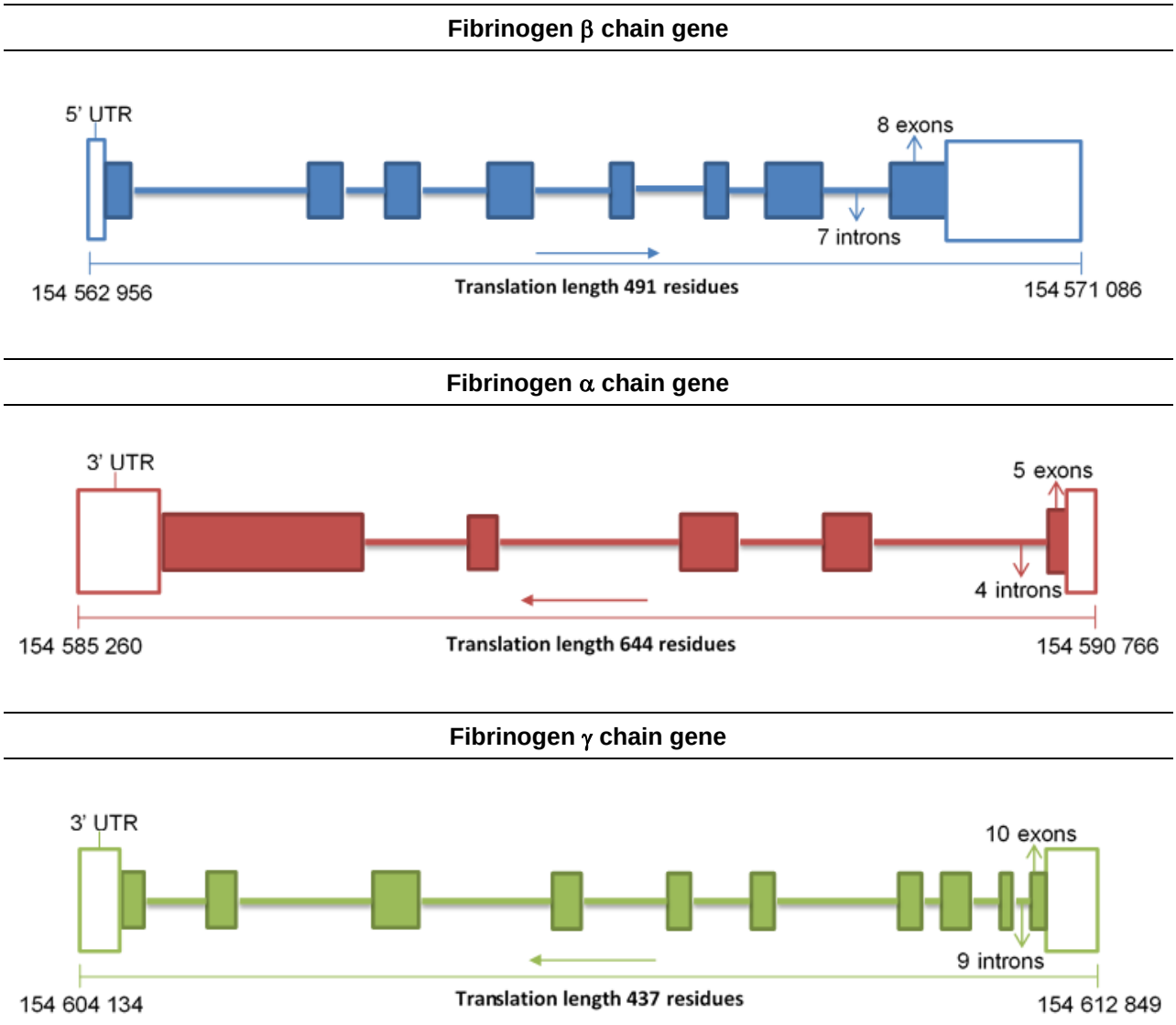


Figure 2.3 Composition of the fibrinogen gene cluster

Adapted from Ensembl, Release 86 (Yates et al., 2016)

Transcription and translation of the genes depicted above form the functional fibrinogen protein. The following section will discuss the transcriptional regulation of the fibrinogen gene cluster, and how two distinct mechanisms of transcriptional control enable fibrinogen to both chronically regulate blood flow and acutely respond to tissue damage and excessive bleeding.

2.3.2 Transcriptional regulation of fibrinogen

As summarised in the previous section, the fibrinogen protein is composed of three polypeptide chains coded for by three individual genes. These genes are independently transcribed and the co-regulatory nature of their expression has been greatly investigated (Fish & Neerman-Arbez, 2012). As all three chains are required to form the functional protein, theoretically, regulatory change in any of the chains would reflect on the entire protein. This has been proven in *in vitro* research where the induction of the expression of one of the polypeptide chains resulted in the up-regulation of the entire protein (Otto *et al.*, 1987; Roy *et al.*, 1990; Roy *et al.*, 1994). In addition, a circadian rhythm to hepatic fibrinogen expression has been observed, with a peak in fibrinogen concentrations and expression observed in the early morning, in the presence of light stimuli (Bremner *et al.*, 2000; Sakao *et al.*, 2003).

The β chain gene has a transcriptional orientation opposing that of the α and γ chain genes that are in tandem, and has previously been indicated to be rate-limiting in the transcriptional process (Yu *et al.*, 1983; Yu *et al.*, 1984; Roy *et al.*, 1990; Herrick *et al.*, 1999). Fibrinogen is expressed primarily in hepatocytes, in which two distinct forms of transcriptional control have been observed: Firstly, basal fibrinogen expression that controls day-to-day blood flow and viscosity, and secondly acute phase response (APR) transcriptional regulation, that regulates fibrinogen expression during physiological trauma in order to replenish fibrinogen lost either by excessive bleeding or vessel and wound repair (Ritchie & Fuller, 1983; Fuller *et al.*, 1985; Gabay & Kushner, 1999; Fish & Neerman-Arbez, 2012).

The basal expression of fibrinogen relies on three primary promoter cis-acting transcription factors, hepatic nuclear factor 1 (HNF1), CCAAT enhancer-binding protein (C/EBP) and an upstream stimulatory factor (USF) (Figure 2.2). HNF1, positively regulated by hepatocyte nuclear factor-4 α, and C/EBP are involved in the basal transcription of the α and β chain genes (Courtois *et al.*, 1987; Dalmon *et al.*, 1993; Hu *et al.*, 1995) and are located -59 to -47 and -132 to -124 (*FGB*) (Dalmon *et al.*, 1993) and -91 to -79 and -142 to -132 (*FGA*) (Hu *et al.*, 1995) base pairs (bps) from the transcription start site (TSS). Transcription of the γ chain is dependent on the USF located -77 to -66 bps from the TSS (Mizuguchi *et al.*, 1995).

APR transcriptional regulation relies on the elevated secretion of IL-6 by macrophages that follows in response to pro-inflammatory cytokines such as TNF-α and IL-1β (Fish & Neerman-Arbez, 2012). Increased IL-6 stimulates the hepatic production of acute phase proteins, such as C-reactive protein, heptoglobulin and fibrinogen (Redman & Xia, 2000). The up-regulation of fibrinogen expression during the APR is under the transcriptional control of the IL-6 and

glucocorticoid signalling pathways. IL-6 responsive elements (IL-6REs) have been observed upstream of all three of the fibrinogen genes and are functional through their interaction with the signal transducer and activator of transcription-3 (STAT-3) transcription factor (Fish & Neerman-Arbez, 2012). Variants within the *STAT-3* gene have been identified in a recent GWAS to affect fibrinogen concentration, reiterating the role of STAT-3 in fibrinogen regulation (de Vries *et al.*, 2016). The IL-6REs are located between -150 and -82 bps upstream of the *FGB* TSS (Anderson *et al.*, 1993; Dalmon *et al.*, 1993), between -127 and -122 bps upstream of the *FGA* TSS (Hu *et al.*, 1995), and between -306 and -301 upstream of the *FGG* TSS (Mizuguchi *et al.*, 1995). The IL-6-RE at the *FGA* and *FGB* promoter is only functional when the nearby C/EBP site (*FGA*) (Hu *et al.*, 1995) or HNF1 (*FGB*) (Dalmon *et al.*, 1993) site is intact and functional. The IL-6RE at the γ chain, however, is independent from the C/EBP site's activity (Mizuguchi *et al.*, 1995).

In addition, glucocorticoid responsive enhancer elements have been identified in the β and γ chain promoter regions between -2900 and -1500, and -1116 to -1102 upstream of the TSS, respectively (Anderson *et al.*, 1993; Dalmon *et al.*, 1993; Asselta *et al.*, 1998). Glucocorticoids also influence STAT-3 by down-regulating the expression of suppressor of cytokine signalling 3, thereby relieving the inhibitory effect thereof on STAT-3 phosphorylation (Dittrich *et al.*, 2012).

Repression of the response of the fibrinogen genes to IL-6 has been reported for TNF-α and interleukin-1β that are upstream of IL-6 in the acute phase response. Phosphorylation of the STAT-3 transcription factor is prolonged through IL-1β in a nuclear factor kappa β dependent mechanism, inhibiting the APR response (Albrecht *et al.*, 2007).

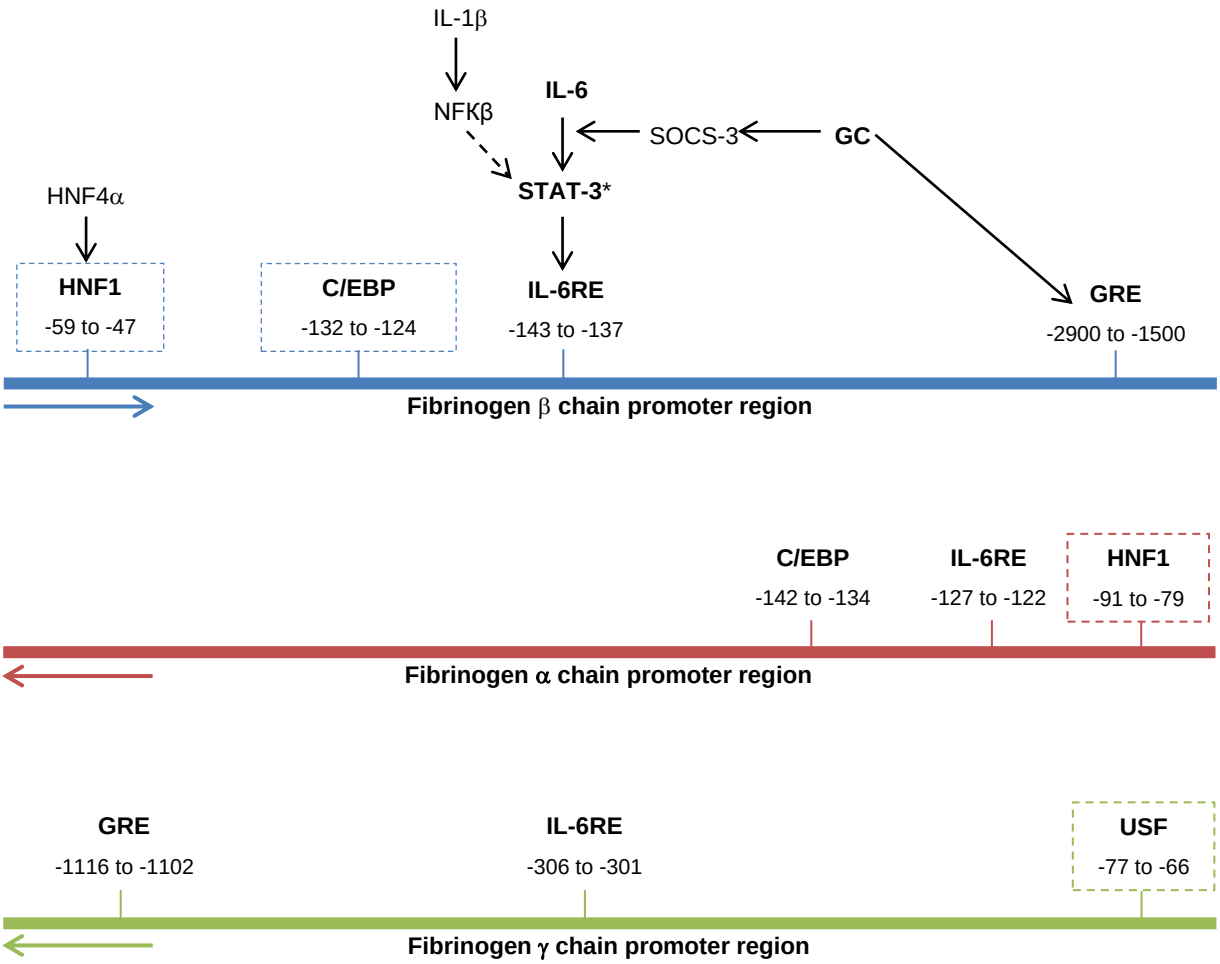


Figure 2.4 Transcriptional regulation of the fibrinogen gene cluster

Adapted from Redman and Xia (2000) and Fish and Neerman-Arbez (2012)

Mechanisms depicted at the top of the diagram are applicable to the same transcription factors in the α and γ chain promoters. Dashed outline = involvement in basal transcriptional regulation; Solid arrow = positive regulation/induction; Dashed arrow = repression; * activated through phosphorylation; C/EBP = CCAAT enhancer binding protein element; GRE = Glucocorticoid responsive enhancer element; HNF1 = hepatocyte nuclear factor-1; HNF4α = hepatocyte nuclear factor-4 α, IL-6RE = interleukin-6 responsive element; SOCS-3 = suppressor of cytokine signaling-3; STAT-3 = signal transducer and activator of transcription-3; USF = upstream stimulatory factor

It has been shown that polymorphisms within the promoter regions of the fibrinogen genes can affect the IL-6-induced expression of fibrinogen by causing either greater or lesser increases in the hepatic expression of fibrinogen (Verschuur *et al.*, 2005; Morozumi *et al.*, 2009). Apart from these effects, several polymorphisms within and outside of the fibrinogen genes' promoter regions have been described to associate with the fibrinogen phenotype in terms of protein concentrations or disease risk, independent of cytokine interactions (Jacquemin *et al.*, 2008; Dehghan *et al.*, 2009). Polymorphisms that are associated with the fibrinogen phenotype and are investigated in this study will be discussed in the following section.

2.3.3 Common genetic variation within the fibrinogen gene cluster

The fibrinogen cluster is highly polymorphic, with hundreds of SNPs and insertions/deletions identified to date (Yates *et al.*, 2016). Focused investigations have concentrated on only a few SNPs, thought to be biologically significant in terms of either alterations in protein function or structure, or the rate of expression. These SNPs are investigated in terms of either fibrinogen concentrations as an intermediate phenotype, or in terms of cardiovascular outcomes and disease risk (Behague *et al.*, 1996; de Maat *et al.*, 1998; Lim *et al.*, 2003). Results have been controversial (Keavney *et al.*, 2006; Sabater-Lleal *et al.*, 2013). Several SNPs, as summarised in Table 2.2, have for instance been associated with increased fibrinogen concentrations, which are thought to be independently associated with CVD risk and cardiovascular-related pathology (Section 2.2.3). However, as reported in Mendelian randomisation studies, the same SNPs that are associated with increased fibrinogen concentrations are not necessarily associated with increased CVD risk, placing a question mark over the causative role of fibrinogen in CVD (Ken-Dror *et al.*, 2012). The causality of fibrinogen in CVD, as well as the efficacy of Mendelian randomisation to refute causality claims, remains a debated topic (Keavney *et al.*, 2006; Meade *et al.*, 2006; Sabater-Lleal *et al.*, 2013).

In Table 2.2 some of the SNPs that have been most prevalent in the context of the fibrinogen phenotype are summarised. For the purpose of this literature review, only the polymorphisms reported on further in this dissertation will be included, as the totality of literature on SNP associations is too broad for the scope of this dissertation.

Table 2.1 The association of selected fibrinogen polymorphisms with total fibrinogen concentrations and (patho)physiological outcomes

Variant name	Gene Location	Association with total fibrinogen concentrations			(Patho)physiological affects
		Increase	Decrease	None	
rs7439150	*FGB* 5' UTR	de Vries *et al.* (2016)			
rs1800789 -1420G/A	*FGB* 5' UTR	Kathiresan *et al.* (2006) Dehghan *et al.* (2009) Sabater-Lleal *et al.* (2013)			↑ CAD severity in MI patients (Behague *et al.*, 1996) ↓ CAD risk (Theodoraki *et al.*, 2010)
rs1800791 -854G/A	*FGB* 5' UTR	Behague *et al.* (1996) van't Hooft *et al.* (1999)		Kathiresan *et al.* (2006) Mannila *et al.* (2006) Carty *et al.* (2008) Jacquemin *et al.* (2008)	
rs1800790 -455G/A	*FGB* 5' UTR	Humphries *et al.* (1987) Behague *et al.* (1996) Tybjærg-Hansen *et al.* (1997) de Maat *et al.* (1998) Van Der Bom *et al.* (1998) van't Hooft *et al.* (1999) Liu *et al.* (2005) Kathiresan *et al.* (2006) Carty *et al.* (2010) Ken-Dror *et al.* (2012)		Connor *et al.* (1992) Wang *et al.* (1997)	↑ CAD severity in MI patients (Behague *et al.*, 1996) ↑ CAD risk in patients with NIDDM (Carter *et al.*, 1996) ↔ CAD risk (Gardemann *et al.*, 1997; Wang *et al.*, 1997) ↔ IHD risk (Tybjærg-Hansen *et al.*, 1997) ↑ CAD progression (de Maat *et al.*, 1998) ↑ Ischemic stroke risk (Nishiuma *et al.*, 1998) ↔ MI risk (Gardemann *et al.*, 1997; McCarthy *et al.*, 2004) ↑ CAD risk (McCarthy *et al.*, 2004)

Variant name	Gene Location	Association with total fibrinogen concentrations			(Patho)physiological affects
		Increase	Decrease	None	
rs1800788 -249C/T 1643C/T	*FGB* 5' UTR	Liu *et al.* (2005)		Mannila *et al.* (2005) Uitte de Willige *et al.* (2005) Kathiresan *et al.* (2006) Jacquemin *et al.* (2008)	↑ CVD risk (Liu *et al.*, 2005) ↔ CAD risk (Behague *et al.*, 1996) ↔ MI risk (Mannila *et al.*, 2005)
rs1800787 -148C/T	*FGB* 5' UTR	Humphries *et al.* (1987) Schadt *et al.* (2008) Wassel *et al.* (2011)		Mannila *et al.* (2005)	↓ CAD risk (Theodoraki *et al.*, 2010) ↔ MI risk (Mannila *et al.*, 2005)
rs4220	*FGB* Exon 8	Kathiresan *et al.* (2006) Danik *et al.* (2009) Ken-Dror *et al.* (2012)		Mannila *et al.* (2005)	↔ MI risk (Mannila *et al.*, 2005)
rs4463047	3' *FGB* Intergenic	Sabater-Lleal *et al.* (2013)	Wassel *et al.* (2011)	Mannila *et al.* (2005)	↔ MI risk (Mannila *et al.*, 2005)
rs6050 Thr312Ala 6534A/G	*FGA* Exon 2		Carty *et al.* (2010)	Mannila *et al.* (2005)	↔ Stroke risk (Carter *et al.*, 1999) Altered fibrin gel structure (Carter *et al.*, 1999; Lim *et al.*, 2003) ↑ thromboembolism risk (Lim *et al.*, 2003) ↔ MI risk (Mannila *et al.*, 2005)

Variant name	Gene Location	Association with total fibrinogen concentrations			(Patho)physiological affects
		Increase	Decrease	None	
rs2070011 -58G/A	*FGA* 5' UTR		Ken-Dror *et al.* (2012)	Mannila *et al.* (2005)	↔ MI risk (Mannila *et al.*, 2005)
rs2066865 10034C/T	*FGG* 3' UTR		Carty *et al.* (2008)	Jacquemin *et al.* (2008)	
rs1049636 9340T/C	*FGG* 3' UTR		Jacquemin *et al.* (2008) Danik *et al.* (2009)	Mannila *et al.* (2005)	↔ MI risk (Mannila *et al.*, 2005)

Associations reported refer to the association with the minor allele; A = adenine; C = cytosine; CAD = coronary artery disease; CVD = cardiovascular disease; *FGA* = fibrinogen alpha chain gene; *FGB* = fibrinogen beta chain gene; *FGG* = fibrinogen gamma chain gene; G = guanine; IHD = ischaemic heart disease; MI = myocardial infarction; NIDDM = non-insulin dependent diabetes mellitus; rs = reference sequence; T = thymine; UTR = untranslated region; ↑ = increase; ↓ = decrease; ↔ no association.

In addition to the SNPs listed above, two *FGB* promoter region SNPs were also investigated in this dissertation, FGB-rs2227385 and *FGB*-rs2227388. These are novel variants identified through sequencing of the *FGB* promoter region in selected Tswana individuals (Kotzé *et al.*, 2015). The 1000 genomes project has identified the same variants in individuals of other African countries, including Nigeria, Sierra Leone, Kenya, Gambia and Nigeria, with minor allele frequencies (MAFs) of 0.5 to 5% (Yates *et al.*, 2016). The MAF of these variants in the current study population is 3% and 16%, respectively (Kotzé *et al.*, 2015). No other publications on these novel variants, to our knowledge, exist.

As summarised in Table 2.2, most of the investigated fibrinogen variants are in and around the β chain gene, specifically within the promoter area. These SNPs are non-coding variants and do not alter the amino acid sequence of the protein, but exert a regulatory role (Roche *et al.*, 2003). Polymorphisms within this region can, therefore, affect the expression of fibrinogen and thereby associate with increased or decreased protein concentrations (Roche *et al.*, 2003).

Most of the studies summarised in Table 2.2 were epidemiological research conducted in adult men and women of European descent. *In vitro* research was performed only by van't Hooft *et al.* (1999). There is some inconsistency in the literature. Results obtained by Behague *et al.* (1996), van't Hooft *et al.* (1999), Mannila *et al.* (2005), Liu *et al.* (2005), Carty *et al.* (2010), Wassel *et al.* (2011) and Ken-Dror *et al.* (2012) differed on the larger consensus, although the discrepancy seems to be accounted for by differences in ethnicity, disease status and possibly study population size. Regardless of these differences, the notion that polymorphisms are able to contribute protectively or pathogenically to the fibrinogen phenotype in terms of expression and functionality is clear. Lack of consensus regarding functionality and isolated effects remains troublesome and clear mechanisms have not been elucidated, as a large degree of LD between these SNPs makes the identification of possible functional variants problematic.

In addition to the polymorphic effects on the fibrinogen phenotype, several alternative modifications to the fibrinogen genes have also been documented. These variants will be discussed in the following section. The fibrinogen γ' variant will be discussed in detail, as it will be used as a phenotype outcome throughout this dissertation.

2.3.4 Common fibrinogen splice variants

Fibrinogen is a highly heterogenous protein, of which over a million non-identical variations exist owing to polymorphic variation (Section 2.3.3), alternative splicing, proteolytic degradation, and post-translational modification (Henschen-Edman, 1995; Blombäck, 1996). Fibrinogen occurs primarily in a high MW (HMW) form (340 kDa), although proteolysis of one or both of the α

chains' COOH-termini results in low (LMW) and very low (VLMW) MW variants (Holm & Godal, 1984; Holm *et al.*, 1986). These variants are depicted in Figure 2, where the alterations in the αC domains are shown in red. In any individual, the distribution of molecular weight variants is typically 70% (HMW), 26% (LMW) and 4% (VLMW), although HMW fibrinogen can increase to up to 85% of total fibrinogen during the acute phase (Jensen *et al.*, 2000). Apart from the alterations to the αC domain, α chain transcript variants also exist. The major α chain transcript consist of five exonic regions (αA), although 1-3% of the α transcripts (αE) possess an additional exon and COOH-terminal (Mosesson, 2005; Standeven *et al.*, 2005).

Variation in the β chain, due to a 5-residue deletion close to the centre of the β chain polypeptide, has also been reported. This deletion results in a minor β chain isoform, 554 Dalton lower in MW, that comprises about 10% of the β chains (Brennan *et al.*, 2009). The minor isoform is associated with longer clotting times due to hindered polymerisation and a lower functional fibrinogen concentration resulting in faster clot lysis (Brennan *et al.*, 2009).

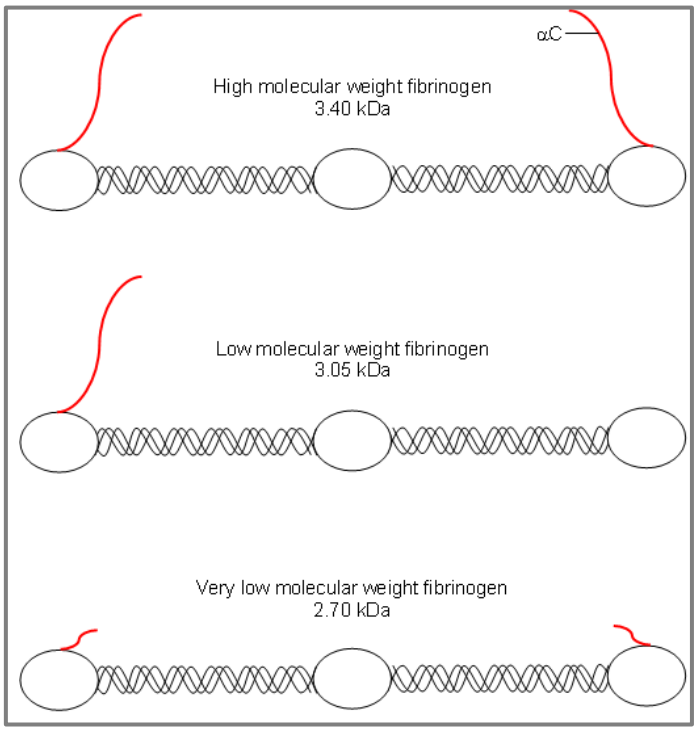


Figure 2.5 High, low and very low molecular weight fibrinogen molecules

Adapted from Weijers (2011)

Lastly, splice variation in the γ chain also exists. This review will focus specifically on the minor γ fibrinogen chain, γ'. In addition to being one of the phenotype outcomes in the current study, γ' fibrinogen is independently associated with altered clot structure (Collet *et al.*, 2004) and

increased CVD risk (Drouet *et al.*, 1999; Lovely *et al.*, 2002; Mannila *et al.*, 2007; Cheung *et al.*, 2008; Cheung *et al.*, 2009) and is, therefore, a critical variant to investigate.

The γ' fibrinogen chain arises owing to alternative messenger ribonucleic acid (mRNA) splicing in the 9th intronic region (Chung & Davie, 1984). To explain this adequately, consider the main fibrinogen γ chain variant, γA (Table 2.1). The γA fibrinogen transcript (*FGG*) consists of 10 exons (upon splicing out the introns). The γA chain terminates at a stop codon after the addition of a poly-A tail to the 10th exon (Fornace *et al.*, 1984). The γ' chain, however, occurs when alternative splicing results in the translation of the 20 amino acid 9th intron, containing two sulphated tyrosines, which replaces the four amino acids originating from exon 10 (Fornace *et al.*, 1984). Polyadenylation prior to the stop codon occurs similarly to the γA chain (Wolfenstein-Todel & Mosesson, 1980; Fornace *et al.*, 1984). Alternative splicing of the γ chain results in a γ' chain, 16 amino acids longer than what is coded for by only the *FGG* exons (Wolfenstein-Todel & Mosesson, 1980; Chung & Davie, 1984; Meh *et al.*, 1996). The mechanism described above is depicted below (Figure 2.4).

The γ chains of the fibrinogen protein occur mostly (85 to 92%) in a homodimeric manner with two γA chains, and rarely (< 1%) as two γ' chains (Wolfenstein-Todel & Mosesson, 1980). The heterodimeric variation γA/γ' constitutes 8 to 15% of total fibrinogen concentrations (Mosesson *et al.*, 1972). Fibrinogen consisting of one or more γ' chains has alternative physiological effects to that of homodimeric γA/γA fibrinogen, and is clinically associated with opposing effects on thrombosis in arterial and venous circulation. Fibrinogen γ' is prothrombotic through its ability to interact with thrombin by binding to thrombin exosite II, thereby localising thrombin to the fibrin clot where it remains active and is protected against degradation (Lovely *et al.*, 2003; Uitte de Willige *et al.*, 2009). Additionally, fibrinogen γ' has been shown to co-purify with FXIII and is associated with increased FXIII concentrations within the fibrin clot (Siebenlist *et al.*, 2005). Lastly, fibrinogen γ' alters fibrin clot properties and is associated with slower polymerisation, thinner fibrin fibres, altered intrafibrillar packing, and a prolonged clot lysis time (Allan *et al.*, 2012; Domingues *et al.*, 2016). On the contrary, fibrinogen γ' is associated with antithrombotic properties such as the inhibition of platelet aggregation owing to the elimination of a platelet binding site through the alternative splicing mechanism (Peerschke *et al.*, 1986; Farrell, 2004), increased activated protein C sensitivity, and decreased FV and FVIII activity (Macrae *et al.*, 2016). Fibrinogen γ' is associated with an increased risk of arterial thrombosis where the prothrombotic effects of the γ' chain prevails, whereas the opposite, antithrombotic effects, dominate venous circulation, where γ' is inversely associated with thrombotic risk (Uitte de Willige *et al.*, 2009; Macrae *et al.*, 2016).

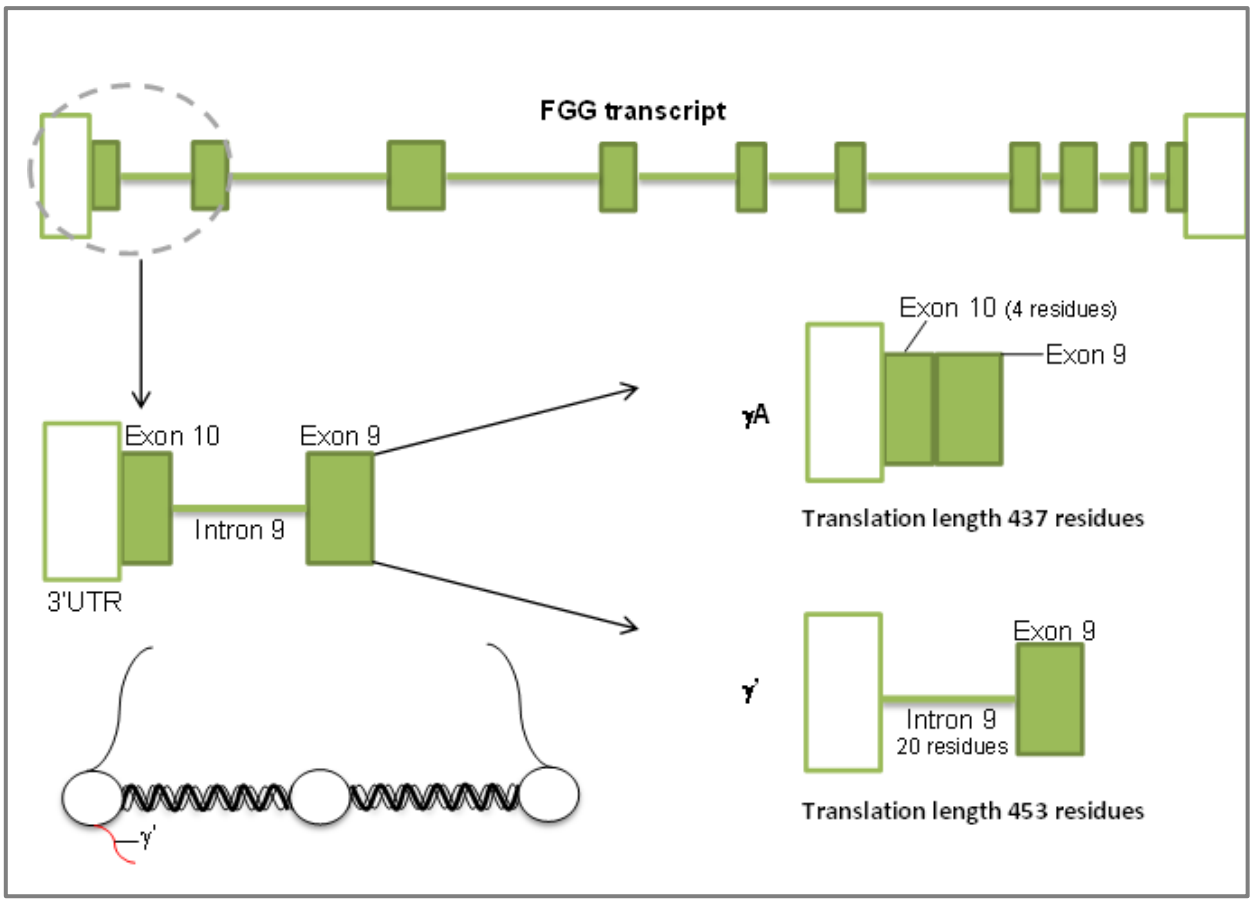


Figure 2.6 Alternative splicing mechanism of the fibrinogen γ chain

Adapted from Uitte de Willige et al. (2009)

The fibrinogen phenotype is heritable and subject to all the transcriptional and genetic processes discussed thus far. The following section will discuss the heritability of fibrinogen, including a brief summary of heritability studies to date, as well as the question regarding the missing heritability of fibrinogen.

2.3.5 The heritability of fibrinogen

Fibrinogen is a multifactorial phenotype, determined and influenced by a combination of various genetic and environmental contributors. A meta-analysis conducted by the Fibrinogen Studies Collaboration, using prospective data of more than 150 000 adults, concluded that the totality of modifiable factors investigated in the included studies only contributed to half of the fibrinogen variance (Kaptoge *et al.*, 2007). The other 51%, as mentioned previously, is reported to be genetic composition, as reported by various heritability studies. A brief summary of the findings of heritability studies conducted with fibrinogen as phenotype is presented in Table 2.2.

Table 2.2 Studies performed to estimate the heritability of fibrinogen

Study reference	Study population	Study population size	Estimated heritability	Environmental contribution to the variance in fibrinogen concentrations	Quantification of fibrinogen performed by
Hamsten *et al.* (1987)	Swedish	458 (MI survivors and their families) 407 (control subjects)	51%	Obesity and smoking had a negligible effect	Polymerisation time
Berg and Kierulf (1989)*	Norwegian	118 (59 twin pairs)	27%	Not reported	Clauss method
Reed *et al.* (1994)*	American	166 (83 twin pairs)	30%	Greater than genetic contribution, particularly smoking and diabetes	Clauss method
Friedlander *et al.* (1995)	Israeli	452 (82 pedigrees)	20 – 48%	10–64 % (environmental contribution increases with age)	Clauss method
Livshits *et al.* (1996)	Israeli	808 (204 pedigrees)	39%	Not reported	Kinetic method
Pankow *et al.* (1998)	American	2 029 (512 pedigrees)	34%	Age, anthropometry, lifestyle and metabolic factors contributed 19–29 %	Clauss method
Souto *et al.* (2000)	Spanish	397 (21 pedigrees)	34%	14%	Clauss method
de Lange *et al.* (2001)*	European	1 002 (501 twin pairs)	44%	56%	Clauss method
Yang *et al.* (2003)	European	1 595 (330 pedigrees)	24%	Not reported	Clauss method and modified method of Ratnoff and Menzie

*Twin study design; Clauss method = Clauss (1957); Modified method of Ratnoff and Menzie = Ratnoff and Menzie (1951) and Swaim and Feders (1967); Kinetic method = Hemker *et al.* (1979); Polymerisation time = Vermylen *et al.* (1963)

Heritability estimates have ranged from a modest 20% (Friedlander *et al.*, 1995) in an Israeli population to 51% (Hamsten *et al.*, 1987) in Swedish individuals. Differing results could be owing to a difference in study population ethnicities and inclusion criteria (e.g. apparently healthy or CVD patients classified by various disease definitions), as well as inconsistent statistical and biochemical methodologies, such as the quantification of fibrinogen concentration. Fibrinogen concentrations can be measured through either functional or antigen methodologies. The results obtained from these two methods are well correlated, but still somewhat discrepant, as these assays measure alternative properties of the same protein (Kamath & Lip, 2003; Lowe *et al.*, 2004). Functional assays quantify the 'clotting fibrinogen' by measuring the rate of the conversion of fibrinogen to stable fibrin, by fitting turbidity graphs to determine the concentration of the functional fibrinogen against the clotting time of known fibrinogen concentrations (Clauss, 1957). The antigen assays, however, are designed to measure the total fibrinogen concentration present in the plasma (Lowe *et al.*, 2004). As non-functional fibrinogen variants and splice variants with alternate clotting properties have been reported, it is important to note that functional assays may under-report true total concentrations, and antigen assays may over-report on physiologically relevant/functional fibrinogen. These differences in measurement of fibrinogen concentration may therefore, at least in part, explain the different heritability estimates. Twin studies are the greatest predictors of heritability, and in the context of fibrinogen, estimations from twin studies vary from 20 to 44%. Although the studies discussed above report different heritability estimates and vary in conclusion, the moderate to great heritable component of the fibrinogen phenotype remains clear.

Subsequently, various GWAS have been conducted in an effort to identify common SNPs that contribute to the above-mentioned heritability. These GWAS include up to 10.7 million SNPs in populations of more than 120 000 individuals and have, as yet, only been able to allocate 1.3 to 3.7% of the genetic variability of fibrinogen (Sabater-Lleal *et al.*, 2013; Huffman *et al.*, 2015; de Vries *et al.*, 2016). These investigations have included all the autosomal chromosomes as well as the X-chromosome, and used all of the available imputation methods (Sabater-Lleal *et al.*, 2013; de Vries *et al.*, 2016). The SNPs contributing to the 3.7% were mostly located outside the fibrinogen gene cluster (as will be discussed in Section 2.4.3), and consisted of a large number of variants, each contributing a relatively small amount.

The discrepancy between heritability and SNP associations affirms the research question regarding the missing heritability of the fibrinogen phenotype. The main aim of this dissertation is to address the question regarding fibrinogen's missing heritability by investigating the fibrinogen phenotype in terms of variability in concentration and functionality in an African

population. The African continent is the ancestral origin of human genetics and is imperative in the clarification of the genetic aetiology of common variants, particularly those that carry pathological relevance. Three approaches will be adopted in the effort to address the missing heritability of fibrinogen, and the theoretical background of these approaches will be discussed in the following section.

2.4 ADDRESSING THE MISSING HERITABILITY OF FIBRINOGEN

The large gap between genetic association and heritability studies raised the question of the missing heritability of the genetic architecture of fibrinogen. Furthermore, information on the fibrinogen phenotype in Africans remains largely unexplained, although the phenotype is pathological in the African setting (Pieters *et al.*, 2011; Lammertyn *et al.*, 2015). Identifying possible causal genes and variants in the greatly diverse African genome could advance not only the understanding of thrombogenesis in Africa, but also the global understanding of the genotypic regulation of the fibrinogen phenotype.

The theoretical background of the three main approaches to addressing the missing heritability of fibrinogen that will be incorporated further in this study will be discussed in the following three sub-sections. These include firstly the basic principles of genetic linkage, LD and haplotypes; secondly, gene-environment interactions, with specific focus on the role of IL-6; and lastly pleiotropic and polygenic co-regulatory mechanisms within a candidate gene approach.

2.4.1 Genetic linkage, linkage disequilibrium and haplotypes

Genetic linkage refers to the co-segregation of alleles based on the genomic distance between them (Flint-Garcia *et al.*, 2003). Genetic linkage is reflected in the associated inheritance of polymorphisms in close proximity, as genes that are physically situated close to one another are often inherited without recombination events. Linkage disequilibrium (also known as allelic association) is the non-random inheritance of alleles owing to shared inheritance, natural selection and migration (Devlin & Risch, 1995; Jorde, 2000; Pritchard & Przeworski, 2001). Simplified, it is the co-occurrence of specific alleles, more than would be expected by chance. Theoretically, complete LD implies that the information of a specific allele can be imputed based on the knowledge of an SNP in LD with it (Collins *et al.*, 1997; Goldstein, 2001). Loci situated closely together (high linkage) may result in high LD between them, although high linkage is not a requirement for LD (Flint-Garcia *et al.*, 2003).

LD is statistically quantified primarily by D' and r^2 values ranging from 0 to 1; where 0 indicates independent loci and 1 complete LD (Lewontin, 1964; Hill & Robertson, 1968; Jorde, 2000). D

is an indication of the difference between the observed and expected frequency of the co-occurrence of alleles under complete linkage equilibrium, which is the product of the respective MAFs (Lewontin, 1964; Devlin & Risch, 1995). The D-value is scaled to a $|D'|$ value in order to produce comparable outcome measures between 0 and 1. D' is the calculated D value divided by the maximum probable D (Lewontin, 1964; Hedrick, 1987; Lewontin, 1988). As the calculation of D' involves the MAF, the sample size greatly influences the reliability of lower D' values (Jorde, 2000). The r^2 variable is the statistical correlation between the two loci involved (Hill & Robertson, 1968), and is an indication of both recombination and mutation (Wall & Pritchard, 2003). It is calculated by dividing the square of the D-value by the product of the four genotype frequencies involved (Pritchard & Przeworski, 2001). The inverse of the r^2 value is useful as a predictor of sufficient population and control group size estimates can be inferred through this value (Carlson *et al.*, 2003; Wall & Pritchard, 2003).

LD between SNPs in large genomic regions often form block-like patterns across these chromosomal regions, referred to as haplotype blocks (Daly *et al.*, 2001; Patil *et al.*, 2001; Reich *et al.*, 2001; Gabriel *et al.*, 2002). The term "haplotype" refers to the particular combination of alleles carried along a gene or chromosome (International HapMap Consortium, 2005) . The term is based on "haploid", and refers to the inherited single chromosomal set compiled from both the paternal and maternal chromosome. Haplotype blocks are regions with consistently high LD and low population haplotype diversity, which are separated by regions of great recombination, also referred to as "recombination hotspots" (Goldstein, 2001; Jeffreys *et al.*, 2001; Patil *et al.*, 2001; Cardon & Abecasis, 2003). Haplotype blocks are identified when: (1) the average D' is closer to 1 than the pre-determined D' threshold throughout a selected chromosomal region (Reich *et al.*, 2001); (2) a few selected SNPs are able to predict the SNPs contained in the larger chromosomal region adequately (Patil *et al.*, 2001), or (3) the diversity of haplotype combinations is low (Daly *et al.*, 2001). The confidence intervals of the D' statistic are used to generate haplotype blocks within the genome (Gabriel *et al.*, 2002). Upon removal of all SNPs with a MAF of less than 5%, each SNP pair comparison is termed "strong LD", "strong recombination" or "inconclusive". A block is then generated if more than 95% of the informative SNPs are evident of "strong LD" (Gabriel *et al.*, 2002; Barrett *et al.*, 2005). Alternative algorithms, including the "four gamete rule" (Wang *et al.*, 2002) and "solid spine of LD" (Barrett *et al.*, 2005), are also used to define haplotype blocks, although as subsequent chapters report the method suggested by Gabriel *et al.* (2002), these methods will not be elaborated on.

Haplotypes have greatly advanced genetic association studies to narrow the search for causal variants, as well as to uncover a larger stretch of genome with fewer genotyped SNPs (Risch & Merikangas, 1996; Goldstein, 2001; Patil *et al.*, 2001; Zhang *et al.*, 2002). Associations are,

therefore, firstly identified with specific haplotypes, where after functional research will be performed on single SNPs. Haplotypes are labelled using tagging SNPs that are efficient proxies for the entire haplotype, and provide an excellent platform for imputation (Johnson *et al.*, 2001; Patil *et al.*, 2001; Gabriel *et al.*, 2002). The limitation of this approach is that with high LD and few haplotypes, it is often difficult to differentiate between functional SNPs and bystanders in high LD with the actual causal variant (Crawford & Nickerson, 2005). In the context of fibrinogen, this has been particularly problematic in the promoter area of the β chain where almost complete LD exists between *FGB* -1420G/A, -455G/A, -993C/T and -148C/T (Green, 2001; Baumann & Henschen, 1994; van't Hooft *et al.*, 1999; Mannila *et al.*, 2005).

The aim of this investigation is, therefore, to overcome this barrier by using a study population known to have low LD and high recombination. Genetic variability in Africans is greater than in any other ethnic group (Chen *et al.*, 1995; Teo *et al.*, 2010). An investigation by Schuster *et al.* (2010) reported greater genetic diversity within a single ethnic group in Africa (Khoisan compared to Khoisan) than in unrelated ethnic groups elsewhere (Asian and European). In addition, preliminary results obtained in the investigation of the fibrinogen gene cluster in the South African Tswana population revealed much higher recombination rates than those reported for Europeans (Kotzé *et al.*, 2015).

In terms of haplotypes, clear differences in haplotype blocks between individuals of African descent (Yoruban and admixed African Americans), and non-African individuals (Asian and European) have been reported (Gabriel *et al.*, 2002). Haplotype blocks in 54 chromosomal regions each spanning 250 kb, in various population groups, were investigated by Gabriel *et al.* (2002), using the confidence bounds of D' where "strong LD" was defined as: $0.98 \geq D' \geq 0.70$, and strong evidence of recombination as a $D' < 0.9$. Recombination rates were three to four times higher in individuals of African descent, and were affected by distance more severely, as observed when recombination rates were 50% at an 8 kb region in Africans but only at 22 kb in the non-African population groups. Gabriel *et al.* (2002) concluded that more than three times as many SNPs would be needed to generate a fully powered haplotype association study in African population groups. This observation is consistent with many others, as reviewed by Wall and Pritchard (2003), reporting that LD in non-African populations is less complex and extends over wider genomic regions, probably because of population bottlenecks occurring at the time when humans first left the African continent (Frisse *et al.*, 2001; Reich *et al.*, 2001; Wall, 2001; Tishkoff & Williams, 2002). Although lack of LD is limiting to imputation and genome-wide investigations, the possible limitation will be used as a strength in the attempt to identify functionality in a small genomic region. This dissertation is the first investigation of haplotypes in the fibrinogen genes in an African population, has the potential to generate novel

SNP combinations and haplotypes and could advance the investigation of fibrinogen's missing heritability.

The first approach to be pursued in this dissertation is, therefore, a focused investigation of the fibrinogen gene SNPs that have shown functional relevance in Europeans, as well as the haplotypes of which these SNPs form part. The second approach, investigating the interaction of environmental contributors, specifically IL-6, with fibrinogen SNPs will be discussed in the following section.

2.4.2 Gene-environment interactions

Apart from the clear genetic influence, strong and recurring associations of various external stimuli with fibrinogen concentrations have been reported in large-scale epidemiological studies, and have been thoroughly reviewed in various publications (Folsom *et al.*, 1991; Dotevall *et al.*, 1994; de Maat, 2001; Kamath & Lip, 2003; Pulanić & Rudan, 2005; Pieters & Vorster, 2008; Arbustini *et al.*, 2013). Briefly, fibrinogen concentrations are positively associated with age, female gender (also menopause, pregnancy and oral contraceptives), BMI, low-density lipoprotein, triglycerides, physiological trauma such as surgery, infection, strenuous exercise and diabetes mellitus (blood glucose and insulin concentrations). In contrast, fibrinogen concentrations are inversely associated with socio-economic status, level of education, regular exercise and fitness, moderate alcohol consumption, high-density lipoprotein and the consumption of fish, vitamin A, C and E (Folsom *et al.*, 1991; Dotevall *et al.*, 1994; de Maat, 2001; Kamath & Lip, 2003; Pulanić & Rudan, 2005; Pieters & Vorster, 2008; Arbustini *et al.*, 2013).

The environmental control of fibrinogen has been acknowledged by heritability studies, reporting 14% to 64% of the fibrinogen phenotype to be environmentally determined (Table 2.3). Almost half of the variance in fibrinogen concentrations in the large meta-analysis performed by the Fibrinogen Studies Collaboration was determined by physical activity, BMI, smoking and alcohol consumption habits (Kaptoge *et al.*, 2007). BMI, IL-6 concentrations, smoking and insulin concentrations contributed to 30.6% of the variation in the study performed by Mannila *et al.* (2005), while in the GWAS conducted by Sabater-Lleal *et al.* (2013), smoking and BMI alone contributed to 5.3%. Environmental factors and genetics each seemed to contribute to about half of the variance in fibrinogen.

Theoretically, all the modifiable and non-modifiable contributors, including both environmental and genetic factors identified to date, should represent the full variance of fibrinogen. This is, however, not the case (Manolio *et al.*, 2009; Sabater-Lleal *et al.*, 2013). The large contribution

of both factors inspired a research question based on what the interactive effect of these two highly influential determinants would be on the ability to explain the missing heritability of fibrinogen. Could the polymorphic variance within the fibrinogen genes react to environmental queues, and thereby have a greater regulatory effect? There is ample evidence from the literature that suggests that gene-environment interactions can influence protein phenotypes (Hunter, 2005; Manolio *et al.*, 2009; Manuck & McCaffery, 2014).

Some evidence in terms gene-environment interactions influencing fibrinogen concentration is also available. Greater increases in fibrinogen concentrations have been reported in the presence of the FGB -455 A allele in current smokers, upon intensive exercise, and after surgery (Green *et al.*, 1993; Thomas *et al.*, 1996; Cotton *et al.*, 2000; Brull *et al.*, 2002; Baumert *et al.*, 2014). Similarly, an interactive effect of the *FGB* -1420 A allele and smoking (Baumert *et al.*, 2014) and *FGA* -rs6050 G allele and BMI (Jeff *et al.*, 2015) on fibrinogen concentrations has been reported. These interactions could all be linked through a common causal pathway: inflammation. Livshits *et al*. accurately hypothesised in their heritability study in 1996 that loci with regulatory functions in terms of fibrinogen exist. Although these loci had not been adequately described at the time of their publication, their large effect on the observed fibrinogen variance was clear. The effect of these loci on the transcriptional regulation of fibrinogen has now been fully elucidated (Section 2.3.2). As mentioned previously, the transcriptional regulation of fibrinogen is mediated by cytokine interaction, and is observed in the ability of IL-6 and glucocorticoids to induce fibrinogen expression greatly (Dalmon *et al.*, 1993; Vasse *et al.*, 1996).

As discussed in Section 2.3.2, the three fibrinogen genes all have response elements contained in their respective promoter areas that respond particularly to IL-6 during the APR. IL-6 is the primary inflammatory mediator of fibrinogen synthesis and genotype-IL-6 interactions potentially influencing the expression of fibrinogen have been reported in other ethnic groups (Krobot *et al.*, 1992; Folsom, 1995; Cushman *et al.*, 1996; Burger *et al.*, 2004; Sinha *et al.*, 2005; Wannamethee *et al.*, 2005; Kaptoge *et al.*, 2007). In Europeans, the minor allele at FGB -455 and FGB -854 is associated with greater IL-6-induced fibrinogen expression, whereas the opposite has been reported for FGB -1420A and -148T (Verschuur *et al.*, 2005; Morozumi *et al.*, 2009).

The second approach to address the missing heritability of fibrinogen is, therefore, to investigate the interactive effect of IL-6 and fibrinogen SNPs in determining fibrinogen phenotypes in an African population. This approach offers the opportunity to identify novel interactions, as Africans differ in terms of both the fibrinogen genotype and the IL-6 and

fibrinogen phenotypes. Unique interactions could be observed, as the current study population is evident of a state of low-grade chronic inflammation (Kotzé *et al.*, 2014; Kotzé *et al.*, 2015). In addition, two novel β chain promoter region variants are investigated in this dissertation that are not polymorphic elsewhere, but could be located in an IL-6 responsive region.

The following section will review the theoretical background for the third approach, the regulation of complex phenotypes by a wide range of polymorphic variance able to regulate fibrinogen pleiotropically and polygenically through common pathways.

2.4.3 Pleiotropic and polygenic co-regulation

A "cross-phenotype" association refers to the observation of a single locus or genetic region that is associated with more than one, often unrelated, phenotype (Solovieff *et al.*, 2013). "Pleiotropy" is based on the same principle; although pleiotropic relationships are indicative of each locus exerting a true independent genetic effect and not simply both loci being associated with a common pathway (Parkes *et al.*, 2013; Gratten & Visscher, 2016). The investigation of cross-phenotype and pleiotropic relationships is valuable, as biochemical phenotypes fluctuate in response to various external influences, whereas the genome remains static. The molecular associations may, therefore, present more accurate observations of a true and constant phenotype.

Previous GWAS have identified numerous cross-phenotype and pleiotropic effects in relation to the fibrinogen phenotype (Table 2.4). Fibrinogen is regulated by the small accumulative effect of a large number of SNPs across the genome. Although the genetic architecture of the fibrinogen phenotype is highly polygenic and reflects the influence of numerous loci, these loci are all involved in only a few pathways or physiological mechanisms, including the inflammatory, immunological, coagulation and metabolic pathways (Dehghan *et al.*, 2009; Wassel *et al.*, 2011; Sabater-Lleal *et al.*, 2013; Baumert *et al.*, 2014; de Vries *et al.*, 2016). Table 2.4 provides a summary of the main findings of GWAS that have investigated the fibrinogen phenotype, indicating the loci that are involved in the above-mentioned pathways in the footnote. Very few loci outside of these pathways have been identified. Figure 2.5 depicts the variants that have been associated with the fibrinogen phenotype. These summaries provide some indication of the widespread genomic control of the fibrinogen phenotype and include 233 variants outside the fibrinogen gene cluster (Yates *et al.*, 2016) Even though progress has been made in the identification of these variants, the total amount of variance explained remains negligible.

Table 2.3 Summary of genome-wide association studies with fibrinogen as phenotype

Reference	**Participants** N (Ethnicity)	**Variants included**	**Summary of main finding**	**Variance explained** (%)	**Fibrinogen quantified *via* a combination of:**	**Mean fibrinogen/range of fibrinogen means among included studies** (g/L)
de Vries *et al.* (2016)	120 246 (EUR)	≈ 10.7m SNPs ≈ 1.2m in/dels	41 significant loci identified, of which 18 were new. Prominent loci include *FGB*, *IRF1* *STAT3*, *HNF4*, *SH2B3*	3%	Clauss method Immunophelometric Prothrombin conversion time	2.63 – 4.56
Huffman *et al.* (2015)	76 316 (AFR, EUR, ASI, HIS)	Low frequency and rare variants (NR)	6 significant loci identified in the following genes: *FGB*, *FGG*, *PDLIM4*, *HNF4A*	1.3% (EUR) 0.12% (AFR)	Clauss method Immunophelometric Prothrombin conversion time	2.51 – 4.55 (EUR) 2.66 – 3.80 (AFR) 3.29 (ASI) 3.59 (HIS)
Sabater-Lleal *et al.* (2013)	81 668 (EUR) 8289 (AA) 1366 (HIS)	≈ 2.5m SNPs	985 SNPs in 23 genes 8 known (*IL6R, NLRP3, IL1RN, CPS1, PCCB, FGB, IRF1* and *CD300LF*), and 15 novel *JMJD1C, LEPR, PSMG1, CHD9, SPPL2A, PLEC1, FARP2, MS4A6A, TOMM7/IL6, ACTN1, HGFAC, IL1R1, DIP2B* and *SHANK3/CPT1B*	3.7%	Clauss method Immunophelometric method	2.67 – 4.56
Wassel *et al.* (2011)	23 634 (EA) 6 657 (AA)	47 539 SNPs	In EA 32 SNPs spanning 11 genes *FGA*, *FGB*, *FGG*, *IL6R*, *CPS1*, *PCCB*, *HDLBP*, *IL1RN*, *NLRP3*, *IRF1-SCL22A5* In AA 6 SNPs spanning *FGA*, *FGB*, *FGG*	1.8 – 3.8% (AA) 2.2 – 15% (EA)	Clauss method Immunophelometric method Ratnoff and Menzie	2.97 – 3.57 (EA) 3.20 – 3.63 (AA)
Danik *et al.* (2009)	17 686 (EUR)	337 343 SNPs	19 SNPs in 5 chromosomal regions near *IL-6R*, *CPS1, FGA*, *FGB*, *FGG, IRF1, SLC22A5*, *SLC22A4*, *SLC9A3R1, NAT9*, *CD300LF*	1.93%	Immunoturbimetric assay	3.51 (median)

Dehghan *et al.* (2009)	22 096 (EUR)	2 661 766 SNPs	73 SNPs clustered around 4 genes: *FGB*, *IRF1*, *PCCB*, *NLRP3**	< 2%	Clauss method Immunophelometric Ratnoff and Menzie Prothrombin conversion time	2.70 – 3.20

AFR = African; *ACTN1 = actinin alpha 1;* ASI = East Asian; *CD300LF = CD300 molecule-like family member F*[δ]*; CHD9 = chromodomain helicase DNA binding protein 9*[δ]*; CPS1 = carbamoyl-phosphate synthase 1*[δ]*; DIP2B = Disco interacting protein 2 homolog B ;* EA = European American; EUR = European; *FARP2 = FERM, ARH/RhoGEF and pleckstrin domain protein 2; FGA = fibrinogen alpha chain gene*[#]*; FGB = fibrinogen beta chain gene*[#]*; FGG = fibrinogen gamma chain gene*[#]*; HDLBP* = high density lipoprotein binding protein[δ]*; HGFAC = hepatocyte growth factor activator*[δ]*;* HIS = Hispanic; *HNF4A = hepatocyte nuclear factor 4 alpha*[δ]*; IL1R1 = interleukin 1 receptor, type I*; IL1RN = interleukin 1 receptor antagonist*; IL6R = interleukin 6 receptor*; IRF1 = interferon regulatory factor 1*[&] *; JMJD1C = jumonji domain containing 1C* [δ]*; LEPR = leptin receptor*[&]*; MS4A6A = Membrane Spanning 4-Domains A6A; NAT9* = N-Acetyltransferase 9[&]*; NLRP3 = nucleotide-binding domain and leucine-rich repeat containing Pyrin Domain Containing*; PCCB = propionyl-coa carboxylase beta subunit*[δ]*; PDLIM4 = PDZ and LIM domain protein 4; SH2B3 = SH2B adaptor protein 3**[&]*; PLEC1 = plectin 1l PSMG1 = proteasome assembly chaperone 1; SHANK3 = SH3 and multiple ankyrin repeat domains 3; CPT1B = carnitine palmitoyltransferase 1b*[&]*; SLC22A4 = solute carrier family 22 member 4; SLC22A5 =* solute carrier family 22 member 5*; SLC9A3R1 = solute carrier family 9 member a3 regulator 1*; SNPs = single nucleotide polymorphisms*; SPPL2A= signal peptide peptidase like 2A; STAT3* = signal transducer and activator of transcription 3**; TOMM7 = translocase of outer mitochondrial membrane 7;* [#]Related to coagulation; [&]Related to immunity; *Related to the inflammation; [δ]Related to metabolism

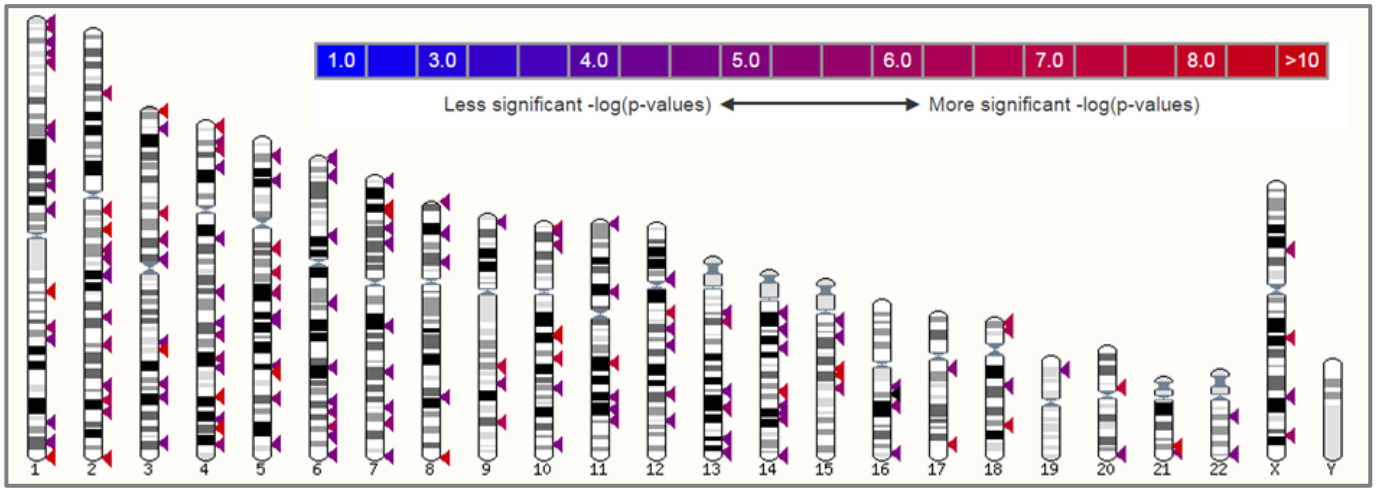


Figure 2.7 Human karyotype indicating loci significantly associated with the fibrinogen phenotype

Image generated through Ensembl release 85 (Yates *et al.*, 2016)

Regardless of the large sample sizes and genomic ranges covered by GWAS, the fibrinogen phenotype still remains largely unexplained. The most successful attempt at exploring the fibrinogen genotype was made by Sabater-Lleal *et al.* (2013), although the authors concluded that large variability still exists, and has remained unexplained by them or any other study thus far. The study performed by Sabater-Lleal *et al*. (2013) included Europeans (n = 91 323), African-Americans (n = 8 289) and Hispanic-Americans (n = 1 366). SNP-associations observed in individuals of European descent were largely reflected in the African-American study population. Although never identical, each European lead SNP was mirrored in an African-American SNP within a 200kB margin. While data from Sabater-Lleal *et al*. (2013) and other GWAS (Dehghan et al., 2009; Wassel et al., 2011; Baumert et al., 2014; de Vries et al., 2016) have been extremely valuable in identifying possible SNPs to be investigated, the unique composition of the African genome, compared to previously investigated population groups, does not allow for the extrapolation of data from European or ad-mixed African-American populations (Tishkoff & Williams, 2002; Campbell & Tishkoff, 2008). The investigation of pleiotropic and cross-phenotype associations in Africans using GWAS is additionally hindered by several methodological challenges, as reviewed by Teo *et al.* (2010), including low MAFs, complex LD patterns and the lack of ethnic-specific chips for whole genome analysis.

Regardless of the large sample sizes and genomic ranges covered by GWAS, the fibrinogen phenotype still remains largely unexplained. The most successful attempt at exploring the fibrinogen genotype was made by Sabater-Lleal *et al.* (2013), although the authors concluded that large variability still exists, and has remained unexplained by them or any other study thus far. GWAS data, although extremely valuable in identifying possible SNPs to be investigated, have been examined mostly in European and ad-mixed African-American populations (Wassel *et al.*, 2011; Sabater-Lleal *et al.*, 2013) and cannot be generalised to the black South African population because of the unique composition of the African genome (Tishkoff & Williams, 2002; Campbell & Tishkoff, 2008).

Therefore, the third objective of this dissertation investigates pleiotropic associations through a candidate gene approach. A candidate gene approach will provide an opportunity to investigate variation beyond that contained in global GWAS chips, including rare and novel variants. Furthermore, targeted genes with possible biological significance will be investigated in a more focussed manner. Specific genes or polymorphisms will be selected based on their phenotype's association with fibrinogen concentrations or the functional fibrinogen phenotype, namely turbidity-derived fibrin clot properties. This will serve as a biological filter to ensure a physiologically sound mechanism by which these SNPs could exert their effects. These SNPs will be selected from genotyping performed within the PURE population where the association

of the intermediate phenotypes are known. Genotyping was performed upon sequencing, to ensure the identification of rare and low MAF variants that might not have been identified in global GWAS (Wassel *et al.*, 2011). The resulting associations will be investigated in terms of specific regulatory sequence divergences ensuing polygenic regulation. As gene expression is under the control of transcriptional enhancers consisting of transcription factor binding and regulatory sites, SNPs will be investigated in terms of their functional significance through these pathways (Doniger & Fay, 2007; Han *et al.*, 2015; Tuğrul *et al.*, 2015).

Finally, in addition to the investigation of single polymorphisms with multiple phenotypic outcomes, there is also value in exploring the effect of a combination of numerous polymorphisms in predicting a single phenotype. This is calculated by means of polygenic risk scores (Dudbridge, 2013). These risk scores have the ability to determine the polygenic effect of harbouring several of the identified minor/risk alleles concurrently. A candidate gene approach in a study population that has not been investigated in terms of the fibrinogen phenotype could lead to the identification of novel associations and mechanisms and to greater understanding of the molecular regulation of the fibrinogen phenotype.

2.5 CONCLUSION

In this literature review, the theoretical background used to conceptualise and conduct this study is provided. It is clear that although fibrinogen is a physiological necessity, a higher fibrinogen concentration, as observed in the black South African population, does have pathological consequences. The involvement of fibrinogen in thrombogenesis is discussed, along with the genetic regulation thereof, including transcriptional regulation, influential polymorphisms and common splice variants.

This study originates from the discrepant literature regarding heritability *versus* genetic association data and aims to address the missing heritability of the fibrinogen phenotype by following the three approaches elaborated on in this chapter. This literature review is followed by two original articles. In Chapter 3, the investigation of the first and second approach, namely the independent (approach 1) and IL-6 interactive (gene-environment interactions, approach 2) association of selected fibrinogen gene variants and their haplotypes in terms of the fibrinogen phenotype, is presented. Thereafter, the candidate gene analysis of variants able to contribute to the fibrinogen phenotype through pleiotropic and polygenic regulation is reported (approach 3).

CHAPTER 3

INDEPENDENT AND IL-6-INTERACTIVE ASSOCIATIONS OF SELECTED FIBRINOGEN POLYMORPHISMS IN PREDICTING FIBRINOGEN AND CLOT-RELATED PHENOTYPES

This chapter includes:

- Guide for authors , British Journal of Haematology (Impact factor: 5.8);
- The original article titled: “Independent and IL-6-interactive associations of selected fibrinogen polymorphisms in predicting fibrinogen and clot-related phenotypes”.

GUIDE FOR AUTHORS: BRITISH JOURNAL OF HAEMATOLOGY

http://onlinelibrary.wiley.com/journal/10.1111/(ISSN)1365-2141/homepage/ForAuthors.html

Instructions to Authors

All papers should be submitted on the world wide web at http://mc.manuscriptcentral.com/bjh. Authors will need their entire manuscript in electronic format. Full instructions, a user ID and password are available at the site. Microsoft Word, Powerpoint and Excel files are automatically converted to PDF format for compatibility. Other recommended files types are RTF, GIF and JPEG. Users may have difficulties viewing other file types. Where possible, files should be combined and artwork embedded in the main manuscript file to aid users in downloading and viewing papers. Technical support for submission can be obtained from the ScholarOne Support website at http://mcv3support.custhelp.com or by telephoning +1 434 817 2040 ext 167.

The British Journal of Haematology requires the submitting author (only) to provide an ORCID iD when submitting his or her manuscript.

Editorial correspondence should be sent to BJH Office, 11 Auchinbaird, Sauchie, Alloa FK10 3HB, UK, Telephone and Fax +44 (0) 1259 220869, e-mail lorna@bjhaem.co.uk. Authors unable to submit their manuscript online should contact the BJH Office.

Submission of papers

The British Journal of Haematology invites papers on original research in clinical, laboratory and experimental haematology. All papers should include only new data which have not been published elsewhere. All authors are expected to disclose any commercial affiliations as well as consultancies, stock or equity interests and patent-licensing arrangements that could be considered to pose a conflict of interest regarding the submitted article. Specifics of such disclosures will remain confidential. If appropriate, general statements in the acknowledgements regarding such disclosures may be recommended by the editors.

Wiley Blackwell will dispose of all hardcopy or electronic material submitted two issues after publication.

NEW: Pre-submission English-language editing

Authors for whom English is a second language may choose to have their manuscript professionally edited before submission to improve the English. A list of independent suppliers of editing services can be found at http://authorservices.wiley.com/bauthor/english_ language.asp. All services are paid for and arranged by the author, and use of one of these services does not guarantee acceptance or preference for publication.

Ethical policy and guidelines

BJH encourages its contributors and reviewers to adopt the standards of the International Committee of Medical Journal Editors. BJH will not consider papers that have been accepted for publication or published elsewhere. Copies of existing manuscripts with potentially overlapping or duplicative material should be submitted together with the manuscript, so that the Editors can judge suitability for publication.

The Editors reserve the right to reject a paper on ethical grounds. Please read the Ethical Policies of BJH by clicking here.

Disclosure and competing interests statement

Authors are required to disclose financial interests in any company or institution that might benefit from their publication. A competing interest exists when a primary interest (such as patients' welfare or the validity of research) might be influenced by a secondary interest (such as financial gain or personal rivalry).

It may arise for the authors of a British Journal of Haematology article when they have a financial interest that may influence their interpretation of their results or those of others. Financial interests are the easiest to define and they have the greatest potential to influence the objectivity, integrity or perceived value of a publication. They may include any or all, but are not limited to, the following:

Personal financial interests: Stocks or shares in companies that may gain or lose financially through publication; consultant fees or fees from speakers bureaus other forms of remuneration from organisations that may gain or lose financially; patents or patent applications whose value may be affected by publication.

Funding: Research support from organisations that might gain or lose financially through publication of the paper.

Employment: Recent, present or anticipated employment of you or a family member by any organization that may gain or lose financially through publication of the paper. Any such competing interest that authors may have should be declared. The aim of the statement is not to eradicate competing interests, as they are almost inevitable. Papers will not be rejected because there is a competing interest, but a declaration on whether or not there are competing interests will be added to the paper.

Patent rights

Consultancy work.

All authors must disclose competing interests, or state "none" via the Journal's ScholarOne Manuscripts website.

All sources of funding must be disclosed in the Acknowledgments section of the paper. List governmental, industrial, charitable, philanthropic and/or personal sources of funding used for the studies described in the manuscript. Attribution of these funding sources is preferred.

Examples:

This work was supported by a grant from the National Institutes of Health, USA (DKxxxx to AB).

This work was supported by the Crohn's and Colitis Foundation of Canada (grant to AB and CD).

This work was supported by a grant from Big Pharma Inc. (to AB) and equipment was donated by Small Pharma Inc. EF received a graduate studentship award from the University of xxxxx.

For papers where there are no competing interests, all authors must include the statement 'Competing interests: the authors have no competing interests.'

We will also ask reviewers to provide a statement of competing interests.

AUTHORSHIP

All authors must fulfil the following three criteria:

Substantial contributions to research design, or the acquisition, analysis or interpretation of data,
Drafting the paper or revising it critically, and • Approval of the submitted and final versions. In the Acknowledgments section of the paper all authors, must indicate their specific contributions to the work described in the manuscript. Some examples include

X performed the research

Y designed the research study

Z contributed essential reagents or tools

A analysed the data

B wrote the paper.

An author may list more than one contribution, and more than one author may have contributed to the same element of the work. E.g. 'A performed the research, A and C analysed the data and wrote the paper, E contributed the knockout mice for the study and G designed the research study and wrote the paper'.

Role of professional medical writers in peer-reviewed publications

Please ensure that you follow the guidelines by the European Medical Writers Association on the role of medical writers. The guidelines emphasise the importance of respecting widely recognised authorship criteria, and in particular of ensuring that all people listed as named authors have full control of the content of papers. The role of professional medical writers must be transparent. Please name any professional medical writer among the list of contributors to any article for British Journal of Haematology (not only original research papers), and specify in the acknowledgements and statement of competing interests for the article who paid the writer. Writers and authors must have access to relevant data while writing papers.

Copyright

If your paper is accepted, the author identified as the formal corresponding author for the paper will receive an email prompting them to login into Author Services; where via the Wiley Author Licensing Service (WALS) they will be able to complete the license agreement on behalf of all authors on the paper.

For authors signing the copyright transfer agreement

If the OnlineOpen option is not selected the corresponding author will be presented with the copyright transfer agreement (CTA) to sign. The terms and conditions of the CTA can be previewed in the samples associated with the Copyright FAQs below:

CTA Terms and Conditions http://exchanges.wiley.com/authors/faqs---copyright-_301.html

For authors choosing OnlineOpen

If the OnlineOpen option is selected the corresponding author will have a choice of the following Creative Commons License Open Access Agreements (OAA):

Creative Commons Attribution Non-Commercial License OAA

Creative Commons Attribution Non-Commercial -NoDerivs License OAA

To preview the terms and conditions of these open access agreements please visit the Copyright FAQs hosted on Wiley Author Services http://exchanges.wiley.com/authors/faqs---copyright-_301.html and visit http://www.wileyopenaccess.com/details/content/12f25db4c87/ Copyright--License.html.

If you select the OnlineOpen option and your research is funded by certain funders [e.g. The Wellcome Trust and members of the Research Councils UK (RCUK) or the Austrian Science Fund (FWF)], you will be given the opportunity to publish your article under a CC-BY license supporting you in complying with your Funder requirements. For more information on this policy and the Journal's compliant self-archiving policy please visit: http://www.wiley.com/go/funderstatement.

For RCUK, Wellcome Trust, FWF authors click on the link below to preview the terms and conditions of this license:

Creative Commons Attribution License OAA
To preview the terms and conditions of these open access agreements please visit the Copyright FAQs hosted on Wiley Author Services http://exchanges.wiley.com/authors/faqs---copyright-_301.html and visit http://www.wileyopenaccess.com/details/content/12f25db4c87/Copyright--License.html.

Research papers

The majority of papers published in the Journal report original research into scientific and clinical haematology. All papers are subject to review and authors are urged to be brief; long papers with many tables and figures may require shortening if they are to be accepted for publication.

Short reports

Short reports which offer significant insight into scientific and clinical haematological processes may be published. They may include up to 1500 words of text, two figures or tables or one of each, and up to 15 references. A summary of up to 100 words should be followed by continuous text, subdivided if appropriate. Short reports could include important preliminary observations, short methods papers, therapeutic advances, and any significant scientific or clinical observations which are best published in this format. Publication of initial results which will lead to more substantial papers will generally be discouraged. Although submission of case reports is not encouraged, these will be considered if the report includes novel scientific material or is of especial clinical interest. Authors will receive proofs.

Annotations and reviews

These are normally invited contributions but suitable papers may be submitted to the Editor for consideration for this purpose. Previous issues of the Journal should be consulted for style of contribution and length.

Letters to the Editor

Correspondence which relates to papers which have recently appeared in the Journal may be published. The Editor reserves the right to invite response from the original authors for publication alongside. In addition, letters dealing with more general scientific matters of interest to haematologists will be considered. Letters should be as short as possible (but no more than 1000 words of text, two figures or tables or one of each, and up to 10 references).

Correspondence to the journal is accepted on the understanding that the contributing author licences the publisher to publish the letter as part of the journal or separately from it, in the exercise of any subsidiary rights relating to the journal and its contents.

Images in Haematology

Authors can submit for consideration an illustration (or, where appropriate, two or more related images) which is interesting, instructive and visually attractive, with a few lines of explanatory text and a maximum of six contributors. If there are more than three authors then the corresponding author should outline the contribution of each author that justifies their inclusion. The images (e.g. a clinical photograph, radiology, cytology, histology, a laboratory test) should be submitted in a digital format online at http://bjh.manuscriptcentral.com. High-quality glossy prints, transparencies, or digital files (see http://authorservices.wiley.com/bauthor/illustration.asp) should be sent to the BJH Office on acceptance. Because of space constraints, there are usually no references in the 'Images' section. However, it is at the discretion of the Images Editor to accept one reference if it appears essential to the contribution in question.

Announcements

Information about scientific meetings that are likely to be of general interest to readers of the Journal may be published at the discretion of the Editor. These should be sent to the Editor as early as possible prior to the event. Text should be as concise as possible, with a maximum of 150 words.

Preparation of manuscripts

Manuscripts should be formatted with wide margins and bear the title of the paper with the name and address of the author(s), together with the name of the hospital, laboratory or institution where the work has been carried out. Authorship should be restricted to individuals who have made a significant contribution to the study. The name, full postal address and e-mail address of the author to whom readers should address correspondence and reprint requests should be given on the first page; this will appear as a footnote in the journal and the publishers will send proofs to this author at the given address unless contrary instructions are written on the manuscript. Correspondence during the peer-review process will be with the author indicated during submission. A running short title of not more than 60 characters and spaces should be included. An informative summary of not more than 200 words must be included at the beginning of the paper and supplied when prompted during the online submission process. Papers should normally be divided into summary, introduction, methods (and/or materials), results, discussion, acknowledgements and references. SI units should be used throughout. Human DNA, gene, protein, and DNA restriction and modification enzyme nomenclature should be standardized as follows:

1) Human genes and alleles should be italicized capitals;
2) Human protein designations are the same as the gene symbol (i.e., written in upper case), but not italicized;
3) When distinguishing between mRNA, genomic DNA and cDNA, the relevant term should be given after the gene symbol, e.g. BCL2 cDNA;

4) D numbers should be as described in the Guidelines for Human Gene Nomenclature (Wain et al, 2002, Genomics, 79, 464-470; http://www.gene.ucl.ac.uk/nomenclature/guidelines.html)
5) Nomenclature for DNA restriction and modification enzymes and their genes should follow Roberts et al (2003, Nucleic Acids Research, 31, 1805-1812; http://nar.oupjournal.org/cgi/ content/full/31/7/1805).
The Editor reserves the right to make textual changes.

Keywords

Five keywords must be supplied after the summary.

Headings

The main categories of headings are side capitals, side italics and shoulder italics. If necessary, small capitals may be used for subsidiary main headings. For examples see articles in a recent issue of the Journal.

Illustrations

Illustrations should be referred to in text as, e.g., Fig 2, Figs 2, 4–7, using Arabic numbers. Individual figure files should bear a reference number corresponding to a similar number in the text, prints should be marked on the back with the name(s)of the author(s) and the title of the paper. Where there is doubt as to the orientation of an illustration the top should be marked with an arrow. Photographs and photomicrographs should be unmounted glossy prints and should not be retouched.

Where printed, diagrams should be on separate sheets. Lines should be of sufficient thickness to stand reduction. Each illustration should be accompanied by a legend clearly describing it. In the full-text online edition of the journal, figure legends may be truncated in abbreviated links to the full-screen version. Therefore the first 100 characters of any legend should inform the reader of key aspects of the figure.

It is the policy of the Journal for authors to pay the full cost for the reproduction of their colour artwork. Therefore, please note that if there is colour artwork in your manuscript when it is accepted for publication, Wiley-Blackwell require you to complete and return a colour work agreement form before your paper can be published. This form can be downloaded as a PDF from the internet here.

Once completed, the original hard copies of the form should be posted to the following address:

Customer Services (OPI)
John Wiley & Sons Ltd, European Distribution Centre
New Era Estate
Oldlands Way
Bognor Regis
West Sussex
PO22 9NQ
United Kingdom

If you are unable to access the internet, or are unable to download the form, please contact the Editorial Office and they will be able to email or fax a form to you. Any article received by Wiley Blackwell with colour work will not be published until the form has been returned.

Electronic artwork

We would like to receive the artwork accompanying accepted manuscripts in electronic form. Please save vector graphics (e.g. line artwork) in encapsulated Postscript format (EPS), and bitmap files (e.g. half-tones) in tagged image file format (TIFF). Detailed information on our digital illustration standards is available at http://authorservices.wiley.com/bauthor/ illustration.asp.

Tables

Tables should be as few as possible and should include only essential data; they should be printed on separate sheets and should be given Roman numerals.

References

We recommend the use of a tool such as Reference Manager for reference management and formatting. Reference Manager reference styles can be searched for at: http//www.refman.com/support/rmstyles.asp. Only papers closely related to the author's work should be cited. References should be made by giving the author's surname with the year of publication in parentheses. Where the reference contains more than two authors it should be given at each mention in the text with only the first surname plus et al, e.g. Jones et al (1948). If several papers by the same author (s) and from the same year, or by the same author but different subsequent authors in the same year are cited, a, b, c, etc., should be put after the year of publication, e.g. Jones et al (1948a, b). All references should be brought together at the end of the paper in alphabetical order, with all authors, titles of journals spelt out in full, and with both first and last page numbers given. The style to be used is that of any recent issue of the Journal.

Supporting information

Supporting Information can be a useful way for an author to include important but ancillary information with the online version of an article. Examples of Supporting Information include additional tables, data sets, figures, movie files, audio clips, 3D structures, and other related nonessential multimedia files. Supporting Information should be cited within the article text, and a descriptive legend should be included. It is published as supplied by the author, and a proof is not made available prior to publication; for these reasons, authors should provide any Supporting Information in the desired final format.

For further information on recommended file types and requirements for submission, please visit: http://authorservices.wiley.com/bauthor/suppinfo.asp

Production office

Wiley, 1 Fusionopolis Walk, #07-01 Solaris South Tower, Singapore 138628
Tel.: +65 6643 8465 / Fax: +65 6643 8008

Proofs

Proofs (except for Correspondence) The corresponding author will receive an email alert containing a link to a web site. A working e-mail address must therefore be provided for the corresponding author. The proof can be downloaded as a PDF (portable document format) file from the site. Acrobat reader will be required in order to read this file. This software can be downloaded (free of charge) from the following Web site: http://www.adobe.com/products/acrobat/readstep2.html. This will enable the file to be opened, read on screen, and printed out in order for any corrections to be added. Further instructions will be sent

with the proof. Hard copy proofs will be posted if no e-mail address is available. Excessive changes made by the author in the proofs, excluding typesetting errors will be charged separately.

NEW: Online production tracking is now available for your article through Wiley Blackwell's Author Services

Author Services enables authors to track their article - once it has been accepted - through the production process to publication online and in print. Authors can check the status of their articles online and choose to receive automated e-mails at key stages of production. The author will receive an e-mail with a unique link that enables them to register and have their article automatically added to the system. Please ensure that a complete e-mail address is provided when submitting the manuscript. Visit http://authorservices.wiley.com/bauthor for more details on online production tracking and for a wealth of resources including FAQs and tips on article preparation, submission and more.

Early View

The British Journal of Haematology is covered by Wiley Blackwell's Early View service. Early View articles are complete full-text articles published online in advance of their publication in a printed issue. Articles are, therefore, available for publication as soon as they are ready, rather than having to wait for the next scheduled print issue. Early View articles are complete and final. They have been fully reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in a final form, no changes can be made after online publication. The nature of Early View articles means that they do not yet have volume, issue or page numbers, so Early View articles cannot be cited in the traditional way. They are, therefore given a Digital Object Identifier (DOI), which allows the article to be cited and tracked before it is allocated to an issue. After print publication, the DOI remains valid, and can be continued to be used to cite and access the article.

Offprints

Free access to the final PDF offprint or your article will be available via Author Services only. Please therefore sign up for author services if you would like to access your article PDF offprint and enjoy the many other benefits the service offers. Paper offprints of the printed published article may be purchased if ordered via the method stipulated on the instructions that will accompany the proofs.

The British Journal of Haematology is covered by Current Contents, Chemical Abstracts, Current Clinical Cancer, CABS, ISI/BIOMED, Science Citation Index and ASCA.

Internet

Information on this journal and other Wiley Blackwell publications is on the Wiley Online Library homepage at: http://wileyonlinelibrary.com/journal/bjh

ARTICLE

Independent and IL-6-interactive associations of selected fibrinogen polymorphisms in predicting fibrinogen and clot-related phenotypes

H. Toinét Cronjé[a], Cornelie Nienaber-Rousseau[a], Lizelle Zandberg[a], Zelda de Lange[a], Fiona R. Green[b], Marlien Pieters[a]*

Running head: Genetic variance of fibrinogen in Africans

[a] Centre of Excellence for Nutrition, North-West University, Potchefstroom, South Africa

[b] Division of Cardiovascular Sciences, School of Medical Sciences, Faculty of Biology, Medicine & Health, University of Manchester, United Kingdom

Corresponding Author:

Prof. Marlien Pieters
Centre of Excellence for Nutrition
North-West University
Potchefstroom Campus
Private Bag X6001
Potchefstroom
2520, South Africa

Tel no: +27 18 299 2462
Fax no: +27 18 299 2464
E-mail address: marlien.pieters@nwu.ac.za

Word count:

Summary	197
Text	5187

SUMMARY

Interleukin-6 (IL-6) promotes the expression of fibrinogen, and polymorphic variation within the fibrinogen genes alters the magnitude of this expression. The identification of functional fibrinogen polymorphisms has been hindered by the high linkage disequilibrium (LD) reported in the fibrinogen gene cluster of Europeans. This study investigated two novel and 12 previously identified fibrinogen polymorphisms hypothesised to be of functional relevance, in 2010 Tswana individuals. We aimed to identify functional polymorphisms that contribute to total and γ' fibrinogen concentrations and clot-related phenotypes independently and through their interaction with IL-6, by utilising the high fibrinogen and IL-6 concentrations and the low LD reported in black South Africans. Fibrinogen was significantly associated with IL-6, thereby mediating associations of IL-6 with clot formation and structure, although confounding the association of IL-6 with clot lysis time. None of the common European fibrinogen haplotypes was present in this study population. Single nucleotide polymorphisms of particular functional relevance were FGB-rs7439150, -1420G/A and -148C/T that were significantly associated with fibrinogen concentration and altered clot properties, with several of these associations influenced by IL-6 concentrations. In addition, harbouring more minor alleles across the fibrinogen gene cluster concurrently led to a greater increase in IL-6-induced fibrinogen expression.

Key words: genotype-phenotype association, environment-gene interaction, inflammation, turbidity, gamma prime fibrinogen

INTRODUCTION

Fibrinogen is central to blood coagulation and the inflammatory response. As a haemostatic protein, fibrinogen is activated by thrombin to form fibrin, the main constituent of a blood clot (Sidelmann *et al,* 2000). As an acute phase reactant, fibrinogen is up-regulated upon stimulation by inflammatory cytokines, primarily interleukin-6 (IL-6), in response to physiological trauma such as infection/inflammation (Davalos and Akassoglou 2012). Both fibrinogen and IL-6 are prospectively associated with cardiovascular disease (CVD) risk (Danesh *et al,* 2005; Kaptoge *et al,* 2013; Tzoulaki *et al,* 2007), although causality is under debate (Hinds *et al,* 2016; Keavney *et al,* 2006; Meade *et al,* 2006; Sabater-Lleal *et al,* 2013).

The fibrinogen phenotype is heritable and several single nucleotide polymorphisms (SNPs) within the fibrinogen α, β and γ chain genes (*FGA*, *FGB*, *FGG*) have been identified as contributors to this heritability in Europeans (de Lange *et al,* 2001; Green 2001; Hamsten *et al,* 1987; Jacquemin *et al,* 2008). Fibrinogen expression is regulated on two levels: under basal conditions and during the acute phase response (Fish and Neerman-Arbez 2012; Fuller and Zhang 2001). This acute phase-induced increase in fibrinogen is largely mediated by IL-6, through the JAK/STAT pathway. In addition, sequences responsive to IL-6, and crucial for full IL-6-induced fibrinogen expression, have been identified upstream of the fibrinogen genes (Fish and Neerman-Arbez 2012; Fuller and Zhang 2001).

Genetic variation in the fibrinogen genes is hypothesised to alter the magnitude of fibrinogen expression in response to IL-6 (Jacquemin *et al,* 2008; Morozumi *et al,* 2009; Verschuur *et al,* 2005). Thus far, the focus of previous investigations have been on polymorphisms within the *FGB* promoter region that have been implicated in both epidemiological (Morozumi *et al,* 2009) and *in vitro* (Verschuur *et al,* 2005) investigations to interact with IL-6 in influencing fibrinogen concentrations. Greater IL-6-induced fibrinogen expression has been reported for the *FGB* -455A and -854A alleles (Brull *et al,* 2002; Cotton *et al,* 2000; Montgomery *et al,* 1996; Morozumi *et al,* 2009; Verschuur *et al,* 2005), while the opposite has been observed for *FGB* -1420A and -148T (Morozumi *et al,* 2009; Verschuur *et al,* 2005). Apart from the *FGB* polymorphisms, IL-6 interacted with *FGA*-rs2070011T *in vivo* to enhance fibrinogen expression, whereas *FGA*-rs6050, *FGG*-rs2066865 and *FGG*-rs1049636 had no effect (Jacquemin *et al,* 2008).

In addition to regulating total fibrinogen concentration, IL-6 is able to up-regulate the production of γ' fibrinogen, a common splice variant, comprising 8 to 15% of total fibrinogen concentration (Meh *et al,* 1996; Rein-Smith *et al,* 2013; Wolfenstein-Todel and Mosesson 1980). IL-6 specific

responsive elements have been characterised in the *FGG* promoter region (Duan and Simpson-Haidaris 2003; Zhang *et al,* 1995), and *in vitro* research reported greater increases in γ' compared to total fibrinogen in the presence of IL-6, suggesting alterations to the alternative splicing of the γ fibrinogen gene during inflammation (Rein-Smith *et al,* 2013). A higher concentration of the fibrinogen γ' isoform is a risk factor for arterial thrombosis, although research on the individual and IL-6-interactive effects of fibrinogen polymorphisms predicting γ' fibrinogen concentrations is lacking (Alexander *et al,* 2011; Cheung *et al,* 2008; Drouet *et al,* 1999; Uitte de Willige *et al,* 2009). Increased IL-6, total and γ' fibrinogen are also independently associated with altered clot properties, including faster clot formation, increased fibrin density, thinner fibrin fibres and decreased clot permeability (Bester and Pretorius 2016; Machlus *et al,* 2011; Macrae *et al,* 2016; Pieters *et al,* 2013; Undas *et al,* 2008). The functional effects of the interactions between genetic variants and IL-6 on clot properties have, however, not been described previously.

Tight LD within the fibrinogen SNPs in European individuals has made the identification of functional SNPs difficult. Studies have implicated the *FGB* -455G/A, -854G/A and - 148C/T polymorphisms to be functional in Europeans, although some of the literature thus far is contradictory, particularly when comparing epidemiological (Morozumi *et al,* 2009) and experimental work (Verschuur *et al,* 2005). The African population is known for its complex LD pattern and great genetic diversity, more so than any other population group (Chen *et al,* 1995; Teo *et al,* 2010). A previous publication from the South African arm of the prospective urban and rural epidemiology (PURE) study has indicated that this is also the case for the fibrinogen genes in the Tswana population (Kotzé *et al,* 2015). The lack of complete LD among polymorphisms in the fibrinogen genes in individuals of African descent is utilised in this current investigation to identify potentially functional polymorphisms in the fibrinogen gene cluster. In addition, increased fibrinogen and IL-6 concentrations have been observed in black South Africans when compared to their white counterparts (Lammertyn *et al,* 2015; Pieters and Vorster 2008). Individuals from the South African PURE population specifically have markers indicative of chronic low-grade inflammation, as reported by increased fibrinogen, IL-6, homocysteine and C-reactive protein (Pieters *et al,* 2011), making this population ideal for the investigation of IL-6-regulated gene interactions.

The objective of this study was to investigate selected fibrinogen polymorphisms in terms of their independent and IL-6-interactive effects on total and γ' fibrinogen concentrations, and the downstream functional effects in terms of clot formation, structure and lysis, thereby contributing

to the broader understanding of potential factors that regulate fibrinogen-related, clinically relevant phenotypes.

MATERIALS AND METHODS

Study population and ethical considerations

This study is affiliated with the international PURE (Teo *et al,* 2009) study and is a cross-sectional investigation of the baseline data collected in South Africa Details pertaining to participant selection and recruitment were published previously (Kotzé *et al,* 2015). In short, the study population consisted of 2010 apparently healthy, self-identified black Tswana-speaking South African adults. Participants were between the ages of 35 and 70 years, and were not suffering from any acute or chronic diseases. Ethical approval for the research was obtained from the Health Research Ethics Committee of the North-West University (NWU-00016-10-A1) and the study was conducted in accordance with the revised version (2000) of the Helsinki Declaration of 1975.

Blood collection and storage

Fasting blood samples were collected between 7:00 and 11:00 am, and centrifuged within 30 minutes of collection at 2000 x g for 15 minutes. Blood samples for haemostatic variables were collected in 3.2% sodium citrate tubes, for lipid and IL-6 analyses in tubes without anti-coagulants and for glycated haemoglobin ($HbA1_c$) in sodium fluoride tubes. Upon blood collection and centrifugation, samples were snap-frozen and stored at -80°C until analysis.

Biochemical analyses

$HbA1_c$ concentrations were quantified using a hexokinase method from Synchron® systems (Beckman Coulter Co., Fulleron, CA, USA), and the D-10 haemoglobin testing system (Bio-Rad, Hercules, California, USA). A Sequential Multiple Analyser Computer, using the Konelab™ auto-analyser (Konelab 20i, Thermo Fischer Scientific, Vantaa, Finland) was used to measure high-density lipoprotein cholesterol (HDL-c) concentrations. Plasminogen activator inhibitor type 1 activity ($PAI\text{-}1_{act}$) was quantified by an indirect enzymatic method (Spectrolyse PAI-1, Trinity Biotech, Bray, Ireland). Serum IL-6 was measured by means of the Elecsys, through ultra-sensitive enzyme immunoassays (Elecsys 2010, Roche, Basel, Switzerland).

Total fibrinogen concentrations were quantified using the modified Clauss method on the Dade Behring BCS coagulation analyser (Multifibrin U-test Dade Behring, Deerfield, IL, USA). An enzyme-linked immunosorbent assay using a 2.G2.H9 mouse monoclonal coating antibody

against human γ' fibrinogen (Santa Cruz Biotechnology, Santa Cruz, USA) and a goat polyclonal horseradish peroxidase-conjugated antibody against human fibrinogen (Abcam Cambridge, USA) was used to measure γ' fibrinogen concentrations. Fibrinogen γ' is expressed both as absolute concentration and as a percentage of total fibrinogen (γ' ratio).

Plasma fibrinolytic potential was determined by measuring turbidity with a spectrophotometer (A405) (de Lange *et al,* 2013). Tissue plasminogen activator (tPA; Actilyse, Boehringer Ingelheim, Ingelheim, Germany) was added to plasma clots induced by tissue factor (TF; Dade Innovin, Siemens Healthcare Diagnostics Inc., Marburg, Germany) according to the method of Lisman *et al.* (Lisman *et al,* 2005). The tPA and TF concentrations were modified to obtain clot lysis times (CLTs) between 60 and 100 minutes. Final concentrations in the plasma clots were 125 x diluted TF, 100 ng/mL tPA, 17 mmol/L $CaCl_2$, and 10 mmol/L phospholipid vesicles (Rossix, Mölndal, Sweden). Resultant turbidity curves were analysed using Origin® software version 8.5 (Origin lab®, 2010). CLT (minutes) was calculated as the difference between the time at the midpoint of clear and maximum turbidity (clot formation) and the midpoint between maximum and clear turbidity (clot lysis). In addition, lag time (minutes) was calculated as an indicator of the time required for the activation of the coagulation cascade by TF and for protofibrils to reach sufficient length to allow lateral aggregation. The slope of the curve ($x10^{-3}$ au/s) during clot formation was used as a representation of the rate of lateral aggregation of the fibrin protofibrils. Lastly, maximum absorbance, the increase in absorbance at the peak of the curve (nm), was used as an indicator of fibre diameter.

DNA isolation, SNP selection and genotyping

Genomic DNA was extracted from buffy coat using the QIAGEN FlexiGene™ kit. Unsuccessful extractions were repeated by means of the Maxwell® 16 DNA purification kit. The quality of the DNA was determined using the NanoDrop™ spectrophotometer (ND-1000, Wilmington, DE, USA). Fourteen SNPs, spanning the three fibrinogen genes, were selected for genotyping based on literature reporting the independent and IL-6-induced association of these SNPs with total or γ' fibrinogen concentrations (Cook *et al,* 2001; Green 2001; Jacquemin *et al,* 2008; Lim *et al,* 2003; Mannila *et al,* 2006; Reiner *et al,* 2006; Sabater-Lleal *et al,* 2013; Uitte de Willige *et al,* 2005; van't Hooft *et al,* 1999). Furthermore, the promoter region of *FGB* was sequenced in a subgroup of 28 randomly selected individuals for the identification of novel SNPs. ABI Prism®, BIGDye® Terminator version 3.1 Ready Reaction Cycle Sequencing Kits (Applied Biosystems, CityFoster, CA, USA) were used, and electropherograms were aligned using BioEdit (version 7.1.3.0, Ibis Biosciences, Carlsbad, CA, USA).

The polymorphisms selected for genotyping were *FGB* - rs7439150, rs2227385 (novel), rs1800789 (-1420G/A), rs2227388 (novel), rs1800791 (-854G/A), rs1800790 (-455G/A), rs1800788 (-249C/T), rs1800787 (-148C/T), rs4220, rs4463047, *FGA* - rs6050, rs2070011 (2224G/A), and *FGG* - rs2066865 and rs1049636 (9340T/C). Three methods were used to genotype these polymorphisms, of which two (Thermo Fischer Scientific® Taqman based assays and the Illumina® VeraCode GoldenGate assay technology using a BeadXpress® platform), have been described previously (Kotzé *et al,* 2015). In addition, *FGB*-rs1800789, rs1800790, rs4463047 and rs7439150 were genotyped by competitive allele-specific polymerase chain reactions (KASP) with supplies obtained from LGC Limited. Custom-designed assays and synthetic controls were manufactured by KBioscience (LGC, Middlesex, TW11 0LY, UK) and are presented in the online supporting information (Table 3.5).

A two-step 61-to-55°C touchdown polymerase chain reaction (PCR) protocol was performed in a Hydrocycler 4™ water bath thermal cycler (LGC, Middlesex, TW11 0LY, UK). The fluorescent signal of the PCR products was measured by a FLUOstar Omega SNP plate reader (BMG LABTECH Ltd) and the data were analysed by means of KlusterCaller™ V. 3.4.1.36 software (LGC, Middlesex, TW11 0LY, UK).

Statistical analyses

The statistical analyses were performed in three phases. Firstly, the SNPs were investigated in terms of their location, LD and haplotypes using the Ensembl database release 84 (Yates *et al,* 2016) and Haploview version 4.2 (Barrett *et al,* 2005). In addition, their independent associations with phenotype outcomes (total and γ' fibrinogen concentration, lag time, slope, maximum absorbance and CLT) were determined by independent t-tests and analysis of co-variance (ANCOVA) adjusting for confounders, using the statistical package for the social sciences (SPSS®) version 23 (IBM® Corp, 2015).

The second phase involved investigating the association of the fibrinogen phenotypes with IL-6 independent from the polymorphisms. Differences in phenotypes were tested with analysis of variance (ANOVA) and ANCOVA using IL-6 stratified by quartiles as the categorical variable. Tukey's honest significant difference *post-hoc* tests were performed to determine significant inter-group differences.

Lastly, interaction effects between IL-6 and the fibrinogen polymorphisms on the fibrinogen phenotypes were determined by creating interaction terms and entering them into an ANCOVA with full factorial analysis. Adjustments were made for age, gender, body mass index (BMI), human immunodeficiency virus (HIV) status, $HbA1_c$ and HDL-c based on Pearson correlations

and independent t-tests revealing these variables as possible confounders. When CLT was used as an outcome variable, it was also adjusted for PAI-1_{act}. Multiple testing was accounted for by setting significance at a $p < 0.01$ following Benjamini and Hochberg adjustments (Benjamini and Hochberg 1995). Q-Q plots were used to evaluate the normality of the standardised residuals. Interaction data are reported as the slope of the linear regression line between IL-6 and the fibrinogen phenotypes split according to genotype.

In order to determine the possibility of a cumulative effect of the SNPs on the IL6-fibrinogen relationship, polymorphisms indicating significant interactions were grouped by means of the generation of a simple genetic 'risk score'. Each genotype forming part of the model (significant interaction upon adjustment for multiple testing), was given a score per individual genotype. A score of one was given for major allele homozygotes; two was given to the heterozygotes and three to minor allele homozygotes. Where the minor allele frequency (MAF) was low and only two groups were compared (see below), a score of two was given all minor allele carriers.

A power calculation indicated that the smallest group of 71 individuals (predicted by promoter sequencing variation) would provide 80% power to detect a medium effect size (Cohen's d-value of 0.3) (Faul *et al,* 2007). Analyses for SNPs having fewer than 71 individuals in the minor allele homozygote group were therefore performed using two groups only, combining heterozygotes and minor allele homozygotes.

Any analyses (t-tests, ANOVA, full factorial analysis ANCOVA) yielding significant results in terms of total or γ' fibrinogen were followed by adjustments for these proteins during statistical testing using clot properties as outcome variables. These adjustments were made to identify SNP/IL-6-clot property outcomes that were not mediated by total and/or γ' fibrinogen concentration.

RESULTS

Description of individual polymorphisms

Ten of the 14 investigated polymorphisms are situated in and around the β chain gene, of which eight are within the promoter area, one in exon 8 (Arg478Lys, previously reported as Arg448Lys) and one in the 3' untranslated region (UTR). Two α chain variants, one in exon 2 (Thr331Ala, previously reported as Thr312Ala) and one in the promoter area, and two γ chain variants, both in the 3' UTR, are also included. The relative positions of these SNPs, spanning a 50 Kb region on chromosome 4, are illustrated in Figure 3.1. All the variants were in Hardy-Weinberg equilibrium (Hardy 1908).

LD and haplotype construction

Figure 3.2 depicts the LD pattern (illustrated by D' values upon an r^2 colour scheme) observed for the 14 SNPs. Lower recombination rates are indicated by increased numerical values (D') and darker coloured blocks (r^2). Blocks containing no numerical value indicate a D' of 1.0. Two haplotype blocks *FGB*-rs7439150*rs2227385*rs1800789 and *FGG*-rs2066865*rs1049636 were constructed *via* the method suggested by Gabriel *et al,* (2002). No complete LD (D' and r^2 = 1.0) was observed, as the LD pattern was disturbed owing to differing MAFs that led to relatively low r^2 values. Acknowledging the limited LD, all further analyses were performed using individual polymorphisms only.

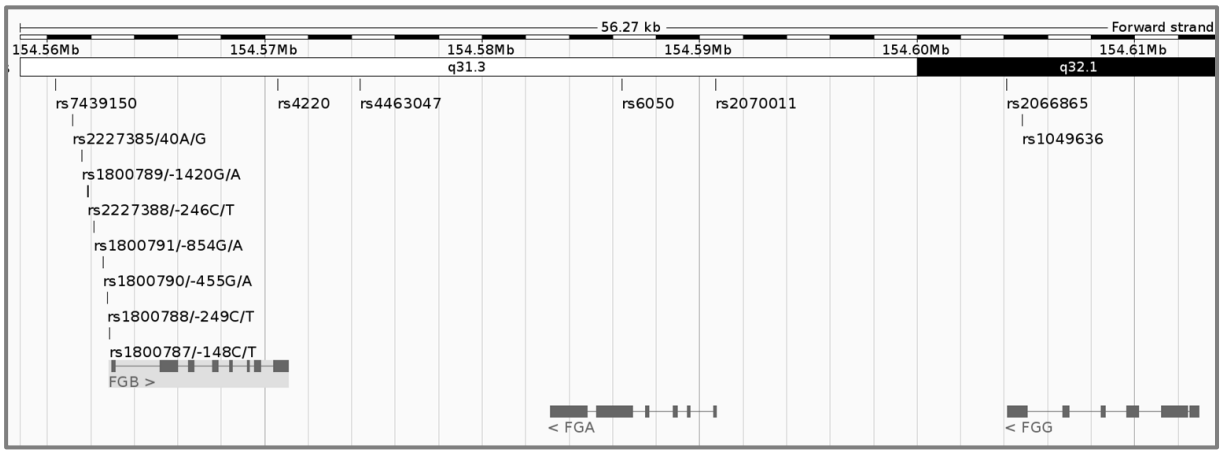


Figure 3.1 Fourteen polymorphisms spanning the fibrinogen gene cluster

Image generated through the Ensembl database (Yates et al, 2016)

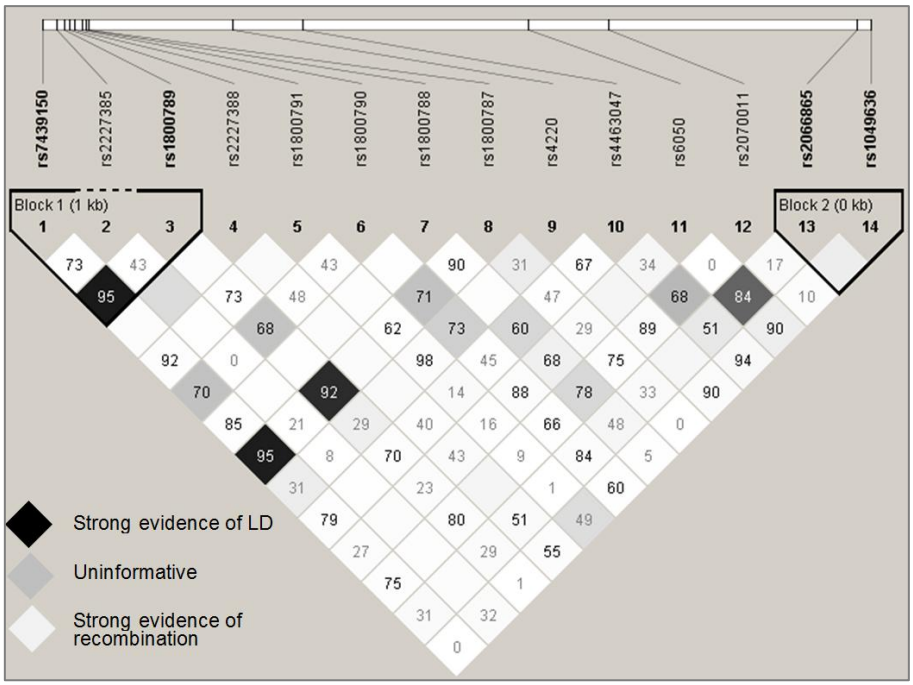


Figure 3.2 Pairwise LD structure of 14 fibrinogen SNPs, illustrated by D' values on an r^2 colour scheme

Empty boxes indicate D' values of 1.0. Increased numerical values indicate stronger evidence of LD. Image generated through Haploview software (Barrett et al, 2005)

Associations of individual polymorphisms with fibrinogen-related phenotypes

Ten of the 14 SNPs have been reported previously (Kotzé *et al,* 2015) and the results are therefore shown in the online supporting information only (Table 3.6 and 3.7). Briefly, *FGB* -854A and *FGG*-rs1049636C were positively associated with total fibrinogen (p = 0.04 and 0.0009), and *FGA*-rs2070011A with higher γ' fibrinogen concentrations (p = 0.008). In terms of clot properties, *FGB* -148T was associated with a larger fibre diameter (p = 0.001), whereas *FGA*-rs2070011A and *FGG*-rs1049636C, lost significance in terms of their effect on maximum absorbance (fibre diameter) upon adjustment for γ' and total fibrinogen, respectively (p = 0.06 and 0.33).

Associations of each of the four newly investigated promoter region SNPs with the outcome phenotypes are shown in Table 3.1. *FGB* -1420A was associated significantly with total fibrinogen concentrations (p = 0.02) and the same trend was observed for *FGB*-rs7439150A and *FGB* -455A (p ≈ 0.05). Larger fibre diameter and longer CLT were observed in the presence of the *FGB* -1420A allele (p = 0.04 and 0.02). In addition, *FGB* –455A and *FGB*-rs7439150A were positively associated with γ' fibrinogen concentrations (p = 0.02 and 0.005), and *FGB*-rs7439150A with fibre diameter (p = 0.02). Associations with clot properties

were greatly mediated by the primary fibrinogen associations, as most of the significance was lost upon adjustment for total and/or γ' fibrinogen, apart from the association between *FGB* -1420*A* and CLT, which remained (p=0.03).

Table 3.1 Association of selected upstream *FGB* polymorphisms with fibrinogen-related phenotypes

	FGB-rs7439150$^{\delta}$	FGB-rs1800789$^{\delta}$	FGB-rs1800790	FGB-rs4463047
SNP pseudonym		***FGB*-1420 G>A**	***FGB*-455 G>A**	
Minor allele frequency (%)	A = 6.95	A = 6.78	A = 3.32	T = 11.4
Genotype groups (n)	GG[1] (1548) GA/AA[2] (215)	GG[1] (1513) GA/AA[2] (292)	GG[1] (1652) GA/AA[2] (108)	CC[1] (1424) CT/TT[2] (357)
Total fibrinogen (g/L)	3.66 ± 2.13 3.97 ± 2.29	3.64 ± 2.12 * 3.96 ± 2.29 *	3.66 ± 2.12 4.05 ± 2.34	3.71 ± 2.18 3.52 ± 1.99
Fibrinogen γ' (mg/mL)	0.37 ± 0.25* 0.41 ± 0.32*	0.38 ± 0.25 0.42 ± 0.33	0.38 ± 0.26 * 0.46 ± 0.37 *	0.38 ± 0.27 0.37 ± 0.22
Fibrinogen γ' (%)	12.1 ± 8.07 11.6 ± 8.82	12.3 ± 8.27 11.6 ± 7.80	12.1 ± 8.01 13.2 ± 10.1	12.1 ± 8.27 12.2 ± 7.90
Lag time (min)	6.47 ± 1.97 6.59 ± 1.96	6.47 ± 1.96 6.55 ± 1.99	6.48 ± 1.99 6.49 ± 1.98	6.47 ± 1.97 6.52 ± 1.98
Slope ($x10^{-3}$ au/s)	9.58 ± 4.42 10.0 ± 3.99	9.54 ± 4.35 10.3 ± 4.09	9.65 ± 4.36 9.75 ± 3.95	9.67 ± 4.41 9.60 ± 4.32
Maximum absorbance (nm)	0.43 ± 1.59 * 0.46 ± 0.17 *	0.43 ± 0.16 * 0.46 ± 0.17 *	0.43 ± 0.16 0.46 ± 0.19	0.43 ± 0.16 0.44 ± 0.16
Clot lysis time (min)	56.9 ± 11.2 58.0 ± 12.0	56.8 ± 11.2 *# 58.3 ± 12.1 *#	57.0 ± 11.3 58.7 ± 11.1	57.3 ± 11.3 56.3 ± 11.3

A = adenine; C = cytosine; *FGB* = fibrinogen beta chain gene; G = guanine; rs = reference sequence; T = thymine; Lag time = time required for the activation of the coagulation cascade by TF and for protofibrils to reach sufficient length to allow lateral aggregation; Slope = rate of lateral aggregation of fibrin protofibrils; Maximum absorbance = indicator of fibre diameter.

Data presented as mean ± SD; group 1 (major allele homozygotes) first and group 2 (minor allele carriers) below.

*p < 0.05; # p < 0.05 upon adjustment for total and γ' fibrinogen; $^{\delta}$ Strong evidence of LD (r^2 = 0.89; D' = 0.95).

Age, gender, BMI, HIV-status, $HbA1_c$ and HDL-c were confounders. Associations with CLT were adjusted for $PA\text{-}I1_{act}$ additionally.

The association of IL-6 with fibrinogen-related phenotypes

The basic descriptive characteristics of the study population have been reported in a previous publication (Kotzé *et al,* 2015). Descriptive statistics for fibrinogen and clot properties, including their association with IL-6 (presented as quartiles), are shown in Table 3.2. The mean IL-6 concentration was 6.50 ± 21.0 pg/mL. Total and γ' fibrinogen concentrations were positively, although not linearly, associated with IL-6 with significantly higher concentrations in the fourth IL-6 quartile compared to the first three quartiles. The γ' fibrinogen ratio did not reach significance (p = 0.55), implying that the fibrinogen γ' association is probably a reflection of the association of IL-6 with total fibrinogen. Positive associations were also observed for lag time, slope and maximum absorbance, although subsequent adjustment for total and γ' fibrinogen led to loss of significance. CLT decreased as IL-6 increased, with total fibrinogen concentration confounding this effect. Adjustments for total and γ' fibrinogen, and thereafter PAI-1_{act} (as a main modulator of CLT) significantly increased this negative association (Figure 3.3).

Table 3.2 Outcome phenotypes descriptive statistics and association with IL-6- quartiles

Variable	Whole group	Interleukin-6 quartiles (pg/mL)				p-value unadjusted (adjusted)
		Quartile 1 < 0.76	**Quartile 2** 1.50 – 2.84	**Quartile 3** 2.85 – 5.75	**Quartile 4** 5.76 – 424	
Total fibrinogen (g/L)	3.69 ± 2.18	3.02 ± 1.62 [ab]	3.32 ± 1.86 [cd]	3.95 ± 2.17 [ace]	4.53 ± 2.66 [bde]	<0.001
Fibrinogen γ' (mg/L)	0.38 ± 0.27	0.32 ± 0.22 [ab]	0.36 ± 0.25 [c]	0.39 ± 0.21 [ad]	0.46 ± 0.37 [bcd]	<0.001
Fibrinogen γ' (%)	12.1 ± 8.25	12.5 ± 8.56	12.4 ± 7.58	11.8 ± 7.77	11.8 ± 8.94	0.552
Lag time (min)	6.46 ± 1.97	6.24 ± 1.99 [a]	6.46 ± 2.03	6.64 ± 1.93 [a]	6.38 ± 1.94	0.033 (0.126)
Slope ($x10^{-3}$ au/s)	9.70 ± 4.42	9.19 ± 3.80 [a]	9.60 ± 4.37	9.86 ± 4.30	10.5 ± 5.15 [a]	<0.001 (0.897)
Maximum absorbance (nm)	0.43 ± 0.16	0.41 ± 0.14 [ab]	0.43 ± 0.15	0.45 ± 0.15 [a]	0.45 ± 0.18 [b]	<0.001 (0.183)
CLT (min)	57.3 ± 11.2	58.3 ± 10.8 [1]	56.9 ± 10.9	56.8 ± 11.5	56.2 ± 11.9 [1]	0.066 (0.002)

Data presented as mean ± SD

Adjusted p-value = adjusted for total and γ' fibrinogen; CLT = clot lysis time; Lag time = time required for the activation of the coagulation cascade by TF and for protofibrils to reach sufficient length to allow lateral aggregation; Slope = rate of lateral aggregation of fibrin protofibrils; Maximum absorbance = indicator of fibre diameter. [abcde] Means with the same symbol differ significantly for individual outcome variables; [1] Means with the same numerical value differ significantly for individual outcome variables upon adjustment.

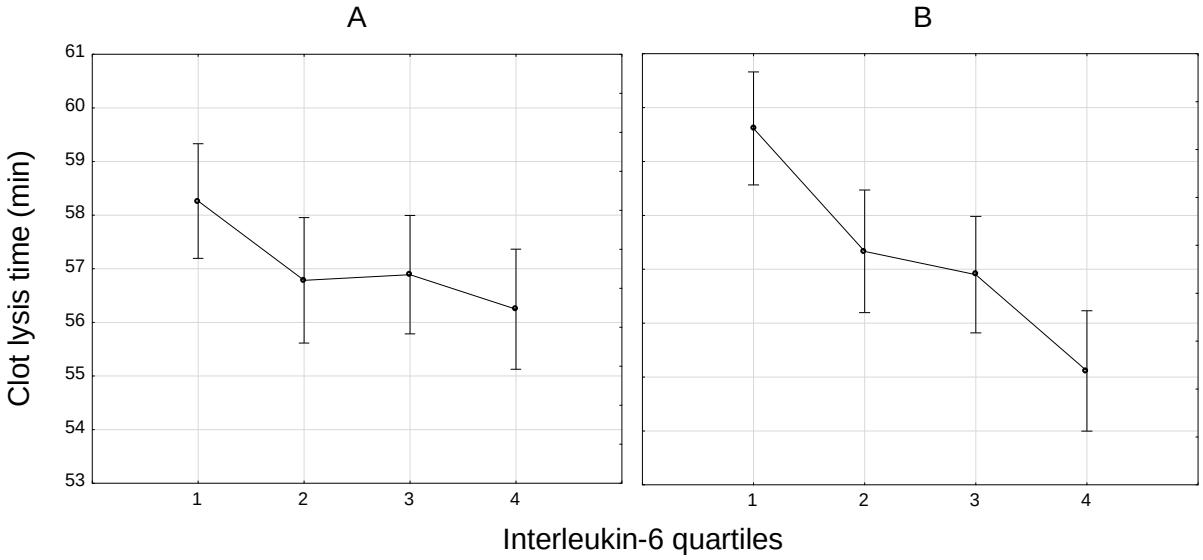


Figure 3.3 CLT across IL-6 quartiles before (A) and after (B) adjustment for total fibrinogen, γ' fibrinogen and PAI-1_{act}

Vertical bars denote 95% confidence interval. P-values: 0.066 (A); <0.00001 (B).
Images are equally scaled

Genotype-IL-6 interactions in terms of phenotype predictions

Interaction analyses were performed for each polymorphism with IL-6 in relation to each of the fibrinogen measures and clot properties (Table 3.3 and 3.4). Three significant SNP-IL-6-interactions were observed in predicting γ' fibrinogen, one of which (*FGA*-rs2070011) showed a corresponding effect with total fibrinogen concentrations. IL-6 concentration was associated with a steeper increase in fibrinogen concentration in minor allele carriers of six SNPs and a shallower increase in total fibrinogen in *FGA*-rs6050G. For *FGG*-rs2066865, a shallower increase was observed in the presence of one minor allele, with a steeper increase in the presence of two minor alleles, more so than for the wild type. Six SNPs interacted with IL-6 in determining fibre diameter (maximum absorbance), in which the presence of the minor allele resulted in a stronger positive association between IL-6 and maximum absorbance than the respective major alleles. Adjustment for fibrinogen concentration largely nullified these associations, with only two associations remaining significant (*FGB* - rs7439150 and -148C/T). *FGB* -1420G/A, which is in high LD with *FGB* - rs7439150 and -148C/T, however, lost its significance after adjustment for fibrinogen owing to its independent association with fibrinogen concentration. No significant interactions were observed for lag time, slope or CLT. Removal of individuals with IL-6 concentrations higher than 100 pg/mL (n = 8) resulted in a loss of significance for many of the interactions presented in Table 3.3 and 3.4. The remaining IL-6 interactions were with *FGG*-rs1049636 in terms of fibrinogen concentration and with *FGB* - rs7439150, -455G/A, -148C/T in terms of γ' fibrinogen (data reported in Table 3.8).

Table 3.3 Genotype-IL-6 interactions modulating total and γ' fibrinogen concentrations

Phenotype	Gene	IL-6 interaction with SNP	Interaction p-value		Major allele homozygotes		Heterozygotes or minor allele carriers (where only two groups)		Minor allele homozygotes (where three groups)	
					N	Slope (95% CI)	N	Slope (95% CI)	N	Slope (95% CI)
Total fibrinogen (g/L)	*FGB*	rs7439150[δ]	0.001	GG	1293	0.015 (0.009 – 0.020)	174	0.072 (0.040 – 0.104)		
	FGB	rs1800789[δ]	<0.001	GG	1263	0.012 (0.006 – 0.018)	240	0.071 (0.050 – 0.093)		
	FGB	rs1800788	<0.001	CC	1419	0.017 (0.012 – 0.023)	154	0.083 (0.045 – 0.121)		
	FGB	rs1800787[δ]	0.001	CC	1164	0.016 (0.010 – 0.022)	152	0.077 (0.043 – 0.111)		
	FGB	rs4220	0.003	GG	1121	0.016 (0.010 – 0.022)	224	0.061 (0.035 – 0.087)		
	FGA	rs6050	<0.001	AA	676	0.039 (0.028 – 0.049)	526	0.016 (0.006 – 0.027)	132	1×10^{-4}(-0.010 – 0.010)
	FGA	rs2070011	0.010	GG	925	0.015 (0.009 – 0.021)	416	0.042 (0.024 – 0.060)		
	FGG	rs2066865	<0.001	CC	746	0.040 (0.029 – 0.050)	482	0.005 (-0.003 – 0.012)	108	0.111 (0.067 – 0.156)
	FGG	rs1049636	<0.001	TT	975	0.014 (0.008 – 0.020)	371	0.110 (0.082 – 0.138)		
Fibrinogen γ' ratio (%)	*FGB*	rs1800790	0.010	GG	1309	0.014 (-0.009 – 0.037)	93	0.242 (0.051 – 0.433)		
	FGB	rs4463047	0.005	CC	1139	0.007 (-0.017 – 0.031)	283	0.134 (0.060 – 0.207)		
	FGA	rs2070011	0.008	GG	890	0.010 (-0.017 – 0.038)	400	0.128 (0.057 – 0.199)		

A = adenine; C = cytosine; CI = confidence interval; G = guanine; IL-6 = interleukin-6; rs = reference sequence; T = thymine; rs1800789 = -1420G/A; rs1800788 = -249C/T; rs1800787 = -148C/T; rs1800790 = -455G/A; rs2070011 = 2224G/A, rs1049636 = 9340T/C; δVariants in high linkage disequilibrium ($r^2 > 0.82$; D' > 0.92)

Table 3.4 Genotype-IL-6 interactions modulating clot properties

Phenotype	Gene	IL-6 interaction with SNP	Interaction p-value			Major allele homozygotes		Heterozygotes or minor allele carriers (where only two groups)		Minor allele homozygotes (where three groups)	
			Unadjusted	Adjusted		N	Slope (95% CI)	N	Slope (95% CI)	N	Slope (95% CI)
Maximum absorbance (nm)	*FGB*	rs7439150 $^{\delta}$	<0.001	0.01	GG	1274	$1x10^{-4}$ ($-3x10^{-4}$ – 0.001)	179	0.003 (0.002 – 0.003)		
	FGB	rs1800789 $^{\delta}$	<0.001	0.82	GG	1245	$2x10^{-4}$ ($-3x10^{-4}$ – 0.001)	241	0.002 (0.002 – 0.003)		
	FGB	rs1800788	0.001	0.04	CC	1407	0.001 ($2x10^{-4}$ – 0.001)	153	0.004 (0.002 – 0.007)		
	FGB	rs1800787 $^{\delta}$	<0.001	0.01	CC	1157	$1x10^{-4}$ ($-3x10^{-4}$ – 0.001)	157	0.003 (0.002 – 0.003)		
	FGA	rs6050	0.001	0.04	AA	663	$-1x10^{-5}$ (-0.001 – 0.001)	531	0.002 (0.001 – 0.002)	133	$1x10^{-5}$(-0.001 – 0.001)
	FGG	rs1049636	0.003	0.23	TT	970	0.001 ($3x10^{-4}$ – 0.001)	369	0.003 (0.001 – 0.005)		

A = adenine; C = cytosine; CI = confidence interval; G = guanine; IL-6 = interleukin-6; rs = reference sequence; T = thymine; rs1800789 = -1420G/A; rs1800788 = -249C/T; rs1800787 = -148C/T; rs1049636 = 9340T/C

Adjusted = adjusted for total fibrinogen (g/L); Slope = rate of lateral aggregation of fibrin protofibrils; Maximum absorbance = indicator of fibre diameter.

δVariants in high linkage disequilibrium ($r^2 > 0.82$; D' > 0.92)

In addition to the individual interaction analyses, the seven variants resulting in a steeper fibrinogen increase in the presence of IL-6 (Table 3.3) were grouped in a genetic 'risk score' to determine whether there were additive effects when these genotypes occurred together in an individual. Each individual's score was composed of the sum of the seven polymorphisms allocated a value of either one (major allele homozygote) or two (minor allele carrier). Final scores ranked from 7 to 13, therefore scores represented the presence of zero to six minor allele groups. The risk score showed a significant interaction with IL-6 ($p < 0.001$) in the prediction of fibrinogen concentrations. Figure 3.4 schematically represents the slope of each score in the relationship of IL-6 and fibrinogen. The addition of a minor allele group to the 'risk score' increased the slope of the IL-6-fibrinogen association. The 12th and 13th risk score were removed from Figure 3.4 as the sizes of these groups were too small to provide adequate power (n = 24 and 5, respectively). The interaction, however, remained significant upon removal of these groups ($p < 0.001$). The regression coefficient for the line indicating the risk score of 8 differed significantly from the score of 7 ($p = 0.04$), with 9, 10 and 11 differing from 8 ($p = 0.01$, 0.048, 0.005, respectively) but not from each other, indicating a possible threshold to the additive effect when more than three risk alleles occurred together.

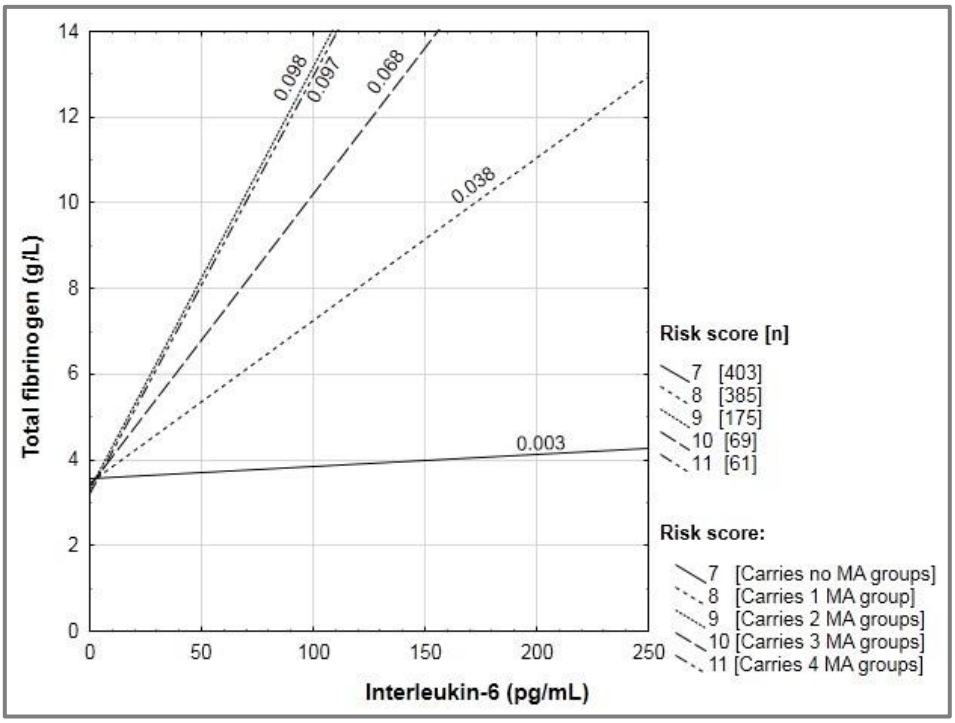


Figure 3.4 Association of total fibrinogen concentrations with circulating interleukin-6 by minor allele (MA) risk score groups

[ ] indicates the number of minor allele (MA) groups predicting the score.

DISCUSSION

High LD in the fibrinogen gene cluster in Europeans continues to hinder the identification of functional polymorphisms contributing to fibrinogen concentration and functionality (specifically fibrin clot properties). This becomes even more complex when taking the role of transcriptional enhancers, such as IL-6, into consideration. Although little is known about the fibrinogen genotype in Africans, preliminary evidence suggests distinct genetic differences from that of Europeans, including higher recombination rates and polymorphic variance. In addition, this population is known for altered fibrinogen and IL-6 phenotypes. This study therefore aimed to identify possible functional fibrinogen SNPs and novel IL-6-interactions related to total and γ' fibrinogen, as well as indicators of clot formation, structure and lysis by using the unique African genetic and phenotypic profile. Our data revealed that none of the common European fibrinogen haplotypes is present in this African population. We demonstrate that in a population with chronic low-grade inflammation, IL-6 interacted with several of the fibrinogen SNPs to influence fibrinogen concentration and even resulted in altered fibrin clot properties. A novel finding with important implications for the causal role of fibrinogen in CVD was that IL-6 not only interacted with individual SNPs, but that an additive effect was also observed when harbouring more than one risk allele concurrently.

The MAF of 12 of the 14 investigated SNPs was significantly lower in this Tswana population than that reported globally (Yates *et al,* 2016). No variation has been reported for the *FGB*-rs2227385 and *FGB*-rs2227388 SNPs outside Africa (Yates *et al,* 2016). As predicted, a high recombination rate was observed in this study population, resulting in no complete LD among any of the SNPs. However, strong LD between *FGB* -1420G/A and -148C/T was observed, similar to the finding of Dehghan *et al,* (2009), although the LD observed for *FGB*-rs7439150 with these SNPs is novel. This lack of full LD differs from numerous reports of almost complete LD between the -1420G/A, -933C/T, -455G/A and -148C/T SNPs in Europeans (Baumann and Henschen 1994; Behague *et al,* 1996; Green 2001; Morozumi *et al,* 2009; Thomas *et al,* 1994; van't Hooft *et al,* 1999; Verschuur *et al,* 2005). None of the frequently occurring haplotypes, consisting of -1420G/A, -933C/T, -854G/A, -455G/A, -249C/T and -148C/T, (Green 2001; Morozumi *et al,* 2009; Verschuur *et al,* 2005), was present in this study population, highlighting the genetic diversity in Africans. In addition, *FGB*-933C/T revealed no genetic variation in this population.

In the present study, *FGB* -854A, -1420A and *FGG*-rs1049636C were associated with higher total fibrinogen concentrations. Higher γ' fibrinogen concentrations (but not γ' ratio) were observed in the presence of *FGB* -455A, rs7439150A and *FGA*-rs2070011A. Furthermore, a

positive association between *FGB* - 148T and fibre diameter, and *FGB* - 1420A and CLT was observed. IL-6 correlated positively with total and γ' fibrinogen, thereby accelerating clot formation and increasing fibre diameter. CLT was, however, negatively associated with IL-6, and both total and γ' fibrinogen concentration confounded this association. Five *FGB*, one *FGA* and one *FGG* SNP significantly interacted with circulating IL-6 to enhance the IL-6-induced expression of total fibrinogen effectively. These interactions proved to be additive, with the presence of more than one minor allele across the gene cluster resulting in greater increases in IL-6-induced fibrinogen expression, although a threshold was reached when more than three minor allele variants occurred together in an individual. Lastly, these interactive associations reflected functional effects in terms of the rate of lateral aggregation and fibre diameter, thus indicating a possible mechanism by which the fibrinogen SNPs, during the acute phase, could enhance thrombotic risk.

Individual SNP associations for ten of the polymorphisms investigated have been reported previously (Kotzé *et al,* 2015). The associations of the four additionally genotyped SNPs, although not all reaching significance, are in agreement with the existing literature. Increased fibrinogen concentrations in the presence of *FGB*-rs7439150A, - 1420A and -455A have been reported in case-control and genome-wide association analyses (Behague *et al,* 1996; Brown and Fuller 1998; Carty *et al,* 2010; Danik *et al,* 2009; de Vries *et al,* 2016; Dehghan *et al,* 2009; Kathiresan *et al,* 2006; Ken-Dror *et al,* 2012; Klovaite *et al,* 2013; Sabater-Lleal *et al,* 2013; Van Der Bom *et al,* 1998). The lack of association between rs4463047 and total fibrinogen concentrations is suggestive of it being a non-functional SNP, as the only two association studies to date reported inconsistent associations: positive (Sabater-Lleal *et al,* 2013) and negative (Wassel *et al,* 2011), probably reflecting differences in LD with a functional SNP elsewhere. The novel association of -1420A with faster CLT, independent of fibrinogen concentration, deserves further investigation.

Most of the research into both the heritability of and variant associations with fibrinogen has been conducted in European populations. The lack of LD, and common haplotypes alongside the inability to reproduce results of independent SNP associations in this current investigation, shows that heritability estimates obtained in Europeans cannot be extrapolated to Africans. This necessitates the replication of heritability studies in African ethnic sub-groups. The 14 SNPs investigated here contributed a mere 0.5% to the variance in total fibrinogen.

Total fibrinogen and IL-6 concentrations were generally higher than those reported in Europeans (Carty *et al,* 2010; Jacquemin *et al,* 2008). IL-6 correlated positively with fibrinogen concentration, thereby accelerating clot formation and increasing fibre diameter, consequently

contributing to clot pathology and CVD (Undas and Ariëns 2011). In the present study, higher IL-6 was associated with a faster CLT, independent of fibrinogen. This is probably the result of increases in other biomarkers influencing CLT, such as plasminogen concentrations leading to faster clot lysis. Previous reports have positively associated IL-6 with increased plasminogen transcription (Jenkins *et al,* 1997; Kida *et al,* 1997).

Fibrinogen γ' was also positively associated with IL-6, although the non-significant association of IL-6 with the total to γ' fibrinogen ratio revealed the observation to be largely a reflection of the relationship with total fibrinogen. This finding suggests that in this study population, characterised by chronic low-grade inflammation, the influence of IL-6 on total and γ' fibrinogen is probably due to up-regulation of the entire fibrinogen gene cluster, while alterations in the *FGG* alternative splicing mechanism may be more relevant in pronounced inflammatory conditions such as have been reported for CVD (Alexander *et al,* 2011; Cheung *et al,* 2009).

IL-6-induced fibrinogen gene expression can be altered by polymorphic variation (Morozumi *et al,* 2009; Verschuur *et al,* 2005), therefore it is important to take genotype-IL-6 interactions into consideration when determining the influence of genetic variance on the fibrinogen phenotype, particularly in populations with prevalent inflammatory conditions. Nine SNPs interacted significantly with IL-6 in predicting fibrinogen concentrations and three SNPs with γ' fibrinogen, highlighting the importance of considering IL-6 levels when investigating the influence of genetic variance on fibrinogen. These interactions furthermore led to altered fibrin clot properties, in particular increased fibre diameter, which has been associated with CVD (Mills *et al,* 2002). The number of significant genotype-IL-6 interactions was reduced when individuals with high IL-6 concentrations (>100 pg/mL) were removed from analyses, suggesting that these interactions are physiologically more relevant in the presence of high IL-6 plasma concentrations. The interactions with *FGB* - rs7439150, -455G/A, -148C/T and *FGG*-rs1049636 remained however, indicating their potential relevance to fibrinogen regulation during chronic low-grade inflammatory conditions thought to be associated with increased CAD (Adukauskienė *et al*, 2016). In addition, these interactions had an additive effect that could only be detected because of the lack of LD in the fibrinogen gene cluster in this African study population. The additive effect reached a threshold when more than three risk alleles occurred together. This observation should be investigated in a larger study population to allow analysis of the number of risk alleles, rather than just the presence thereof, thereby exploring dominant/co-dominant/recessive effects. Nonetheless, the observed additive effects reflect true physiological mechanisms, as the particular combinations are harboured concurrently in each individual. This data are in agreement with the findings of Ken-Dror *et al,* (2012) who suggested that the

fibrinogen phenotype is not regulated by one functional SNP only, but by a combination of minor alleles spanning the whole cluster. The lack of a causal contribution of fibrinogen to CVD by Mendelian randomisation studies should, in our opinion, therefore, be interpreted with caution, as these studies investigated single variants only and focussed on fibrinogen concentration alone while omitting qualitative effects such as altered clot properties (De Moerloose *et al,* 2010).

In this Tswana population, of the 14 SNPs investigated, FGB-rs7439150, -1420G/A and -148C/T had the most pronounced effects on the fibrinogen phenotype in terms of both concentration and altered clot properties, with several of these associations altered by IL-6 concentrations. *FGB*-rs7439150 is located in a regulatory feature (ENSR00000175110) spanning 623 base pairs (4:154560017-154560640), and is predicted to have methylation potential. This feature could be the transcriptional enhancer that is essential to the functional role of *FGB*-rs7439150 portrayed in this study (Zerbino *et al*, 2016). In addition, possible functionality of -148C/T has been described by Verschuur *et al*, (2005) owing to its ability to alter fibrinogen's response to IL-6 by interfering with the hepatocyte nuclear factor 3 and CCAAT enhancer binding protein binding sites. Although the association of the *FGB* promoter -1420G/A polymorphism with fibrinogen concentration has been reported before (Dehghan *et al,* 2009; Sabater-Lleal *et al,* 2013), it is not associated with any known regulatory features (Zerbino *et al*, 2016) and further investigation into its functionality is required. Despite the higher genetic recombination rates in Africans, these three SNPs were however found to be in high, albeit not complete LD ($D' > 0.92$, $r^2 > 0.82$), therefore the possibility of inferred functionality can also not be excluded. These variants did, however, have unrelated independent genotype-phenotype associations; *FGB*-rs7439150 with fibrinogen γ, *FGB* -1420G/A with total fibrinogen, and *FGB* -148C/T with maximum absorbance, respectively. Distinct differences in their IL-6-interactive associations (Tables 3.3, 3.4 and 3.8) were also observed, strengthening their hypothesised independent functional contribution to the fibrinogen-related phenotypes.

In conclusion, the genotypic exploration of the fibrinogen phenotype in the Tswana population proved the unique genetic composition of black South Africans. It is clear that international research cannot be extrapolated to Africans and that the risk profile in terms of the relative contribution of genetics and environmental factors might differ. The results demonstrate that IL-6 can modulate the influence of fibrinogen SNPs on fibrinogen concentration individually and in an additive manner, with the presence of more minor alleles across these genes leading to greater increases in IL-6-induced fibrinogen expression in an apparently healthy study population. Therefore, when investigating the effect of fibrinogen genetics on fibrinogen

concentrations and CVD outcome, the possible interactions with modulating factors and the fact that SNP effects seem to be additive should be taken into account.

ACKNOWLEDGEMENTS

Grants from the North-West University (NWU), South African National Research Foundation, Newton Fund, Population Health Research Institute, Medical Research Council, the North West Province Health Department, DNAbiotec Pty Ltd, and the South African Netherlands Partnerships in Development supported the PURE study and the subsequent biochemical analyses. Funding sources were not involved in the conceptualisation, collection, interpretation or writing of the data in any way.

We thank all participants, fieldworkers and supporting staff involved in the PURE data collection. We thank the PURE-SA research team, especially Professor Annamarie Kruger and the office staff of the Africa Unit for Transdisciplinary Health Research, the Faculty of Health Sciences, NWU (Potchefstroom Campus), South Africa. In addition, we thank the PURE-International research team, especially Doctor Yusuf and the PURE-study office staff at the PHRI, Hamilton Health Sciences and McMaster University, ON, Canada. We thank the staff of the Profiles of Resistance to Insulin in Multiple Ethnicities and Regions study from the Centre for Genome Research and DNAbiotec Pty Ltd. Lastly, we would also like to express our sincerest gratitude to Mrs Marike Cockeran from the Medicine Usage in South Africa research unit of the NWU (Potchefstroom Campus), South Africa, for her guidance with the statistical analyses.

AUTHOR CONTRIBUTIONS

HTC isolated the DNA (second round), performed KASP analyses, performed the statistical analyses and wrote the manuscript; CN-R isolated the DNA (first round), performed RFLP analyses and critically reviewed and approved the final manuscript; FRG conceptualised the manuscript with MP, discussed the analyses and their implications and critically reviewed and approved the final manuscript; ZdL performed the global fibrinolytic assay and critically reviewed and approved the final manuscript; LZ designed the KASP primers, assisted with quality control of the KASP data and critically reviewed and approved the final manuscript; MP conceptualised the manuscript, supervised the laboratory analyses of all the haemostatic variables and wrote the manuscript with HTC. The authors have no conflict of interest to declare.

SUPPORTING INFORMATION

Supporting information can be found in the online version of this manuscript:

Table 3.5: Primer and synthetic control sequences used for KASP analyses

Table 3.6: Associations of individual SNPs with fibrinogen γ′, γ′ ratio and total fibrinogen as published by Kotzé *et al,* (2015)

Table 3.7: Associations of individual SNPs with clot-related phenotypes as published by Kotzé *et al,* (2015)

Table 3.8: Genotype-IL-6 interactions modulating total and γ' fibrinogen concentrations upon removal of individuals with IL-6 > 100pg/mL

REFERENCES

Adukauskienė, D., Čiginskienė, A., Adukauskaitė, A., Pentiokinienė, D., Šlapikas, R. & Čeponienė, I. (2016) Clinical relevance of high sensitivity C-reactive protein in cardiology. *Medicina*, **52**, 1-10.

Alexander, K.S., Madden, T.E. & Farrell, D.H. (2011) Association between γ' fibrinogen levels and inflammation. *Thrombosis and Haemostasis,* **105**, 605-609.

Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics,* **21**, 263-265.

Baumann, R.E. & Henschen, A.H. (1994) Linkage disequilibrium relationships among four polymorphisms within the human fibrinogen gene cluster. *Human Genetics,* **94**, 165-170.

Behague, I., Poirier, O., Nicaud, V., Evans, A., Arveiler, D., Luc, G., Cambou, J.P., Scarabin, P.Y., Bara, L. & Green, F. (1996) β Fibrinogen gene polymorphisms are associated with plasma fibrinogen and coronary artery disease in patients with myocardial infarction The ECTIM Study. *Circulation,* **93**, 440-449.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)***, 57**, 289-300.

Bester, J. & Pretorius, E. (2016) Effects of IL-1β, IL-6 and IL-8 on erythrocytes, platelets and clot viscoelasticity. *Scientific Reports,* **6**, 32188.

Brown, E.T. & Fuller, G.M. (1998) Detection of a complex that associates with the Bβ fibrinogen G−455-A polymorphism. *Blood,* **92**, 3286-3293.

Brull, D., Dhamrait, S., Moulding, R., Rumley, A., Lowe, G., Humphries, S. & Montgomery, H. (2002) The effect of fibrinogen genotype on fibrinogen levels after strenuous physical exercise. *Thrombosis and Haemostasis,* **87**, 37-41.

Carty, C.L., Heagerty, P., Heckbert, S.R., Jarvik, G.P., Lange, L.A., Cushman, M., Tracy, R.P. & Reiner, A.P. (2010) Interaction between fibrinogen and IL-6 genetic variants and associations with cardiovascular disease risk in the cardiovascular health study. *Annals of Human Genetics,* **74**, 1-10.

Chen, Y.S., Torroni, A., Excoffier, L., Santachiara-Benerecetti, A.S. & Wallace, D.C. (1995) Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *American Journal of Human Genetics,* **57**, 133-149.

Cheung, E.Y., Uitte de Willige, S., Vos, H.L., Leebeek, F.W., Dippel, D.W., Bertina, R.M. & de Maat, M.P. (2008) Fibrinogen γ' in ischemic stroke: a case-control study. *Stroke,* **39**, 1033-1035.

Cheung, E.Y., Vos, H.L., Kruip, M.J., den Hertog, H.M., Jukema, J.W. & de Maat, M.P. (2009) Elevated fibrinogen γ′ ratio is associated with cardiovascular diseases and acute phase reaction but not with clinical outcome. *Blood,* **114**, 4603-4604.

Cook, D.G., Cappuccio, F.P., Atkinson, R.W., Wicks, P.D., Chitolie, A., Nakandakare, E.R., Sagnella, G.A. & Humphries, S.E. (2001) Ethnic differences in fibrinogen levels: the role of environmental factors and the β-fibrinogen gene. *American Journal of Epidemiology,* **153**, 799-806.

Cotton, J., Webb, K., Mathur, A., Martin, J. & Humphries, S. (2000) Impact of the-455G>A promoter polymorphism in the β fibrinogen gene on stimulated fibrinogen production following bypass surgery. *Thrombosis and Haemostasis,* **84,** 926-927.

Danesh, J., Lewington, S., Thompson, S.G., Lowe, G., Collins, R., Kostis, J., Wilson, A., Folsom, A., Wu, K. & Benderly, M. (2005) Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of the American Medical Association,* **294**, 1799-1809.

Danik, J.S., Paré, G., Chasman, D.I., Zee, R.Y., Kwiatkowski, D.J., Parker, A., Miletich, J.P. & Ridker, P.M. (2009) Novel loci, including those related to Crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17 686 women: The women's genome health study. *Circulation: Cardiovascular Genetics,* **2**, 134-141.

Davalos, D. & Akassoglou, K. (2012) Fibrinogen as a key regulator of inflammation in disease. *Seminars in Immunopathology*, **34**, 43-62.

De Lange, M., Snieder, H., Ariëns, R.A.S., Spector, T.D. & Grant, P.J. (2001) The genetics of haemostasis: a twin study. *Lancet,* **357**, 101-105.

De Lange, Z., Rijken, D.C., Hoekstra, T., Conradie, K.R., Jerling, J.C. & Pieters, M. (2013) In black South Africans from rural and urban communities, the 4G/5G PAI-1 polymorphism influences PAI-1 activity, but not plasma clot lysis time. *PLoS One,* **8**, e83151.

De Moerloose, P., Boehlen, F. & Neerman-Arbez, M. (2010) Fibrinogen and the risk of thrombosis. *Seminars in Thrombosis and Hemostasis*, **36**, 7-17.

De Vries, P.S. & Chasman, D.I. & Sabater-Lleal, M. & Chen, M.H. & Huffman, J.E. & Steri, M. & Tang, W. & Teumer, A. & Marioni, R.E. & Grossmann, V. & Hottenga, J.J. & Trompet, S. & Muller-Nurasyid, M. & Zhao, J.H. & Brody, J.A. & Kleber, M.E. & Guo, X. & Wang, J.J. & Auer, P.L. & Attia, J.R. & Yanek, L.R. & Ahluwalia, T.S. & Lahti, J. & Venturini, C. & Tanaka, T. & Bielak, L.F. & Joshi, P.K. & Rocanin-Arjo, A. & Kolcic, I. & Navarro, P. & Rose, L.M. & Oldmeadow, C. & Riess, H. & Mazur, J. & Basu, S. & Goel, A. & Yang, Q. & Ghanbari, M. & Willemsen, G. & Rumley, A. & Fiorillo, E. & de Craen, A.J. & Grotevendt, A. & Scott, R. & Taylor, K.D. & Delgado, G.E. & Yao, J. & Kifley, A. & Kooperberg, C. & Qayyum, R. & Lopez, L.M. & Berentzen, T.L. & Raikkonen, K. & Mangino, M. & Bandinelli, S. & Peyser, P.A. & Wild, S. & Tregouet, D.A. & Wright, A.F. & Marten, J. & Zemunik, T. & Morrison, A.C. & Sennblad, B. & Tofler, G. & de Maat, M.P. & de Geus, E.J. & Lowe, G.D. & Zoledziewska, M. & Sattar, N. & Binder, H. & Volker, U. & Waldenberger, M. & Khaw, K.T. & McKnight, B. & Huang, J. & Jenny, N.S. & Holliday, E.G. & Qi, L. & McEvoy, M.G. & Becker, D.M. & Starr, J.M. & Sarin, A.P. & Hysi, P.G. & Hernandez, D.G. & Jhun, M.A. & Campbell, H. & Hamsten, A. & Rivadeneira, F. & McArdle, W.L. & Slagboom, P.E. & Zeller, T. & Koenig, W. & Psaty, B.M. & Haritunians, T. & Liu, J. & Palotie, A. & Uitterlinden, A.G. & Stott, D.J. & Hofman, A. & Franco, O.H. & Polasek, O. & Rudan, I. & Morange, P.E. & Wilson, J.F. & Kardia, S.L. & Ferrucci, L. & Spector, T.D. & Eriksson, J.G. & Hansen, T. & Deary, I.J. & Becker, L.C. & Scott, R.J. &

Mitchell, P. & Marz, W. & Wareham, N.J. & Peters, A. & Greinacher, A. & Wild, P.S. & Jukema, J.W. & Boomsma, D.I. & Hayward, C. & Cucca, F. & Tracy, R. & Watkins, H. & Reiner, A.P. & Folsom, A.R. & Ridker, P.M. & O'Donnell, C.J. & Smith, N.L. & Strachan, D.P. & Dehghan, A. (2016) A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. *Human Molecular Genetics,* **25**, 358-370.

Dehghan, A., Yang, Q., Peters, A., Basu, S., Bis, J.C., Rudnicka, A.R., Kavousi, M., Chen, M.H., Baumert, J. & Lowe, G.D. (2009) Association of novel genetic loci with circulating fibrinogen levels a genome-wide association study in 6 population-based cohorts. *Circulation: Cardiovascular Genetics,* **2**, 125-133.

Drouet, L., Paolucci, F., Pasqualini, N., Laprade, M., Ripoll, L., Mazoyer, E., Bal dit Sollier, C. & Vanhove, N. (1999) Plasma γ'/γ fibrinogen ratio, a marker of arterial thrombotic activity: a new potential cardiovascular risk factor? *Blood Coagulation and Fibrinolysis,* **10**, S35-39.

Duan, H.O. & Simpson-Haidaris, P.J. (2003) Functional analysis of interleukin 6 response elements (IL-6REs) on the human γ-fibrinogen promoter binding of hepatic stat3 correlates negatively with transactivation potential of Type II IL-6REs. *Journal of Biological Chemistry,* **278**, 41270-41281.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007) G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods,* **39,** 175-191.

Fish, R.J. & Neerman-Arbez, M. (2012) Fibrinogen gene regulation. *Thrombosis and Haemostasis,* **108**, 419-426.

Fuller, G.M. & Zhang, Z. (2001) Transcriptional control mechanism of fibrinogen gene expression. *Annals of the New York Academy of Science,* **936**, 469-479.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A. & Faggart, M. (2002) The structure of haplotype blocks in the human genome. *Science,* **296**, 2225-2229.

Green, F. (2001) Fibrinogen polymorphisms and atherothrombotic disease. *Annals of the New York Academy of Science,* **936**, 549-559.

Hamsten, A., De Faire, U., Iselius, L. & Blombäck, M. (1987) Genetic and cultural inheritance of plasma fibrinogen concentration. *The Lancet,* **330**, 988-991.

Hardy, G.H. (1908) Mendelian proportions in a mixed population. *Molecular and General Genetics,* **1**, 395-395.

Hinds, D.A., Buil, A., Ziemek, D., Martinez-Perez, A., Malik, R., Folkersen, L., Germain, M., Mälarstig, A., Brown, A. & Soria, J.M. (2016) Genome-wide association analysis of self-reported events in 6135 individuals and 252 827 controls identifies 8 loci associated with thrombosis. *Human Molecular Genetics*, **25**, 1867-1874.

Jacquemin, B., Antoniades, C., Nyberg, F., Plana, E., Müller, M., Greven, S., Salomaa, V., Sunyer, J., Bellander, T. & Chalamandaris, A.-G. (2008) Common genetic polymorphisms and haplotypes of fibrinogen alpha, beta, and gamma chains affect fibrinogen levels and the response to proinflammatory stimulation in myocardial infarction survivors: the AIRGENE study. *Journal of the American College of Cardiology,* **52**, 941-952.

Jenkins, G.R., Seiffert, D., Parmer, R.J. & Miles, L.A. (1997) Regulation of plasminogen gene expression by interleukin-6. *Blood,* **89**, 2394-2403.

Kaptoge, S., Seshasai, S.R.K., Gao, P., Freitag, D.F., Butterworth, A.S., Borglykke, A., Di Angelantonio, E., Gudnason, V., Rumley, A. & Lowe, G.D. (2013) Inflammatory cytokines and risk of coronary heart disease: new prospective study and updated meta-analysis. *European Heart Journal***, 35**, 578-589.

Kathiresan, S., Yang, Q., Larson, M.G., Camargo, A.L., Tofler, G.H., Hirschhorn, J.N., Gabriel, S.B. & O'Donnell, C.J. (2006) Common genetic variation in five thrombosis genes and relations to plasma hemostatic protein level and cardiovascular disease risk. *Arteriosclerosis, Thrombosis, and Vascular Biology,* **26**, 1405-1412.

Keavney, B., Danesh, J., Parish, S., Palmer, A., Clark, S., Youngman, L., Delépine, M., Lathrop, M., Peto, R. & Collins, R. (2006) Fibrinogen and coronary heart disease: test of causality by 'Mendelian randomization'. *International Journal of Epidemiology,* **35**, 935-943.

Ken-Dror, G., Humphries, S.E., Kumari, M., Kivimaki, M. & Drenos, F. (2012) A genetic instrument for Mendelian randomization of fibrinogen. *European Journal of Epidemiology,* **27**, 267-279.

Kida, M., Wakabayashi, S. & Ichinose, A. (1997) Expression and induction by IL-6 of the normal and variant genes for human plasminogen. *Biochemical and Biophysical Research Communications,* **230**, 129-132.

Klovaite, J., Nordestgaard, B.G., Tybjærg-Hansen, A. & Benn, M. (2013) Elevated fibrinogen levels are associated with risk of pulmonary embolism, but not with deep venous thrombosis. *American Journal of Respiratory and Critical Care Medicine,* **187**, 286-293.

Kotzé, R.C., Nienaber-Rousseau, C., De Lange, Z., De Maat, M.P., Hoekstra, T. & Pieters, M. (2015) Genetic polymorphisms influencing total and γ' fibrinogen levels and fibrin clot properties in Africans. *British Journal of Haematology,* **168**, 102-112.

Lammertyn, L., Mels, C.M., Pieters, M., Schutte, A.E. & Schutte, R. (2015) Ethnic-specific relationships between haemostatic and oxidative stress markers in black and white South Africans: The SABPA study. *Clinical and Experimental Hypertension,* **37**, 511-517.

Lim, B.C., Ariëns, R.A., Carter, A.M., Weisel, J.W. & Grant, P.J. (2003) Genetic regulation of fibrin structure and function: complex gene-environment interactions may modulate vascular risk. *The Lancet,* **361**, 1424-1431.

Lisman, T., de Groot, P.G., Meijers, J.C. & Rosendaal, F.R. (2005) Reduced plasma fibrinolytic potential is a risk factor for venous thrombosis. *Blood,* **105**, 1102-1105.

Machlus, K.R., Cardenas, J.C., Church, F.C. & Wolberg, A.S. (2011) Causal relationship between hyperfibrinogenemia, thrombosis, and resistance to thrombolysis in mice. *Blood,* **117**, 4953-4963.

Macrae, F.L., Domingues, M.M., Casini, A. & Ariëns, R.A. (2016) The (patho)physiology of fibrinogen γ′. *Seminars in Thrombosis and Hemostasis*, **42**, 344-355.

Mannila, M.N., Eriksson, P., Ericsson, C.-G., Hamsten, A. & Silveira, A. (2006) Epistatic and pleiotropic effects of polymorphisms in the fibrinogen and coagulation factor XIII genes on plasma fibrinogen concentration, fibrin gel structure and risk of myocardial infarction. *Thrombosis and Haemostasis,* **95**, 420-427.

Meade, T.W., Humphries, S.E. & De Stavola, B.L. (2006) Commentary: fibrinogen and coronary heart disease - test of causality by 'Mendelian' randomization by Keavney et al. *International Journal of Epidemiology,* **35**, 944-947.

Meh, D.A., Siebenlist, K.R. & Mosesson, M.W. (1996) Identification and characterization of the thrombin binding sites on fibrin. *Journal of Biological Chemistry,* **271**, 23121-23125.

Mills, J.D., Ariëns, R.A., Mansfield, M.W. & Grant, P.J. (2002) Altered fibrin clot structure in the healthy relatives of patients with premature coronary artery disease. *Circulation,* **106**, 1938-1942.

Montgomery, H.E., Clarkson, P., Nwose, O., Mikailidis, D., Jagroop, I., Dollery, C., Moult, J., Benhizia, F., Deanfield, J. & Jubb, M. (1996) The acute rise in plasma fibrinogen concentration with exercise is influenced by the G-453-A polymorphism of the β-fibrinogen gene. *Arteriosclerosis, Thrombosis, and Vascular Biology,* **16**, 386-391.

Morozumi, T., Sharma, A. & De Nardin, E. (2009) The functional effects of the− 455*G/A* polymorphism on the IL-6-induced expression of the β-fibrinogen gene may be due to linkage disequilibrium with other functional polymorphisms. *Immunological investigations,* **38**, 311-323.

Pieters, M., De Maat, M.P., Jerling, J.C., Hoekstra, T. & Kruger, A. (2011) Fibrinogen concentration and its role in CVD risk in black South Africans – effect of urbanisation. *Journal of Thrombosis and Haemostasis,* **106**, 448-456.

Pieters, M., Kotze, R.C., Jerling, J.C., Kruger, A. & Ariëns, R.A. (2013) Evidence that fibrinogen γ′ regulates plasma clot structure and lysis and relationship to cardiovascular risk factors in black Africans. *Blood,* **121**, 3254-3260.

Pieters, M. & Vorster, H.H. (2008) Nutrition and hemostasis: a focus on urbanization in South Africa. *Molecular Nutrition and Food Research,* **52**, 164-172.

Rein-Smith, C.M., Anderson, N.W. & Farrell, D.H. (2013) Differential regulation of fibrinogen γ chain splice isoforms by interleukin-6. *Thrombosis Research,* **131**, 89-93.

Reiner, A., Carty, C., Carlson, C., Wan, J., Rieder, M., Smith, J., Rice, K., Fornage, M., Jaquish, C. & Williams, O. (2006) Association between patterns of nucleotide variation across the three fibrinogen genes and plasma fibrinogen levels: The coronary artery risk development in young adults (CARDIA) study. *Journal of Thrombosis and Haemostasis,* **4**, 1279-1287.

Sabater-Lleal, M. & Huang, J. & Chasman, D. & Naitza, S. & Dehghan, A. & Johnson, A.D. & Teumer, A. & Reiner, A.P. & Folkersen, L. & Basu, S. & Rudnicka, A.R. & Trompet, S. & Malarstig, A. & Baumert, J. & Bis, J.C. & Guo, X. & Hottenga, J.J. & Shin, S.Y. & Lopez, L.M. & Lahti, J. & Tanaka, T. & Yanek, L.R. & Oudot-Mellakh, T. & Wilson, J.F. & Navarro, P. & Huffman, J.E. & Zemunik, T. & Redline, S. & Mehra, R. & Pulanic, D. &

Rudan, I. & Wright, A.F. & Kolcic, I. & Polasek, O. & Wild, S.H. & Campbell, H. & Curb, J.D. & Wallace, R. & Liu, S. & Eaton, C.B. & Becker, D.M. & Becker, L.C. & Bandinelli, S. & Raikkonen, K. & Widen, E. & Palotie, A. & Fornage, M. & Green, D. & Gross, M. & Davies, G. & Harris, S.E. & Liewald, D.C. & Starr, J.M. & Williams, F.M. & Grant, P.J. & Spector, T.D. & Strawbridge, R.J. & Silveira, A. & Sennblad, B. & Rivadeneira, F. & Uitterlinden, A.G. & Franco, O.H. & Hofman, A. & van Dongen, J. & Willemsen, G. & Boomsma, D.I. & Yao, J. & Swords Jenny, N. & Haritunians, T. & McKnight, B. & Lumley, T. & Taylor, K.D. & Rotter, J.I. & Psaty, B.M. & Peters, A. & Gieger, C. & Illig, T. & Grotevendt, A. & Homuth, G. & Volzke, H. & Kocher, T. & Goel, A. & Franzosi, M.G. & Seedorf, U. & Clarke, R. & Steri, M. & Tarasov, K.V. & Sanna, S. & Schlessinger, D. & Stott, D.J. & Sattar, N. & Buckley, B.M. & Rumley, A. & Lowe, G.D. & McArdle, W.L. & Chen, M.H. & Tofler, G.H. & Song, J. & Boerwinkle, E. & Folsom, A.R. & Rose, L.M. & Franco-Cereceda, A. & Teichert, M. & Ikram, M.A. & Mosley, T.H. & Bevan, S. & Dichgans, M. & Rothwell, P.M. & Sudlow, C.L. & Hopewell, J.C. & Chambers, J.C. & Saleheen, D. & Kooner, J.S. & Danesh, J. & Nelson, C.P. & Erdmann, J. & Reilly, M.P. & Kathiresan, S. & Schunkert, H. & Morange, P.E. & Ferrucci, L. & Eriksson, J.G. & Jacobs, D. & Deary, I.J. & Soranzo, N. & Witteman, J.C. & de Geus, E.J. & Tracy, R.P. & Hayward, C. & Koenig, W. & Cucca, F. & Jukema, J.W. & Eriksson, P. & Seshadri, S. & Markus, H.S. & Watkins, H. & Samani, N.J. & Wallaschofski, H. & Smith, N.L. & Tregouet, D. & Ridker, P.M. & Tang, W. & Strachan, D.P. & Hamsten, A. & O'Donnell, C.J. (2013) Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation,* **128**, 1310-1324.

Sidelmann, J.J., Gram, J., Jespersen, J. & Kluft, C. (2000) Fibrin clot formation and lysis: basic mechanisms. *Seminars in Thrombosis and Hemostasis*, Vol. **26**, 605-618.

Teo, K., Chow, C.K., Vaz, M., Rangarajan, S. & Yusuf, S. (2009) The prospective urban rural epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries. *American Heart Journal,* **158**, 1-7.

Teo, Y.-Y., Small, K.S. & Kwiatkowski, D.P. (2010) Methodological challenges of genome-wide association analysis in Africa. *Nature Reviews Genetics,* **11**, 149-160.

Thomas, A., Lamlum, H., Humphries, S. & Green, F. (1994) Linkage disequilibrium across the fibrinogen locus as shown by five genetic polymorphisms, *G/A*-455 (HaeIII), *C/T*-148 (HindIII/AluI), T/G+1689 (AvaII), and BclI (beta-fibrinogen) and TaqI (alpha-fibrinogen), and their detection by PCR. *Human Mutation,* **3**, 79-81.

Tzoulaki, I., Murray, G.D., Lee, A.J., Rumley, A., Lowe, G.D. & Fowkes, F.G.R. (2007) Relative value of inflammatory, hemostatic, and rheological factors for incident myocardial infarction and stroke The Edinburgh Artery Study. *Circulation,* **115**, 2119-2127.

Uitte de Willige, S., de Visser, M.C., Houwing-Duistermaat, J.J., Rosendaal, F.R., Vos, H.L. & Bertina, R.M. (2005) Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen γ′ levels. *Blood,* **106**, 4176-4183.

Uitte de Willige, S., Standeven, K.F., Philippou, H. & Ariëns, R.A. (2009) The pleiotropic role of the fibrinogen γ′ chain in hemostasis. *Blood,* **114**, 3994-4001.

Undas, A. & Ariëns, R.A. (2011) Fibrin clot structure and function: a role in the pathophysiology of arterial and venous thromboembolic diseases. *Arteriosclerosis, Thrombosis, and Vacular Biology,* **31**, e88-99.

Undas, A., Szułdrzynski, K., Stepien, E., Zalewski, J., Godlewski, J., Tracz, W., Pasowicz, M. & Zmudka, K. (2008) Reduced clot permeability and susceptibility to lysis in patients with acute coronary syndrome: effects of inflammation and oxidative stress. *Atherosclerosis,* **196**, 551-557.

Van Der Bom, J.G., De Maat, M.P., Bots, M.L., Haverkate, F., De Jong, P., Hofman, A., Kluft, C. & Grobbee, D.E. (1998) Elevated plasma fibrinogen cause or consequence of cardiovascular disease? *Arteriosclerosis, Thrombosis, and Vacular Biology,* **18**, 621-625.

van't Hooft, F.M., von Bahr, S.J., Silveira, A., Iliadou, A., Eriksson, P. & Hamsten, A. (1999) Two common, functional polymorphisms in the promoter region of the β-fibrinogen gene contribute to regulation of plasma fibrinogen concentration. *Arteriosclerosis, Thrombosis, and Vacular Biology,* **19**, 3063-3070.

Verschuur, M., de Jong, M., Felida, L., de Maat, M.P. & Vos, H.L. (2005) A hepatocyte nuclear factor-3 site in the fibrinogen β promoter is important for interleukin 6-induced expression, and its activity is influenced by the adjacent–148*C/T* polymorphism. *Journal of Biological Chemistry,* **280**, 16763-16771.

Wassel, C.L., Lange, L.A., Keating, B.J., Taylor, K.C., Johnson, A.D., Palmer, C., Ho, L.A., Smith, N.L., Lange, E.M. & Li, Y. (2011) Association of genomic loci from a cardiovascular gene SNP array with fibrinogen levels in European Americans and African-Americans from six cohort studies: The candidate gene association resource (CARe). *Blood,* **117**, 268-275.

Wolfenstein-Todel, C. & Mosesson, M.W. (1980) Human plasma fibrinogen heterogeneity: evidence for an extended carboxyl-terminal sequence in a normal gamma chain variant (γ'). *Proceedings of the National Academy of Sciences of the United States of America,* **77**, 5069-5073.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S. & Gil, L. (2016) Ensembl 2016. *Nucleic Acids Research,* **44**, D710-716.

Zerbino, D.R., Johnson, N., Juetteman, T., Sheppard, D., Wilder, S.P., Lavidas, I., Nuhn, M., Perry, E., Raffaillac-Desfosses, Q. & Sobral, D. (2016) Ensembl regulation resources. Database (Oxford), **2016**:bav119.

Zhang, Z., Fuentes, N.L. & Fuller, G.M. (1995) Characterization of the IL-6 responsive elements in the γ fibrinogen gene promoter. *Journal of Biological Chemistry,* **270**, 24287-24291.

Table 3.5 Primer and synthetic control sequences used for KASP analyses

Assay ID	Primer_AlleleFAM#		Primer_AlleleHEX#		Primer_Common*
***FGB*-rs1800790**	CATTATGATATAACATTACTATTGATTTTAATA	**A**	CATTATGATATAACATTACTATTGATTTTAATG	**G**	ACAATGACATAATTCTATTTCAAAAGGGGC
***FGB*-rs1800789**	TCCTGTATATATTTTAATTAATAGCCACATAA	**A**	CCTGTATATATTTTAATTAATAGCCACATAG	**G**	GGTTTTTAAAGGGATACATGATCTGACAGA
***FGB*-rs7439150**	CAGCTAATATGAAGAACACTGCACTA	**A**	CAGCTAATATGAAGAACACTGCACTG	**G**	GGGAAAGGGAGTACTATTTGTTCAGTTAT
***FGB*-rs4463047**	CTTGGTTCATAAATAAAGAGGAACTAGAA	**T**	CTTGGTTCATAAATAAAGAGGAACTAGAG	**C**	CTATAATGGATGTATGCACTTGGGACATA
Control *FGB*-rs1800790	Major_5′TCATAGAATAGGGTATGAATTTGTTATTTTGTTATTTTGATTAATGTCTAAAACAAAAGATAAACA*CATTATGATATAACATTACTATTGATTTTAAT**[G]***GCCCCTTTTGAAATAGAATTATGTCATTGTCAGAAAACATAAGCATTTATGGTATATCATTAATGAGTCACGATTTTAGTGGTTGCCTTGTGAGTAG′3 Minor_5′TCATAGAATAGGGTATGAATTTGTTATTTTGTTATTTTGATTAATGTCTAAAACAAAAGATAAACA*CATTATGATATAACATTACTATTGATTTTAAT**[A]***GCCCCTTTTGAAATAGAATTATGTCATTGTCAGAAAACATAAGCATTTATGGTATATCATTAATGAGTCACGATTTTAGTGGTTGCCTTGTGAGTAG′3				
Control *FGB*-rs1800789	Major_5′TGTTATCAGTTATATTTTCATGGAAAATAT*TCCTGTATATATTTTAATTAATAGCCACATA**[G]***ATATTTGCTTTTTCTGTCAGATCATGTATCCCTTTAAAAACCATTAAAGCTAAAGAAAAAAATCTTAAAATATAATATTTACTCATTGCTAATAAAGGGAACACAAATTTGGAAAGACTCAAGTGAATATTTTT′3 Minor_5′TGTTATCAGTTATATTTTCATGGAAAATAT*TCCTGTATATATTTTAATTAATAGCCACATA**[A]***ATATTTGCTTTTTCTGTCAGATCATGTATCCCTTTAAAAACCATTAAAGCTAAAGAAAAAAATCTTAAAATATAATATTTACTCATTGCTAATAAAGGGAACACAAATTTGGAAAGACTCAAGTGAATATTTTT′3				
Control *FGB*-rs7439150	Major_5′TCAAGCCAGAGAGCTGCTGGGCAGGACAGATTATCTCTGTTTTCTCTGTCACTGATTTACTACCCTTGACTTATGCAATGACAT*CAGCTAATATGAAGAACACTGCACT**[G]***ATGACCTCAAAATAACTGAACAAATAGTACTCCCTTTCCCCTCCACCAATGGCATCAACATTCTAAAAAAATTTAAGCATCTTAAT′3 Minor_5′TCAAGCCAGAGAGCTGCTGGGCAGGACAGATTATCTCTGTTTTCTCTGTCACTGATTTACTACCCTTGACTTATGCAATGACAT*CAGCTAATATGAAGAACACTGCACT**[A]***ATGACCTCAAAATAACTGAACAAATAGTACTCCCTTTCCCCTCCACCAATGGCATCAACATTCTAAAAAAATTTAAGCATCTTAAT′3				
Control *FGB*-rs4463047	Major_5′ATACATATATATATATACACATATATAATTATATATATGTATATATAATTATACAATTAGGACTGAAAGTCTGCAACAAACTATAATGGATGTATGCACTTGGGACATAAGC*[T]**TCTAGTTCCTCTTTATTTATGAACCAAG*AGAAACAGCTAACTCAGGGATTGTTTTTAAACTGACTACAGATTCCCACCATGGG′3 Minor_5′ATACATATATATATATACACATATATAATTATATATATGTATATATAATTATACAATTAGGACTGAAAGTCTGCAACAAACTATAATGGATGTATGCACTTGGGACATAAGC*[C]**TCTAGTTCCTCTTTATTTATGAACCAAG*AGAAACAGCTAACTCAGGGATTGTTTTTAAACTGACTACAGATTCCCACCATGGG′3				

FGB = fibrinogen beta chain gene; * Underlined sequences correspond to the common primer annealing position; # Italic sequences correspond to the discriminating primer annealing position; Heterozygote control comprised 50/50 major/minor mixture.

Table 3.6 Associations of individual SNPs with fibrinogen γ′, γ′ ratio and total fibrinogen as published by Kotzé et al, (2015)

Genotype	Fibrinogen γ' (g/l)	γ' ratio	Fibrinogen (g/l)
rs2070011			
GG	0.30 (0.23-0.44) *	10.1 (7.14-14.7)	2.80 (2.30-5.00)
	(n = 1021)	(n = 987)	(n = 1112)
GA	0.32 (0.23-0.46) ^{0}	10.3 (7.42-15.0)	2.90 (2.20-5.00)
	(n = 428)	(n = 405)	(n = 422)
AA	0.39 (0.33-0.55) * ^{0}	11.1 (6.55-17.6)	3.70 (2.80-6.20)
	(n = 40)	(n = 36)	(n = 37)
p-value	0.008	0.50	0.11
p-value after adjustment for fibrinogen γ'	-	0.52	0.52
rs1049636			
TT	0.31 (0.23-0.45)	10.3 (7.22-14.8)	2.80 (2.20-4.50) *
	(n = 1058)	(n = 1020)	(n = 1069)
TC	0.31 (0.23-0.47)	10.3 (7.05-15.1)	3.00 (2.20-5.30) ^{0}
	(n = 392)	(n = 371)	(n = 383)
CC	0.32 (0.22-0.57)	8.30 (5.67-13.6)	5.00 (2.80-6.25) * ^{0}
	(n = 45)	(n = 43)	(n = 44)
p-value	0.61	0.08	0.0009
p-value after adjustment for fibrinogen	0.93	0.93	-
rs1800791			
GG	0.31 (0.23-0.45)	10.2 (7.21-14.5)	2.90 (2.20-5.00) *
	(n = 1462)	(n = 1402)	(n = 1460)
GA	0.31(0.23-0.46)	10.3 (7.11-15.1)	3.00 (2.30-5.00)
	(n = 274)	(n = 265)	(n = 286)
AA	0.35 (0.28-0.50)	6.97 (5.39-10.2)	5.30 (2.90-6.40) *
	(n = 11)	(n = 11)	(n = 11)
p-value	0.21	0.40	0.04
p-value after adjustment for fibrinogen	0.48	0.48	-

Data reported as median (25th-75th percentile). * ^{0} Means with the same symbol differ significantly. A, adenine; C, cytosine; G, guanine; SNP, single nucleotide polymorphism; T, thymine; γ', gamma prime.

Table 3.7 Associations of individual SNPs with clot-related phenotypes as published by Kotzé et al, (2015)

Genotype	Clot lysis time (minutes)	Lagtime (minutes)	Slope	Maximum absorbance
rs1800787				
CC	57.3 [56.7-57.9]	6.36 [6.25-6.47]	9.83 [9.58-10.1]	0.43 [0.42-0.44] *
	(n = 1317)	(n = 1345)	(n = 1316)	(n = 1318)
CT	56.2 [54.3-58.0]	6.53 [6.23-6.83]	9.93 [9.30-10.6]	0.47 [0.44-0.50] *
	(n = 165)	(n = 170)	(n = 168)	(n = 165)
TT	57.3 [49.8-64.7]	7.06 [5.51-8.60]	9.76 [8.00-11.5]	0.48 [0.42-0.54]
	(n = 9)	(n = 9)	(n = 9)	(n = 9)
p-value	0.27	0.16	0.82	0.001
rs2070011				
GG	57.4 [56.7-58.1]	6.43 [6.30-6.55]	9.69 [9.41-9.96]	0.43 [0.42-0.43] *
	(n = 1034)	(n = 1061)	(n = 1038)	(n = 1035)
GA	56.9 [55.9-57.9]	6.29 [6.11-6.47]	10.2 [9.79-10.7]	0.45 [0.43-0.46] *
	(n = 440)	(n = 446)	(n = 438)	(n = 440)
AA	59.1 [55.2-63.0]	6.91 [6.34-7.48]	9.11 [7.88-10.3]	0.45 [0.41-0.50]
	(n = 43)	(n = 42)	(n = 42)	(n = 43)
p-value	0.93	0.90	0.22	0.01
p-value adjusted for fibrinogen γ'	0.97	0.81	0.21	0.06
rs1049636				
TT	57.3 [56.7-58.0]	6.37 [6.25-6.48]	9.76 [9.49-10.0]	0.43 [0.42-0.44] *
	(n = 1082)	(n = 1106)	(n = 1081)	(n = 1083)
TC	57.2 [56.0-58.3]	6.47 [6.28-6.67]	9.93 [9.48-10.4]	0.44 [0.43-0.46]
	(n = 394)	(n = 401)	(n = 395)	(n = 394)
CC	57.7 [54.3-61.0]	6.58 [6.09-7.08]	10.1 [8.74-11.5]	0.48 [0.43-0.53] *
	(n = 46)	(n = 48)	(n = 48)	(n = 46)
p-value	0.94	0.27	0.42	0.01
p-value adjusted for fibrinogen	0.88	0.84	0.72	0.33
rs5985				
CC	57.0 [56.3-57.7]	6.31 [6.20-6.42] *	9.91 [9.65-10.2]	0.43 [0.42-0.44]
	(n = 1105)	(n = 1130)	(n = 1110)	(n = 1105)
CA	57.8 [56.6-58.9]	6.61 [6.39-6.84] *	9.74 [9.23-10.2]	0.44 [0.42-0.45]
	(n = 360)	(n = 366)	(n = 357)	(n = 361)
AA	58.8 [54.9-62.7]	6.22 [5.61-6.83]	8.67 [7.45-9.90]	0.40 [0.32-0.48]
	(n = 28)	(n = 30)	(n = 28)	(n = 28)
p-value	0.39	0.04	0.30	0.55

Data reported as mean [95% confidence interval]. * Means with the same symbol differ significantly. A, adenine; C, cytosine; G, guanine; SNP, single nucleotide polymorphism; T, thymine; γ', gamma prime.

Table 3.8 Genotype-IL-6 interactions modulating total and γ' fibrinogen concentrations upon removal of individuals with IL-6 > 100pg/mL

Phenotype	Gene	IL-6 interaction with SNP	Interaction p-value		Major allele carriers		Minor allele carriers	
					N	Slope (95% CI)	N	Slope (95% CI)
Total fibrinogen (g/L)	*FGG*	rs1049636	<0.001	TT	969	0.056 (0.041 – 0.071)	371	0.110 (0.082 – 0.138)
Fibrinogen γ' ratio (%)	*FGB*	rs7439150[δ]	0.01	GG	1236	-0.008 (-0.067 – 0.050)	166	0.182 (0.040 – 0.325)
	FGB	rs1800790	0.01	GG	1304	-0.012 (-0.068 – 0.043)	93	0.242 (0.051 – 0.433)
	FGB	rs1800787[δ]	0.005	CC	1120	-0.036 (-0.097 – 0.025)	144	0.169 (0.014 – 0.323)

C = cytosine; CI = confidence interval; G = guanine; IL-6 = interleukin-6; rs = reference sequence; T = thymine; rs1800790 = -455G/A; rs1800787 = -148C/T; rs1049636 = 9340T/C; δVariants in high linkage disequilibrium (r^2 = 0.88; D' = 0.95)

CHAPTER 4

CANDIDATE GENE ANALYSIS OF THE FIBRINOGEN PHENOTYPE REVEALS THE IMPORTANCE OF POLYGENIC CO-REGULATION

This chapter includes:

- Guide for authors , Matrix Biology (Impact factor: 4.5);
- Letter of acceptance for publication in Matrix Biology;
- The original article titled: “Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation”.

The manuscript as published online can be viewed in Annexure A

GUIDE FOR AUTHORS: MATRIX BIOLOGY

https://www.elsevier.com/journals/matrix-biology/0945-053x/guide-for-authors

YOUR PAPER YOUR WAY

We now differentiate between the requirements for new and revised submissions. You may choose to submit your manuscript as a single Word or PDF file to be used in the refereeing process. Only when your paper is at the revision stage, will you be requested to put your paper in to a 'correct format' for acceptance and provide the items required for the publication of your article.

To find out more, please visit the Preparation section below.

INTRODUCTION

Matrix Biology publishes Regular Research Papers, Brief Reports, Mini Reviews, Landmark Essays, Full Length Reviews and Announcements. The section Types of Papers describes the conditions each of these submission types should meet.

Papers that do not conform to the general criteria for publication in *Matrix Biology* will be returned immediately to authors to avoid unnecessary delay in submission elsewhere.

The decision of the Editor is final; the Editor cannot enter into correspondence about a paper that is considered unsuitable for publication.

Revisions should be submitted within 90 days. If we do not receive your revision within 120 days, your paper will be withdrawn and removed from the system.

Types of papers

Regular Research Papers

are studies utilizing most scientific technologies including molecular biology, cell biology, immunochemistry, structural biology, computational biology, theoretical biology, and macromolecular chemistry where the subject is extracellular matrix or is substantially related to matrix and its biological role. The journal will publish articles that are scientifically rigorous, complete within a logical framework, and, most importantly, address molecular or cellular mechanisms that are timely and represent above average significance.

Brief Reports

are short papers that meet the same standards of quality and field of interest as regular papers, but have a narrower scope and a more focused interest. Brief reports should present new information of exceptional novelty, importance and interest to the broad readership of the journal. Reports should have a maximum of 3 figures, and no longer than 30,000 characters, including spaces. There are no limitations for references.

Commentaries

are brief reviews that present a sharply focused view of a rapidly advancing area of research. Authorship is normally by invitation. Generally, commentaries include a manuscript of 6-10 double-spaced pages with 1-2 schematics. References are limited to 20-30.

Mini Reviews

are timely and original analyses of current topics of high interest. The mini-reviews should summarize the general concepts in the field, provide a brief summary of the new information and define unresolved questions about currently available data. The mini reviews must contain up to 3 original graphics of high quality or equivalent images, and a manuscript containing 12-15 double-spaced pages. There are no limitations for references. We do not accept previously-published graphics or images. The Editor-in-Chief should be consulted in advance by anyone wishing to submit an unsolicited mini review.

Full Length Reviews

are comprehensive and critical analyses of topics that have been under investigation for some time. Most of them are invited reviews from experts in the relevant topic. The reviews must contain up to five high-quality, original graphics or equivalent images, and a manuscript containing 20-40 double-spaced pages. No limitations for references. We do not accept previously-published graphics or images. The Editor-in-Chief should be consulted in advance by anyone wishing to submit an unsolicited full-length review.

Announcements

Announcements of national or international meetings may be published in Matrix Biology.

Contact details for submission

Manuscripts for Matrix Biology should be submitted via the journal's online submission page at http://www.evise.com/evise/faces/pages/ navigation/NavController.jspx? JRNL_ACR=MATBIO.

For all technical queries please consult our help site at http://epsupport.elsevier.com/. Here you will be able to learn more about the online submission system via interactive tutorials, explore a range of online submission solutions via our knowledgebase, and find answers to frequently asked questions. You will also find our 24/7 support contact details should you need any assistance from one of our customer service representatives.

Submission checklist

You can use this list to carry out a final check of your submission before you send it to the journal for review. Please check the relevant section in this Guide for Authors for more details.

Ensure that the following items are present:

One author has been designated as the corresponding author with contact details:

E-mail address

Full postal address

All necessary files have been uploaded:

Manuscript:

Include keywords

All figures (include relevant captions)

All tables (including titles, description, footnotes)

Ensure all figure and table citations in the text match the files provided

Indicate clearly if color should be used for any figures in print

Graphical Abstracts / Highlights files (where applicable)

Supplemental files (where applicable)

Further considerations

Manuscript has been 'spell checked' and 'grammar checked'

All references mentioned in the Reference List are cited in the text, and vice versa

Permission has been obtained for use of copyrighted material from other sources (including the Internet)

Relevant declarations of interest have been made

Journal policies detailed in this guide have been reviewed

Referee suggestions and contact details provided, based on journal requirements For further information, visit our Support Center.

BEFORE YOU BEGIN

Ethics in publishing

Please see our information pages on Ethics in publishing and Ethical guidelines for journal publication.

Declaration of interest

All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work. More information.

Submission declaration and verification

Submission of an article implies that the work described has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see 'Multiple, redundant or concurrent publication' section of our ethics policy for more information), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. To verify originality, your article may be checked by the originality detection service CrossCheck.

Contributors

Each author is required to declare his or her individual contribution to the article: all authors must have materially participated in the research and/or article preparation, so roles for all authors should be described. The statement that all authors have approved the final article should be true and included in the disclosure.

Changes to authorship

Authors are expected to consider carefully the list and order of authors **before** submitting their manuscript and provide the definitive list of authors at the time of the original submission. Any addition, deletion or rearrangement of author names in the authorship list should be made only

before the manuscript has been accepted and only if approved by the journal Editor. To request such a change, the Editor must receive the following from the **corresponding author**: (a) the reason for the change in author list and (b) written confirmation (e-mail, letter) from all authors that they agree with the addition, removal or rearrangement. In the case of addition or removal of authors, this includes confirmation from the author being added or removed.

Only in exceptional circumstances will the Editor consider the addition, deletion or rearrangement of authors **after** the manuscript has been accepted. While the Editor considers the request, publication of the manuscript will be suspended. If the manuscript has already been published in an online issue, any requests approved by the Editor will result in a corrigendum.

Copyright

Upon acceptance of an article, authors will be asked to complete a 'Journal Publishing Agreement' (see more information on this). An e-mail will be sent to the corresponding author confirming receipt of the manuscript together with a 'Journal Publishing Agreement' form or a link to the online version of this agreement.

Subscribers may reproduce tables of contents or prepare lists of articles including abstracts for internal circulation within their institutions. Permission of the Publisher is required for resale or distribution outside the institution and for all other derivative works, including compilations and translations. If excerpts from other copyrighted works are included, the author(s) must obtain written permission from the copyright owners and credit the source(s) in the article. Elsevier has preprinted forms for use by authors in these cases.

For open access articles: Upon acceptance of an article, authors will be asked to complete an 'Exclusive License Agreement' (more information). Permitted third party reuse of open access articles is determined by the author's choice of user license.

Author rights

As an author you (or your employer or institution) have certain rights to reuse your work. More information.

Elsevier supports responsible sharing

Find out how you can share your research published in Elsevier journals.

Role of the funding source

You are requested to identify who provided financial support for the conduct of the research and/or preparation of the article and to briefly describe the role of the sponsor(s), if any, in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. If the funding source(s) had no such involvement then this should be stated.

Funding body agreements and policies

Elsevier has established a number of agreements with funding bodies which allow authors to comply with their funder's open access policies. Some funding bodies will reimburse the author for the Open Access Publication Fee. Details of existing agreements are available online.

Open access

This journal offers authors a choice in publishing their research:

Open access

Articles are freely available to both subscribers and the wider public with permitted reuse.

An open access publication fee is payable by authors or on their behalf, e.g. by their research funder or institution.

Subscription

Articles are made available to subscribers as well as developing countries and patient groups through our universal access programs.

No open access publication fee payable by authors.

Regardless of how you choose to publish your article, the journal will apply the same peer review criteria and acceptance standards.

For open access articles, permitted third party (re)use is defined by the following Creative Commons user licenses:

Creative Commons Attribution (CC BY)

Lets others distribute and copy the article, create extracts, abstracts, and other revised versions, adaptations or derivative works of or from an article (such as a translation), include in a collective work (such as an anthology), text or data mine the article, even for commercial purposes, as long as they credit the author(s), do not represent the author as endorsing their adaptation of the article, and do not modify the article in such a way as to damage the author's honor or reputation.

Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

For non-commercial purposes, lets others distribute and copy the article, and to include in a collective work (such as an anthology), as long as they credit the author(s) and provided they do not alter or modify the article.

The open access publication fee for this journal is **USD 3300**, excluding taxes. Learn more about Elsevier's pricing policy: https://www.elsevier.com/openaccesspricing.

Green open access

Authors can share their research in a variety of different ways and Elsevier has a number of green open access options available. We recommend authors see our green open access page for further information. Authors can also self-archive their manuscripts immediately and enable public access from their institution's repository after an embargo period. This is the version that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and in editor-author communications. Embargo period: For subscription articles, an appropriate amount of time is needed for journals to deliver value to subscribing customers before an article becomes freely available to the public. This is the embargo period and it begins from the date the article is formally published online in its final and fully citable form. Find out more.

This journal has an embargo period of 12 months.

Language (usage and editing services)

Please write your text in good English (American or British usage is accepted, but not a mixture of these). Authors who feel their English language manuscript may require editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English may wish to use the English Language Editing service available from Elsevier's WebShop.

Submission

Our online submission system guides you stepwise through the process of entering your article details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process. Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the Editor's decision and requests for revision, is sent by e-mail.

Referees

Please submit the names and institutional e-mail addresses of several potential referees. For more details, visit our Support site. Note that the editor retains the sole right to decide whether or not the suggested reviewers are used.

PREPARATION

New submissions

Submission to this journal proceeds totally online and you will be guided stepwise through the creation and uploading of your files. The system automatically converts your files to a single PDF file, which is used in the peer-review process.

As part of the Your Paper Your Way service, you may choose to submit your manuscript as a single file to be used in the refereeing process. This can be a PDF file or a Word document, in any format or lay- out that can be used by referees to evaluate your manuscript. It should contain high enough quality figures for refereeing. If you prefer to do so, you may still provide all or some of the source files at the initial submission. Please note that individual figure files larger than 10 MB must be uploaded separately.

Figures and tables embedded in text

Please ensure the figures and the tables included in the single file are placed next to the relevant text in the manuscript, rather than at the bottom or the top of the file.

REVISED SUBMISSIONS

Use of word processing software

Regardless of the file format of the original submission, at revision you must provide us with an editable file of the entire article. Keep the layout of the text as simple as possible. Most formatting codes will be removed and replaced on processing the article. The electronic text should be prepared in a way very similar to that of conventional manuscripts (see also the Guide to Publishing with Elsevier). See also the section on Electronic artwork.

To avoid unnecessary errors you are strongly advised to use the 'spell-check' and 'grammar-check' functions of your word processor.

Article structure

Subdivision - numbered sections

Divide your article into clearly defined and numbered sections. Subsections should be numbered 1.1 (then 1.1.1, 1.1.2, ...), 1.2, etc. (the abstract is not included in section numbering). Use this numbering also for internal cross-referencing: do not just refer to 'the text'. Any subsection may be given a brief heading. Each heading should appear on its own separate line.

Introduction

State the objectives of the work and provide an adequate background, avoiding a detailed literature survey or a summary of the results.

Results

Results should be clear and concise.

Discussion

This should explore the significance of the results of the work, not repeat them. A combined Results and Discussion section is often appropriate. Avoid extensive citations and discussion of published literature.

Essential title page information

Title. Concise and informative. Titles are often used in information-retrieval systems. Avoid abbreviations and formulae where possible.

Author names and affiliations. Please clearly indicate the given name(s) and family name(s) of each author and check that all names are accurately spelled. Present the authors' affiliation addresses (where the actual work was done) below the names. Indicate all affiliations with a lower- case superscript letter immediately after the author's name and in front of the appropriate address. Provide the full postal address of each affiliation, including the country name and, if available, the e-mail address of each author.

Corresponding author. Clearly indicate who will handle correspondence at all stages of refereeing and publication, also post-publication. **Ensure that the e-mail address is given and that contact details are kept up to date by the corresponding author.**

Present/permanent address. If an author has moved since the work described in the article was done, or was visiting at the time, a 'Present address' (or 'Permanent address') may be indicated as a footnote to that author's name. The address at which the author actually did the work must be retained as the main, affiliation address. Superscript Arabic numerals are used for such footnotes.

Abstract

A concise and factual abstract is required. The abstract should state briefly the purpose of the research, the principal results and major conclusions. An abstract is often presented separately from the article, so it must be able to stand alone. For this reason, References should be avoided, but if essential, then cite the author(s) and year(s). Also, non-standard or uncommon abbreviations should be avoided, but if essential they must be defined at their first mention in the abstract itself.

An abstract is not required for Landmark Essays.

Keywords

Immediately after the abstract, provide a maximum of 6 keywords, using American spelling and avoiding general and plural terms and multiple concepts (avoid, for example, 'and', 'of'). Be sparing with abbreviations: only abbreviations firmly established in the field may be eligible. These keywords will be used for indexing purposes.

Abbreviations

Define abbreviations that are not standard in this field in a footnote to be placed on the first page of the article. Such abbreviations that are unavoidable in the abstract must be defined at their first mention there, as well as in the footnote. Ensure consistency of abbreviations throughout the article.

Acknowledgements

Collate acknowledgements in a separate section at the end of the article before the references and do not, therefore, include them on the title page, as a footnote to the title or otherwise. List here those individuals who provided help during the research (e.g., providing language help, writing assistance or proof reading the article, etc.).

Formatting of funding sources

List funding sources in this standard way to facilitate compliance to funder's requirements:

Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa]. It is not necessary to include detailed descriptions on the program or type of grants and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.

If no funding has been provided for the research, please include the following sentence:
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Math formulae

Please submit math equations as editable text and not as images. Present simple formulae in line with normal text where possible and use the solidus (/) instead of a horizontal line for small fractional terms, e.g., X/Y. In principle, variables are to be presented in italics. Powers of e are often more conveniently denoted by exp. Number consecutively any equations that have to be displayed separately from the text (if referred to explicitly in the text).

Footnotes

Footnotes should be used sparingly. Number them consecutively throughout the article. Many word processors build footnotes into the text, and this feature may be used. Should this not be the case, indicate the position of footnotes in the text and present the footnotes themselves separately at the end of the article.

Artwork

Electronic artwork

General points

Make sure you use uniform lettering and sizing of your original artwork.

Preferred fonts: Arial (or Helvetica), Times New Roman (or Times), Symbol, Courier.

Number the illustrations according to their sequence in the text.

Use a logical naming convention for your artwork files.

Indicate per figure if it is a single, 1.5 or 2-column fitting image.

For Word submissions only, you may still provide figures and their captions, and tables within a single file at the revision stage.

Please note that individual figure files larger than 10 MB must be provided in separate source files. A detailed guide on electronic artwork is available.

You are urged to visit this site; some excerpts from the detailed information are given here.

Formats

Regardless of the application used, when your electronic artwork is finalized, please 'save as' or convert the images to one of the following formats (note the resolution requirements for line drawings, halftones, and line/halftone combinations given below):

EPS (or PDF): Vector drawings. Embed the font or save the text as'graphics'.

TIFF (or JPG): Color or grayscale photographs (halftones): always use a minimum of 300 dpi. TIFF (or JPG): Bitmapped line drawings: use a minimum of 1000 dpi.

TIFF (or JPG): Combinations bitmapped line/half-tone (color or grayscale): a minimum of 500 dpi is required.

Please do not:

Supply files that are optimized for screen use (e.g., GIF, BMP, PICT, WPG); the resolution is too low.

Supply files that are too low in resolution.

Submit graphics that are disproportionately large for the content.

Color artwork

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

Figure captions

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

Tables

Please submit tables as editable text and not as images. Tables can be placed either next to the relevant text in the article, or on separate page(s) at the end. Number tables consecutively in accordance with their appearance in the text and place any table notes below the table body. Be sparing in the use of tables and ensure that the data presented in them do not duplicate results described elsewhere in the article. Please avoid using vertical rules.

References

Citation in text

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

Web references

As a minimum, the full URL should be given. Any further information, if known (DOI, author names, dates, reference to a source publication, etc.), should also be given. Reference to material which is available on the Internet but has not been published elsewhere should be made in the text only and should not be included in the reference list.

Data references

This journal encourages you to cite underlying or relevant datasets in your manuscript by citing them in your text and including a data reference in your Reference List. Data references should include the following elements: author name(s), dataset title, data repository, version (where available), year, and global persistent identifier. Add [dataset] immediately before the reference so we can properly identify it as a data reference. This identifier will not appear in your published article.

References in a special issue

Please ensure that the words 'this issue' are added to any references in the list (and any citations in the text) to other articles in the same Special Issue.

Reference management software

Most Elsevier journals have their reference template available in many of the most popular reference management software products. These include all products that support Citation Style Language styles, such as Mendeley and Zotero, as well as EndNote. Using the word processor plug-ins from these products, authors only need to select the appropriate journal template when preparing their article, after which citations and bibliographies will be automatically formatted in the journal's style. If no template is yet available for this journal, please follow the format of the sample references and citations as shown in this Guide. Users of Mendeley Desktop can easily install the reference style for this journal by clicking the following link: http://open.mendeley.com/use-citation-style/matrix-biology

When preparing your manuscript, you will then be able to select this style using the Mendeley plug- ins for Microsoft Word or LibreOffice.

Reference style

Text: Indicate references by number(s) in square brackets in line with the text. The actual authors can be referred to, but the reference number(s) must always be given.

Example: '..... as demonstrated [3,6]. Barnaby and Jones [8] obtained a different result ....'

List: Number the references (numbers in square brackets) in the list in the order in which they appear in the text.

Examples:

Reference to a journal publication:

[1] J. van der Geer, J.A.J. Hanraads, R.A. Lupton, The art of writing a scientific article, J. Sci. Commun. 163 (2010) 51–59.

Reference to a book:

[2] W. Strunk Jr., E.B. White, The Elements of Style, fourth ed., Longman, New York, 2000. Reference to a chapter in an edited book:

[3] G.R. Mettam, L.B. Adams, How to prepare an electronic version of your article, in: B.S. Jones, R.Z. Smith (Eds.), Introduction to the Electronic Age, E-Publishing Inc., New York, 2009, pp. 281–304.

Reference to a website:

[4] Cancer Research UK, Cancer statistics reports for the UK. http://www.cancerresearchuk.org/ aboutcancer/statistics/cancerstatsreport/, 2003 (accessed 13.03.03).

[dataset][5] M. Oguro, S. Imahiro, S. Saito, T. Nakashizuka, Mortality data for Japanese oak wilt disease and surrounding forest compositions, Mendeley Data, v1, 2015. http://dx.doi.org/10.17632/ xwj98nb39r.1.

Reference style

References should be listed at the end of the manuscript. They should be listed in the order in which they appear in the text, tables, and figure legends and numbered sequentially. When cited in the text, reference numbers should be in brackets. Only papers that have been published or accepted should be cited in the reference list. The title of the article, the volume number, and first and last pages should be cited. Journal titles should be abbreviated, e.g.,

1. Sanger, F. & Coulson, A. R. (1975). A rapid method for determining sequences in DNA by the primed synthesis with DNA polymerase. J. Mol. Biol. 94, 441-448.
2. Goto, Y., Calciano, L. J. & Fink, A. F. (1990). Acid-induced folding of proteins. Proc. Natl. Acad. Sci. USA 87, 573-577. Articles in books should include the title of the article, the name of the book, editor(s), edition number, first and last page numbers, the name and the location of the publisher, e.g.,
3. Hanks, S. K. & Hunter, T. (1995). The eukaryotic protein kinase superfamily. In The Protein Kinase FactsBook: Protein-Serine Kinases (Hardie, G. & Hanks, S., eds), pp. 747, Academic Press, London.

Journal abbreviations source

Journal names should be abbreviated according to the List of Title Word Abbreviations.

Video

Elsevier accepts video material and animation sequences to support and enhance your scientific research. Authors who have video or animation files that they wish to submit with their article are strongly encouraged to include links to these within the body of the article. This can be done in the same way as a figure or table by referring to the video or animation content and noting in the body text where it should be placed. All submitted files should be properly labeled so that they directly relate to the video file's content. In order to ensure that your video or animation material is directly usable, please provide the files in one of our recommended file formats with a preferred maximum size of 150 MB. Video and animation files supplied will be published online in the electronic version of your article in Elsevier Web products, including ScienceDirect. Please supply 'stills' with your files: you can choose any frame from the video or animation or make a separate image. These will be used instead of standard icons and will personalize the link to your video data. For more detailed instructions please visit our video instruction pages. Note: since video and animation cannot be embedded in the print version of the journal, please provide text for both the electronic and the print version for the portions of the article that refer to this content.

Supplementary material

Supplementary material can support and enhance your scientific research. Supplementary files offer the author additional possibilities to publish supporting applications, high-resolution images, background datasets, sound clips and more. Please note that such items are published online exactly as they are submitted; there is no typesetting involved (supplementary data supplied as an Excel file or as a PowerPoint slide will appear as such online). Please submit the material together with the article and supply a concise and descriptive caption for each file. If you wish to make any changes to supplementary data during any stage of the process, then please make sure to provide an updated file, and do not annotate any corrections on a previous version. Please also make sure to switch off the 'Track Changes' option in any Microsoft Office files as these will appear in the published supplementary file(s). For more detailed instructions please visit our artwork instruction pages.

RESEARCH DATA

Data in Brief

Authors have the option of converting any or all parts of their supplementary or additional raw data into one or multiple Data in Brief articles, a new kind of article that houses and describes their data. Data in Brief articles ensure that your data, which is normally buried in supplementary material, is actively reviewed, curated, formatted, indexed, given a DOI and publicly available to all upon publication. Authors are encouraged to submit their Data in Brief article as an additional item directly alongside the revised version of their manuscript. If your research article is accepted, your Data in Brief article will automatically be transferred over to *Data in Brief* where it will be editorially reviewed and published in the new, open access journal, *Data in Brief*. Please note an open access fee is payable for publication in *Data in Brief*. Full details can be found on the Data in Brief website. Please use this template to write your Data in Brief.

Database linking

Elsevier encourages authors to connect articles with external databases, giving readers access to relevant databases that help to build a better understanding of the described research. Please refer to

relevant database identifiers using the following format in your article: Database: xxxx (e.g., TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN). More information and a full list of supported databases.

AudioSlides

The journal encourages authors to create an AudioSlides presentation with their published article. AudioSlides are brief, webinar-style presentations that are shown next to the online article on ScienceDirect. This gives authors the opportunity to summarize their research in their own words and to help readers understand what the paper is about. More information and examples are available. Authors of this journal will automatically receive an invitation e-mail to create an AudioSlides presentation after acceptance of their paper.

3D molecular models

You can enrich your online articles by providing 3D molecular models (optional) in PDB, PSE or MOL/MOL2 format, which will be visualized using the interactive viewer embedded within the article. Using the viewer, it will be possible to zoom into the model, rotate and pan the model, and change display settings. Submitted models will also be available for downloading from your online article on ScienceDirect. Each molecular model will have to be uploaded to the online submission system separately, via the '3D molecular models' submission category. More information.

3D models

You can enrich your online articles by providing 3D models (optional) in PLY, OBJ or U3D format, which will be visualized using the interactive viewer next to the article. Each 3D model will have to be zipped and uploaded to the online submission system via the '3D models' submission category. Please be advised that the recommended model size before zipping is maximum 150 MB. Multiple models can be submitted. Please provide a short informative description for each model by filling in the 'Description' field when uploading a dataset. Note: all datasets will be available for download from the online article on ScienceDirect. If you have concerns about your data being downloadable, please provide a video instead. More information on OBJ and PLY models or U3D models.

Interactive plots

This journal enables you to show an Interactive Plot with your article by simply submitting a data file. Full instructions.

Virtual Microscope

The journal encourages authors to supplement in-article microscopic images with corresponding high resolution versions for use with the Virtual Microscope viewer. The Virtual Microscope is a web based viewer that enables users to view microscopic images at the highest level of detail and provides features such as zoom and pan. This feature for the first time gives authors the opportunity to share true high resolution microscopic images with their readers. More information and examples. Authors of this journal will receive an invitation e-mail to create microscope images for use with the Virtual Microscope when their manuscript is first reviewed. If you opt to use the feature, please contact virtualmicroscope@elsevier.com for instructions on how to prepare and upload the required high resolution images.

Amino acid and nuclei sequences

Sequences should be computer printed using a monospace font and high quality printer. They should be assembled as a camera-ready figure, not a table, so that the alignment can be maintained.

AFTER ACCEPTANCE

Online proof correction

Corresponding authors will receive an e-mail with a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to MS Word: in addition to editing text, you can also comment on figures/tables and answer questions from the Copy Editor. Web-based proofing provides a faster and less error-prone process by allowing you to directly type your corrections, eliminating the potential introduction of errors.

If preferred, you can still choose to annotate and upload your edits on the PDF version. All instructions for proofing will be given in the e-mail we send to authors, including alternative methods to the online version and PDF.

We will do everything possible to get your article published quickly and accurately. Please use this proof only for checking the typesetting, editing, completeness and correctness of the text, tables and figures. Significant changes to the article as accepted for publication will only be considered at this stage with permission from the Editor. It is important to ensure that all corrections are sent back to us in one communication. Please check carefully before replying, as inclusion of any subsequent corrections cannot be guaranteed. Proofreading is solely your responsibility.

Offprints

The corresponding author will, at no cost, receive a customized Share Link providing 50 days free access to the final published version of the article on ScienceDirect. The Share Link can be used for sharing the article via any communication channel, including email and social media. For an extra charge, paper offprints can be ordered via the offprint order form which is sent once the article is accepted for publication. Both corresponding and co-authors may order offprints at any time via Elsevier's Webshop. Corresponding authors who have published their article open access do not receive a Share Link as their final published version of the article is available open access on ScienceDirect and can be shared through the article DOI link.

AUTHOR INQUIRIES

Visit the Elsevier Support Center to find the answers you need. Here you will find everything from Frequently Asked Questions to ways to get in touch. You can also check the status of your submitted article or find out when your accepted article will be published.

LETTER OF ACCEPTANCE FOR PUBLICATION IN MATRIX BIOLOGY

Date: 13 October 2016

Ms. Ref. No.: MATBIO-D-16-00084R1

Title: Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation

Matrix Biology

Dear Cornelie,

I am pleased to confirm that your paper "Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation" has been accepted for publication in Matrix Biology.

Comments from the Editor and Reviewers can be found below.

When your paper is published on ScienceDirect, you want to make sure it gets the attention it deserves. To help you get your message across, Elsevier has developed a new, free service called AudioSlides: brief, webcast-style presentations that are shown (publicly available) next to your published article. This format gives you the opportunity to explain your research in your own words and attract interest. You will receive an invitation email to create an AudioSlides presentation shortly. For more information and examples, please visit http://www.elsevier.com/audioslides.

Your accepted manuscript will now be transferred to our production department and work will begin on creation of the proof. If we need any additional information to create the proof, we will let you know. If not, you will be contacted again in the next few days with a request to approve the proof and to complete a number of online forms that are required for publication.

Thank you for submitting your work to this journal.

With kind regards,

Thomas H. Barker, Ph.D.

ARTICLE

Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation

H. Toinét Cronjé[1], Cornelie Nienaber-Rousseau[1]*, Lizelle Zandberg, Tinashe Chikowore, Zelda de Lange, Tertia van Zyl, Marlien Pieters

Centre of Excellence for Nutrition, North-West University, Potchefstroom, South Africa

[1]Authors contributed equally to this manuscript

***Correspondence**: Dr Cornelie Nienaber-Rousseau, Private bag x6001, Nutrition, Box 594, Potchefstroom, 2520, South Africa; Tel: + 27 18 299 4169; Fax: + 27 18 299 2464; E-mail: cornelie.nienaber@nwu.ac.za

ABSTRACT

Fibrinogen and its functional aspects have been linked to cardiovascular disease. There is vast discrepancy between the heritability of fibrinogen concentrations observed in twin studies and the heritability uncovered by genome wide association studies. We postulate that some of the missing heritability might be explained by the pleiotropic and polygenic co-regulation of fibrinogen through multiple targeted genes, apart from the fibrinogen genes themselves. To this end we investigated single nucleotide polymorphisms (SNPs) in genes coding for phenotypes associated with total and γ' fibrinogen concentrations and clot properties. Their individual and accumulative associations with the fibrinogen variables were explored together with possible co-regulatory processes as a result of the gain and loss of transcription factor binding sites (TFBS). Seventy-eight SNPs spanning the *APOB*, *APOE*, *CBS, CRP*, *F13A1*, *FGA*, *FGB*, *FGG*, *LDL-R*, *MTHFR*, *MTR, PCSK-9* and *SERPINE-1* genes were included in the final analysis. A novel *PCSK-9* SNP (rs369066144) was identified in this population, which associated significantly ($p = 0.04$) with clot lysis time (CLT). Apart from SNPs in the fibrinogen (*FGA*, *FGB* and *FGG*) and FXIII (*F13A1*) genes, the fibrinogen phenotypes were also associated with SNPs in genes playing a role in lipid homeostasis (*LDL-R, PCSK-9*) together with *CBS* and *CRP* polymorphisms (particularly, *CRP*-rs3093068). The genetic risk scores, presenting accumulative genetic risk, were significantly associated ($p \leq 0.007$) with total and γ' fibrinogen concentrations, lag time, slope and CLT, highlighting the importance of a polygenetic approach in determining complex phenotypes. SNPs significantly associated with the fibrinogen phenotypes, resulted in a total of 75 TFBS changes, of which 35 resulted in a loss and 40 in a gain of TFBS. In terms of co-regulation, V$IRF4.02, V$E2FF and V$HIFF were of particular importance. The investigation into TFBS provided valuable insight as to how sequence divergences in seemingly unrelated genes can result in transcriptional co-regulation of the fibrinogen phenotypes. The observed associations between the identified SNPs and the fibrinogen phenotypes therefore do not imply direct effects on cardiovascular disease outcomes, but may prove useful in explaining more of the genetic regulation of the investigated fibrinogen phenotypes.

Keywords: fibrinolysis; gamma prime fibrinogen; transcription factor binding sites; lipids, turbidimetry

INTRODUCTION

Fibrinogen is a hexameric glycoprotein consisting of two alpha (α), two beta (β) and two gamma (γ) polypeptide chains, and is encoded for by the fibrinogen α, β and γ chain genes (*FGA*, *FGB* and *FGG*) on the q-arm of chromosome four [1]. Several variants of the fibrinogen molecule exist, of which fibrinogen gamma prime (γ') is considered to be a common splice variant contributing to between 8 and 11% of total plasma fibrinogen concentrations [2, 3]. Fibrinogen γ' arises in response to alternative splicing of the carboxyl-terminal region of γ mRNA, resulting in a higher molecular weight γ' than the γA chain [2, 3]. Fibrinogen is an essential haemostatic protein, as the precursor of fibrin in the final stages of blood coagulation [4]. Both total and γ' fibrinogen influence clot structure [5, 6], and increased levels have been related to denser blood clots resistant to clot lysis [7]. Denser clots that remain in the vasculature for a longer time are associated with cardiovascular disease (CVD) outcomes such as myocardial infarction, stroke and coronary artery disease [8]. Increased fibrinogen concentration, irrespective of its functional properties, is also associated with CVD [9].

Family and twin studies investigating the genetics of fibrinogen have reported the heritability of fibrinogen concentrations to be between 30 and 51% [10-16]. Genome wide association studies (GWAS) have only allocated 3.7% of this possible 51% fibrinogen heritability to common single nucleotide polymorphisms (SNPs; single base pair variations occurring at frequencies of > 1% in the population) to date [17]. As these GWAS included millions of SNPs without significant results, investigating pleiotropic and polygenic sequence divergences and the possible co-regulation thereof among candidate genes chosen based on association analyses of phenotypes related to fibrinogen, might prove valuable. There is growing evidence that a significant proportion of the heritability of complex phenotypes, such as fibrinogen, may be explained by a combination of genetic variants, and their combined effects can be calculated in the form of polygenic risk scores [18], making use of a biological filtering approach by taking into account the mechanistic pathways associated with the complex phenotypes. Furthermore, the combination of genetics and molecular biology has greatly facilitated the identification of candidate genes [19, 20]. Multifactorial phenotypes can now be represented as complex interactive networks, which consist of a combination of genetic and non-genetic factors. Therefore, genetic variations in multiple genes in one particular pathway or disease network could lead to synergistic heterozygosity [21]. We have incorporated the concept of synergistic heterozygosity into our hypothesis, through genetic risk score (GRS) analyses, to observe the polygenic effects of harbouring several risk alleles concurrently. In addition, gene expression is generally controlled by transcriptional enhancers, which consist of a cluster of transcription

factor binding sites (TFBS) spaced by spacer sequences and enhancers. SNPs in these regions have potential functional significance, which is not necessarily obvious when observed independently from a functional context [22-24]. Being in regulatory regions, these SNPs may, therefore, associate with the outcome phenotype through polygenic co-regulation rather than being in the causal pathway.

In agreement with international research [25, 26], ethnic differences in fibrinogen concentrations have been reported in the South African population. Black South Africans present with higher total fibrinogen concentrations than their counterparts from European ancestry [27, 28]. A pre-disposition to hypercoagulability has also been observed in the black South African population [29]. Studies on the genetic variation in African individuals have revealed vast genetic diversity in Africa [30, 31]. Furthermore, research conducted in the South African Prospective Urban and Rural Epidemiology (PURE) study population revealed less linkage disequilibrium (LD; neighbouring polymorphisms that are inherited together) in the fibrinogen genes than what has been observed in European populations [32]. The higher variance in fibrinogen concentrations, together with a unique genetic diversity and low LD, presents a promising opportunity to unravel the missing heritability observed for fibrinogen.

To this end, we conducted a candidate gene association study by investigating SNPs of variables associated with fibrinogen-related phenotypes. Candidate genes to be included were further verified through an *in silico* pathway and network analysis to increase our understanding of the polygenic regulation of the fibrinogen phenotype. Polygenetic and pleiotropic co-regulation were explored *via* GRS and TFBS analyses. Protein concentration (both total and γ' fibrinogen) and functionality (plasma clot properties) were used as outcome variables.

RESULTS AND DISCUSSION

The total study population consisted of 1 677 participants included after quality control (QC). Table 1 presents the descriptive characteristics of the study population. Total and γ' fibrinogen concentrations, together with lag time, slope, maximum absorbance and clot lysis time (CLT), were used as phenotype outcomes for all further analyses.

Table 4.1 Descriptive characteristics of the study participants

Variable	Median (25 – 75th percentiles)
Age (years)	48 (41 – 56)
Gender [n (%)]	
Male	628 (37.4)
Female	1049 (62.6)
HIV status [n (%)]	
Positive	269 (16.1)
Negative	1401 (83.9)
Total fibrinogen (g/L)	2.9 (2.2 – 5.0)
Fibrinogen γ' (%)	10.2 (1.2 – 14.7)
Lag time (min)	6.5 (5.1 – 7.8)
Slope ($x10^{-3}$ au/s)	9.0 (6.5 – 12.0)
Maximum absorbance (nm)	0.4 (0.3 – 0.5)
CLT (min)	57.1 (50.9 – 63.9)

CLT = clot lysis time; HIV = human immunodeficiency virus

This study is the first exploration of the pleiotropic and polygenetic co-regulation of candidate SNPs associated with the fibrinogen protein and its functional phenotypes in a large black South African population. Association analyses were conducted to identify appropriate SNPs to include. Table 2 presents the phenotypes that correlated significantly ($r > 0.1$; $p \leq 0.001$) with the fibrinogen variables in this study population. Genes were identified for inclusion in genetic association analyses based on associations of their respective phenotypes with the outcome variables. In addition, the continuous variables presented in Table 2 were treated as covariates and adjusted for in all further analyses. Categorical variables, gender and human immunodeficiency virus (HIV) status were identified as covariates when investigating total and γ' fibrinogen concentrations; however, for CLT we co-varied for gender alone, and for slope and maximum absorbance we only adjusted for HIV status, based on the significance of these associations.

Table 4.2 Phenotypes correlating significantly with total and γ' fibrinogen concentration and clot properties

Total fibrinogen* (g/L)		Fibrinogenγ'* (%)		Clot lysis time* (min)	
Variable	r	Variable	r	Variable	r
Age (years)	0.15	Age (years)	-0.11	BMI (kg/m^2)#	0.28
BMI (kg/m^2)#	0.16	Hcy (µmol/L)	-0.11	Hip C (cm)#	0.27
Hip C (cm)#	0.17	HDL-c (mmol/L)	-0.12	HbA1c (%)	0.22
HbA1c (%)	0.20	PAI-1$_{act}$ (U/mL)	-0.17	TG (mmol/L)$	0.20
LDL-c (mmol/L)	0.13			LDL-c (mmol/L)$	0.17
CRP (mg/L)	0.42			HDL-c (mmol/L)$	0.20
IL-6 (pg/mL)	0.26			PAI-1$_{act}$ (U/mL)	0.31
Lag time* (min)		**Slope* (x10^{-3} au/s)**		**Maximum absorbance* (nm)**	
Variable	r	Variable	r	Variable	r
Age (years)	0.14	PA Index	0.18	Age (years)	0.14
PA Index	-0.18	Hcy (µmol/L)	0.10	BMI (kg/m^2)#	0.10
SBP (mmHg)	0.13	PAI-1$_{act}$ (U/mL)	-0.17	Hip C (cm)#	0.10
		CRP (mg/L)	0.23	CRP (mg/L)	0.26

* $p \leq 0.001$ for all variables; #Adjustments were made only for the highest correlating variable when both BMI and Hip C were correlated to the outcome variable owing to co-linearity; $Only TG and HDL-c adjusted for owing to co-linearity. BMI = body mass index; CRP = C-reactive protein; HbA1c = glycated haemoglobin; Hcy = homocysteine; HDL-c = high-density lipoprotein cholesterol; Hip C = hip circumference; IL-6 = interleukin-6; LDL-c = low-density lipoprotein cholesterol; PAI-1$_{act}$ = plasminogen activator inhibitor type 1 activity; PA index-= physical activity index; SPB = systolic blood pressure; TG = triglycerides

The 13 genes of interest, coding for apolipoprotein-B (Apo-B), apolipoprotein-E (Apo-E), cystathionine beta synthase (CBS), C reactive protein (CRP), factor XIII (FXIII), fibrinogen, low-density lipoprotein receptor (LDL-R), methylenetetrahydrofolate reductase (MTHFR), methionine synthase (MTR), proprotein convertase subtilisin/kexin type 9 (PCSK 9) and plasminogen activator inhibitor type 1 (PAI-1), identified *via* association analyses (Table 2), were uploaded into the Genomatix Pathways System application of the Genomatix Software Suite package. The connectivity between these genes was explored to verify whether they would meaningfully contribute to further statistical models. All the genes included were significantly (13/892, $p = 9.38 \times 10^{-5}$) related to the CVD pathway. The direct and indirect relationships indicative of the shortest route between the genes are depicted in Figure 1. The interconnectivity between these genes demonstrates the rationale for their inclusion to explore the polygenic co-regulation as contributing factors to the fibrinogen phenotypes observed in this study population.

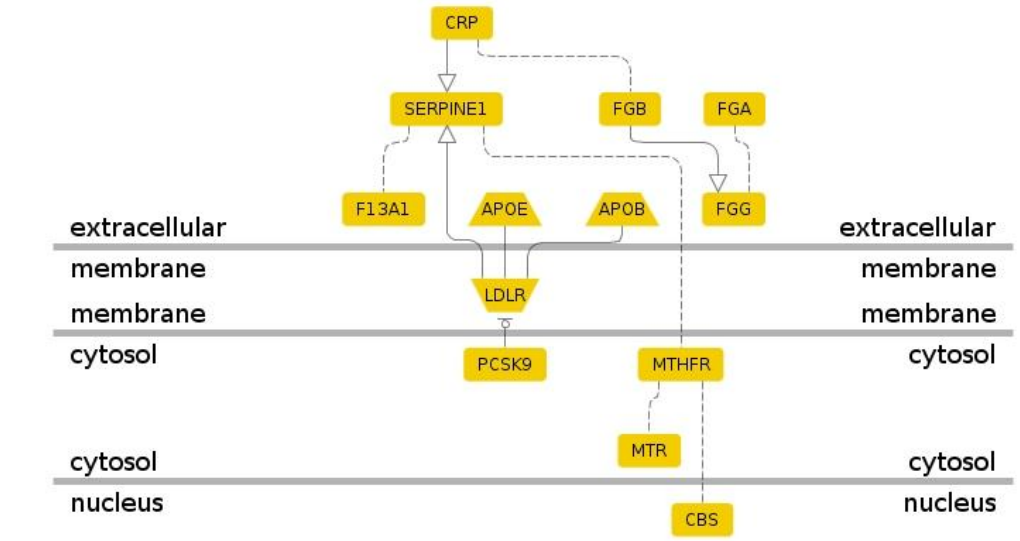


Figure 4.1 The in silico portrayal of the 13 genes included in a CVD network

Indirect relationship
Direct relationship
Suppressive relationship
Directional relationship

Upon evaluation of the results presented in Figure 1 and Table 2, SNP selection commenced (as detailed in the Methods section). One hundred and eight polymorphisms analysed in previous PURE sub-studies were included prior to QC. Four polymorphisms were removed on account of having a minor allele frequency (MAF) < 1%, three SNPs did not adhere to the Hardy-Weinberg equilibrium (HWE), and 15 SNPs were pruned because of linkage with other SNPs in the model ($r^2 > 0.5$; $p < 0.05$). A further eight polymorphisms were removed as a result of insufficient call rates. Final analyses consisted of 78 SNPs with a genotype call rate of 97%, spanning the *APOB*, *APOE*, *CBS*, *CRP*, *F13A1*, *FGA*, *FGB*, *FGG*, *LDL-R*, *MTHFR*, *MTR, PCSK-9 and SERPINE-1* genes. A complete list of polymorphisms included in this study is presented in a supplementary table (Supplementary Table 1).

Individual association of SNPs with fibrinogen variables

Regression analyses were performed for individual phenotype outcomes, with all 78 SNPs and covariates as determined for each phenotype (Table 2) included. The 27 SNPs that presented with significant results, after adjusting for covariates, prior to adjustment for multiple testing, are reported in Table 3. Apart from SNPs directly related to the outcome variables (*i.e*. *FGA*, *FGB*, *FGG* and *F13A1*), loci in the *LDL-R, PCSK-9, CRP* and *CBS* genes also contributed significantly to the fibrinogen phenotypes. After adjusting for multiple testing *PCSK-9*-rs369066144 (which

was not in LD with any of the other *PCSK-9* SNPs) and *CRP*-rs3093068 (also no LD with other *CRP* SNPs) remained significantly associated with CLT (p = 0.04) and γ' fibrinogen (p = 0.02), respectively.

Table 4.3 SNPs significantly associated with the fibrinogen phenotypes

Chr	Gene	SNP	SNP pseudonyms[$]	Location	MAF (%)	β	p
Total fibrinogen (g/L)							
1	*PCSK-9*	rs28362202	c.-26 G>A	promoter exon	7.4	0.41	0.004
1	*PCSK-9*	rs2483205	c.658-7 C>T	intron	47.1	0.17	0.021
1	*PCSK-9*	rs2495477	c.799+3 G>A	intron	27.1	-0.19	0.022
1	*PCSK-9*	rs374941781	c.829 G>C	intron	6.5	0.43	0.003
1	*PCSK-9*	rs584626[#]	c.1354+102 T>C	intron	33.2	0.19	0.013
4	*FGA*	rs2070011	c.-58 G>A	promoter exon	17.0	0.22	0.032
4	*FGG*	rs1049636	c.1299+79 T>C	exon intron	15.6	0.32	0.002
Fibrinogen γ' (%)							
1	*CRP*	rs2808630	g.159711078 A>G	intergenic	14.2	1.00	0.024
1	*CRP*	rs3093068**	g.159711574 C>G	intergenic	37.5	-1.28	0.001
1	*CRP*	rs3093062[#]	c.-409 G>A	promoter intergenic	16.2	-1.11	0.011
1	*CRP*	rs7553007[#]	g.159728759 G>A	intergenic	23.7	1.19	0.001
Lag time (min)							
1	*PCSK-9*	rs499718	c.523+230 C>T	intron	48.5	-1.60	0.025
1	*PCSK-9*	rs505151	c.2009 G>A	exon	30.1	0.18	0.024
6	*F13A1*	rs5985	c.344 G>A	exon	14.2	0.29	0.007
Slope (x10^{-3} au/s)							
1	*PCSK-9*	rs2495482	c.207+15 G>A	intron	13.2	-0.59	0.012
1	*PCSK-9*	rs4927193	c.399+165 T>C	intron	32.4	-0.37	0.020
1	*PCSK-9*	rs28362263	c.1327 A>G	exon intron	3.8	-0.86	0.034
1	*PCSK-9*	rs584626[#]	c.1354+102 T>C	intron	33.2	0.33	0.042
4	*FGA*	rs6050	c.991 A>G	exon intron	30.2	0.41	0.016
19	*LDL-R*	rs2228671	c.81 C>T	exon	3.0	1.15	0.010
19	*LDL-R*	rs3826810	c.*141 G>A	exon	12.9	0.52	0.030
19	*LDL-R*	rs2738465	c.*504 G>A	3' UTR	25.2	0.37	0.033
19	*LDL-R*	rs1433099	c.*666 T>C	3' UTR	47.2	0.31	0.049

Maximum absorbance (nm)							
4	*FGB*	rs4463047	g.154574381 T>C	intergenic	10.5	0.02	0.027
4	*FGA*	rs2070011	c.-58 G>A	promoter exon	17.0	0.02	0.014
19	*LDL-R*	rs17242759	c.67+18 C>A	intron	8.9	0.02	0.045
Clot lysis time (min)							
1	*PCSK-9*	rs494198	c.799+64 A>C	intron	41.9	-0.76	0.043
1	*PCSK-9*	rs369066144**	c.*500 C>T	3'UTR	3.5	-3.61	0.001
21	*CBS*	rs5742905#	c.833 T>C	exon	27.2	-0.97	0.022

$Ensembl release 84 - March 2016, retrieved 27 May 2016 [33] **Remained significant after adjustment for multiple testing; #rs3093062 is in linkage with rs3093058; rs7553007 is in LD with rs1341665, rs2027471, rs1205 and rs2794520; rs584626 is in LD with rs533375 and rs585131; rs5742905 is in LD with *CBS* 844ins68. CBS = cystathionine beta synthase; Chr = chromosome; CRP = C-reactive protein; FGA = fibrinogen alpha chain gene; FGB = fibrinogen beta chain gene; FGG = fibrinogen gamma chain gene; F13A1 = factor XIII; LDL-R = low-density lipoprotein receptor; MAF = minor allele frequency; PCSK-9 = proprotein convertase subtilisin/kexin type 9; rs = reference sequence; SNP = single nucleotide polymorphism; UTR = untranslated region

PCSK-9-rs369066144 (c.500 C>T) was a novel variant identified in the PURE population with a MAF of 3.5%. It is situated in the 3' untranslated region (3'UTR), generally known to be involved in the regulation of gene expression. The functional relevance of this particular variant is still to be determined. Twelve of the 24 *PCSK-9* polymorphisms, together with five of the 19 *LDL-R* polymorphisms under investigation, presented with significant associations with one or more of the outcome phenotypes. Both of these genes encode for proteins involved in blood lipid homeostasis and have been associated with low-density lipoprotein cholesterol (LDL-c) concentrations in the PURE population in which statin use was negligible [34]. *PCSK-9* is the gene coding for the PCSK-9 enzyme that is involved in the degradation of the LDL-R, thereby inhibiting the ability of the LDL-R to remove LDL-c from the bloodstream [35, 36]. The number of significant associations seen between these genes and all fibrinogen phenotypes (apart from γ' fibrinogen) indicates that these lipid mediating genes might provide some insight into the missing heritability observed for fibrinogen and the genetic regulation of clot properties. Literature on the LDL-R and PCSK-9 and the current phenotypes is limited, with only a single study revealing PCSK-9 protein concentrations positively correlated to fibrinogen concentrations [37], and some evidence of an association between LDL-c concentrations and the enhanced onset of clot formation *via* enhanced platelet activation and tissue factor expression [38]. One *LDL-R* polymorphism, rs2228671, has specifically been observed to be an independent predictor of Factor VIII (FVIII), a component of the coagulation cascade, with the T-allele being associated with increased FVIII concentrations [39]. Enhanced functioning of the coagulation cascade could explain the accelerated clot formation indicated by an increase in slope. The current investigation is, to our knowledge, the first exploration of the association of the genetic

determinants of PCSK-9 and LDL-R with fibrinogen and clot properties. Future studies, investigating PCSK-9 and LDL-R concentrations and how they are associated with fibrinogen and its functional aspects, will shed more light on our observations.

As mentioned above, *CRP*-rs3093068 also remained significantly associated with γ' following multiple testing ($p = 0.02$). The fibrinogen γ' phenotype was associated exclusively with four of the included six SNPs of the *CRP* gene. These associations may be explained by co-regulatory processes, rather than being in the fibrinogen causal pathway, as detailed in Section 2.3.

Polygenic association of SNPs with fibrinogen variables

The idea of combining multiple genetic markers into a single score for predicting disease risk has recently become a popular avenue to explore [40]. Generally, individual markers show no detected effect on their own; however, when combined in a score of several genetic factors it has a stronger predictive power for a particular phenotype. With this in mind we made use of a GRS model as a way of measuring the accumulative effect these SNPs have on each phenotype outcome. The model uses a basic weighted polygenic risk score approach where a combination of genetic variants, in this case SNPs that associated significantly with the respective fibrinogen phenotypes (Table 3) prior to adjustment for multiple testing, are evaluated as a combined score allocated to each outcome. The GRS model takes the direction of the β-values obtained above into account, thereby allocating higher scores to the risk allele, not necessarily the minor allele (see section 3.5.1). As the African population has been observed to have a predisposition to hypercoagulability, any allele adding to the risk of hypercoagulability was identified as a risk allele in this population. Being a weighted score, it also considers the magnitude of the contribution of the allele to the phenotype.

Adjustments were made for the specific covariates identified previously (Table 2) for each outcome variable. Table 4 presents the individual contribution of each GRS to the variance of the respective fibrinogen phenotypes, as well as the contribution of the total model. Each GRS was positively associated with its respective fibrinogen phenotype ($p \leq 0.007$) except for maximum absorbance ($p = 0.31$), highlighting the value of the polygenic approach. The use of a GRS is more valuable than investigating SNPs individually, as it considers the accumulative genetic variation present in an individual and calculates the combined effect of a specific set of sequence variations across the genome on the outcome phenotype. Though significant, the fibrinogen phenotypes' variance explained by the GRS, ranged from 4 to 29% only, placing the SNPs' contributions into perspective in terms of other contributing factors such as environmental factors or SNPs not genotyped in this study.

Table 4.4 Contribution of genetic risk scores to phenotypes

Phenotype	GRS β (95% CI)	GRS p-value	GRS total model r^2
Total fibrinogen (g/L)	0.54 (0.34 – 0.75)	< 0.001	0.29
Fibrinogen γ' (%)	0.55 (0.32 – 0.79)	< 0.001	0.04
Lag time (min)	0.12 (0.03 – 0.21)	0.007	0.05
Slope ($x10^{-3}$ au/s)	0.79 (0.54 – 1.05)	< 0.001	0.15
Maximum absorbance (nm)	0.71 (-0.67 – 2.10)	0.31	0.01
CLT (min)	1.07 (0.64 – 1.50)	< 0.001	0.18

CI = confidence interval; CLT = clot lysis time; GRS = genetic risk score

Polygenic co-regulation

In an effort to examine the mechanism behind the cumulative effect observed in the GRS results, we investigated possible co-regulation of the genes included in this study through the gain and loss of TFBS resulting from the SNPs identified. Considering the majority of these variations were not necessarily coding for molecules in the fibrinogen causal pathway, and were found in the non–protein coding regions, the functionality of the particular variation is not as obvious as in the case of the coding region variants. The possible effects of the variation of the non-coding region become evident only when viewed in a particular functional context, such as when investigating TFBS, which the included genes have in common. The gain and loss of TFBS due to SNPs create the possibility for seemingly unrelated genes to be co-regulated *via* the recruitment of a common transcription factor (TF). The importance of the gain and loss of TFBS lies within the resultant secondary signalling response, which follows as a consequence of transcriptional demand. The 27 SNPs in Table 3 gave rise to a total of 75 TFBS changes, of which 35 sequence divergences resulted in a loss of TFBS and 40 formed a new TFBS, hence an overall gain (Supplementary Table 2).

One noticeable feature in terms of fibrinogen concentrations was the loss and gain of the V$IRF4.02 matrix that binds the TF interferon regulatory factor 4 (IRF4). The loss of this TFBS occurs at the *PCSK-9* loci as a result of *PCSK-9*-rs37494178 and the gain of the same V$IRF4.02 matrix at the *FGG* loci was created as a result of *FGG*-rs1049636. The gain and loss of the same TFBS might be indicative of a possible TF interaction between *PCSK-9* and *FGG.* However, the exact mechanism between the TFBS gain and loss and TF recruitment still needs to be elucidated. It is known that the interferon regulatory factor family of proteins has a particularly important function in the regulatory effect of interferons on interferon-inducible genes. IRF4 is a family member specific for lymphocytes and negatively regulates toll-like-

receptor (TLR) signalling that is central to the activation of innate and adaptive immune systems [41, 42]. Recently, several TFs, of which IRF4 is one, have been observed to contribute to cardiac hypertrophy [42]. The role of IRF4 in the context of fibrinogen concentrations and its regulation needs further investigation. Similarly, the V$E2FF family matrix was lost at the *PCSK-9* loci as a result of rs28362202 and gained for *FGA*-rs2070011. The significance of the gain and loss of E2FF in the context of fibrinogen remains unclear. However, E2F TFs such as E2F1 regulate specific biological pathways such as cell responses to inflammation, hypoxia, DNA damage and proliferation [43] and, therefore, might be important in the fibrinogen phenotype context.

In terms of the *CRP* SNPs associated with γ' fibrinogen, the gain of TFBS, V$HIFF, a binding site for hypoxia inducible factor 1 alpha (*HIF1A*) was noted. The induction of *HIF1A* expression is thought to regulate the expression of hepatocyte nuclear factor 4 alpha (*HNF4A*), which activates *HNF1* expression and is followed by an increased expression of γ chain fibrinogen. Therefore, the gain of the V$HIFF site on the *CRP* gene allows the binding of a regulatory component (HIFIA), suggesting a possible co-regulatory pathway for CRP and γ' fibrinogen. HIF1 also plays an essential role in vasculogenesis [44] and has an effect on other acute phase proteins such as interleukin-6 (IL-6). The relationship between *CRP* SNPs and γ' fibrinogen might be modulated through IL-6 and not through CRP directly, since IL-6 stimulates the hepatic production of both fibrinogen and CRP concentrations [45, 46]. As for the phenotypes related to clot properties, no clear mechanism in terms of TFBS was observed. Supplementary Table 2 reports all TFBS gains and losses present for the respective loci.

Although our association results uncovered SNPs that significantly associated with fibrinogen phenotypes, it remains possible that our findings were attributable to other SNPs that were in strong LD with these polymorphisms, which we did not genotype. In addition, the development of a GWAS chip for black South African populations will contribute significantly to the identification of SNPs related to the fibrinogen phenotype, beyond what we were able to determine in this study, including a more extensive array of SNPs involved in the inflammatory pathway being critical in fibrinogen regulation. The concept of the polygenic regulation of fibrinogen phenotypes proved valuable in our study, since it gave a more integrative approach to groupings of SNPs that contributed alone and as a collective to our fibrinogen phenotypes. Future studies gathering transcriptomic data will provide stronger evidence of the possible effects of the gain and/or loss of TFBS explored in this study.

In conclusion, our results revealed not only the regulation of the fibrinogen phenotypes through fibrinogen (*FGA*, *FGB* and *FGG*) and FXIII (*F13A1*) SNPs, but also through SNPs in the *LDL-R*,

PCSK-9, *CBS* and *CRP* genes. The last-mentioned genes have not previously been related to the fibrinogen phenotypes. The accumulating associations of these SNPs as determined through the GRSs proved the importance of a polygenetic approach in determining complex phenotypes and might be of practical importance to an individual's risk in terms of hypercoagulability. Further investigation of TFBS indicated that polygenic transcriptional co-regulation may form the basis of the SNPs' association with the fibrinogen phenotypes. The associations between the SNPs identified here and the fibrinogen phenotypes do not imply any causal link to CVD outcomes, but adds to our current knowledge of the genetic regulation of fibrinogen and may prove useful in explaining more of the variance of the investigated fibrinogen phenotypes.

METHODS

Study population

This is a cross-sectional study embedded in the international multi-centred PURE study [47]. Baseline measurements of the South African arm of the PURE study, consisting of 2 010 self-identified Tswana individuals of the North West province, are reported. Apparently healthy men and women who were between the ages of 35 and 70 and not suffering from any acute or chronic illness were recruited for this study. Data collection was performed in accordance with the Declaration of Helsinki and ethical approval was granted by the ethics committee of the North-West University, South Africa (NWU-00016-10-A1).

Blood sampling and biochemical analyses

Blood samples were collected from the antecubital vein by a registered nurse between 7:00 am and 11:00 am after an overnight fast. Samples were centrifuged within 30 minutes of collection at 2 000 x *g* for 15 minutes. A detailed description of the analyses of the biochemical markers associated with the fibrinogen phenotype, namely IL-6, homocysteine, high-sensitivity CRP, plasminogen activator inhibitor type 1 activity and serum lipids has been published previously [48-50]. The modified Clauss method on the Dade Behring BCS coagulation analyser was used to quantify total fibrinogen (Multifibrin U-test Dade Behring, Deerfield, IL, USA). Fibrinogen γ' concentration was measured with an enzyme-linked immunosorbent assay making use of a 2.G2.H9 mouse monoclonal coating antibody against human γ' fibrinogen (Santa Cruz Biotechnology, Santa Cruz, USA) and a goat polyclonal horseradish peroxidase-conjugated antibody against human fibrinogen (Abcam Cambridge, USA) [6].

Plasma fibrinolytic potential was determined by adding tissue plasminogen activator (tPA) to tissue-factor-induced plasma clots and measuring turbidity with a spectrophotometer [51]. Tissue factor and tPA concentrations were modified slightly for the purpose of obtaining comparable CLTs of about 60-100 minutes. Final concentrations in the clots were: tissue factor diluted 125 times (Dade Innovin, Siemens Healthcare Diagnostics Inc., Marburg, Germany), 17 mmol/l $CaCl_2$, 100 ng/mL tPA (Actilyse, Boehringer Ingelheim, Ingelheim, Germany) and 10 mmol/l phospholipid vesicles (Rossix, Mölndal, Sweden). CLT was calculated as the difference between the time at the midpoint of clear and maximum turbidity (clot formation) and the midpoint between maximum and clear turbidity (clot lysis). In addition to CLT, the following variables were calculated from the turbidity curves: lag time, an indicator of the time required for the activation of the coagulation cascade and for protofibrils to reach sufficient length to allow lateral aggregation; slope, as a representation of the rate of lateral aggregation of fibrin protofibrils; and maximum absorbance, as an indicator of fibre diameter.

Selection of genetic variants

Correlation analyses were used for the identification of phenotypes significantly associated ($r > 0.1$; $p < 0.05$) with one or more of our outcome variables, *i.e.* total and γ' fibrinogen concentrations, lag time, slope, maximum absorbance and CLT. Subsequently polymorphisms previously genotyped in the PURE population, that were located in genes coding for Apo-B, Apo-E, CBS, CRP, FXIII, LDL-R, MTHFR, MTR, PCSK-9 and PAI-1 were included based on these phenotypic associations. Included polymorphisms were selected based on sequencing of the respective genes in 28 individuals in the PURE population. The sequencing data was used to determine genetic variation at the loci involved to establish the viability of genotyping candidate SNPs within this specific population, as well as to identify novel SNPs. In addition, SNPs known from the literature to be associated with the respective phenotypes were also included.

Genotyping

Isolation of genomic deoxyribonucleic acid (DNA) from the leukocyte layer was performed using QIAGEN® Flexigene® DNA extraction kits (QIAGEN® Valencia, CA, USA; catalogue number 51206), and concentrations were determined by the NanoDrop™ spectrophotometer (ND-1000, Wilmington, DE, USA).

Four major genotyping methods were used for the genotyping of the polymorphisms included. The BeadXpress® platform was used to genotype selected loci in the *APOB, APOE, CRP, F13A1, FGA*, *FGB*, *FGG*, *LDL-R* and *PCSK-9* [34, 52, 53] genes. PCR-based restriction fragment length polymorphism (RFLP) analysis was used for the genotyping of the *CBS, MTHFR* and *MTR* polymorphisms [54]. TaqMan-based assays (Thermo Fischer Scientific®, Walthem MA, USA), using the MyIQ5 Bio-Rad® real-time polymerase chain reaction (RT-PCR) thermal cycler (Bio-Rad® Laboratories Inc., Hercules, CA, USA) were used for the genotyping of selected loci within the *FGB*, *SERPINE-1* and *PCSK-9* genes [49, 52]. Competitive allele-specific polymerase chain reaction (KASP) methods with supplies obtained from LGC Limited were used to genotype four additional *FGB* polymorphisms. Supplementary Table 1 presents the genotyping method used for each SNP included.

Statistical analysis

The statistical package for the social sciences (SPSS®) version 23 (IBM® Corp, 2015) and PLINK version 1.9 [55] were used for statistical analyses. Significance was set at a p-value of ≤ 0.05. Kolmogorov-Smirnov testing together with histograms and Q-Q plots were used to evaluate the normality of our data. Most of the variables were not normally distributed and, therefore, non-parametric analyses were performed. Correlations between continuous variables were determined by Spearman rank order correlations. Between-group comparisons for gender and HIV status were performed by conducting Mann-Whitney U tests. The results of these analyses were used to identify the phenotypes associated with the fibrinogen variables.

Quality control and statistical analysis of genetic data

QC included the removal of loci with a MAF below 1%, and failure to meet the requirements of the HWE. Correlated SNPs with LD indicated by r^2 of 0.5 or more where pruned and only one of these genetic variants were used for further statistical analysis. These pruned SNPs were, however, taken into consideration again when determining possible functional effects.

Association tests for the genetic variants that had passed QC were performed through PLINK software. One thousand permutation tests were used to adjust for multiple testing. The variants that presented with significant results prior to multiple testing were then grouped and used to compute a GRS [56]. A weighted method was used to calculate the GRS, which assumes the additive genetic model for each SNP. The additive genetic model performs well when the true genetic model is unknown or wrongly indicated [57]. The genotypes of the selected SNPs were assigned a weight of 0, 1 and 2 depending on the number of minor alleles. This score was then multiplied with the β-value obtained for each variant. The GRS was determined by adding the

scores of the selected SNPs [57, 58]. Linear regression models were used to evaluate the association of the GRS with total and γ' fibrinogen concentrations, lag time, slope, maximum absorbance and CLT, respectively. Both the association and linear regression models for each phenotype outcome were performed without (results not shown) and after inclusion of covariates as determined by the Spearman correlations and Mann-Whitney U tests discussed previously.

In silico network analyses

The Genomatix Software Suite, version 3.6 (Genomatix Software GmbH, Munich, Germany) was applied for the *in silico* network analyses. Genomatix Pathways System was used to explore the network connectivity between the genes of interest and the literature mining extracted from the Genomatic Literature Mining database (Release 02-2016). SNPinspector imbedded in the Genomatix Software Suite package was used to inspect the gain and loss of TFBS and the TFBS information was extracted from the Matrix Library version 9.4 (December 2015) and ElDorado genome database (Release 02-2016). The *in silico* analyses were performed in the default settings as recommended for *Homo sapiens* data mining.

AUTHOR CONTRIBUTION

HTC: isolated the DNA (second round), performed KASP analyses, preparation of files for statistical analyses, statistical analyses using SPSS, wrote the manuscript with CN-R, reviewed and approved the final manuscript.
CN-R: isolated the DNA (first round), performed RFLP analyses, conceptualised the paper, wrote the manuscript with HTC and was involved in the critical review thereof.
LZ: undertook interpretation of the results, gene functionality analyses and interpretation, reviewed and approved the final manuscript.
TC: did statistical analyses using PLINK and SPSS, reviewed and approved the final manuscript.
ZdL: performed the global fibrinolytic assay, critical review of the interpretation of results with regards to clot properties, reviewed and approved the final manuscript.
TvZ: dealt with quality control of the BeadXpress® data, critical review of the interpretation of results with regard to lipid mediators, review and approval of the final manuscript.
MP: supervised the laboratory analyses of all the haemostatic variables, involved in writing the manuscript and critical review thereof.
All authors gave permission for submission to Matrix Biology.

CONFLICT OF INTEREST

The authors declared no conflict of interest.

ACKNOWLEDGEMENTS

This work was supported by grants from the South African National Research Foundation, North-West University (NWU), Population Health Research Institute (PHRI), Medical Research Council, the North West Province Health Department, DNAbiotec Pty Ltd, South African Netherlands Partnerships in Development and the Newton Fund. Funding sources had no involvement in the writing of the report, the collection or interpretation of the data.

We thank all subjects, all supporting staff, the PURE-SA research team, especially Prof Annamarie Kruger, the field workers involved in recruiting the participants, the office staff of the Africa Unit for Transdisciplinary Health Research, the Faculty of Health Sciences, NWU, South Africa, as well as the PURE-International research team, especially Dr Yusuf, and the PURE-study office staff at the PHRI, Hamilton Health Sciences and McMaster University, ON, Canada. We would also like to thank the staff of the Profiles of Resistance to Insulin in Multiple Ethnicities and Regions study from the Centre for Genome Research and DNAbiotec Pty Ltd.

REFERENCES

[1] A. Henschen, F. Lottspeich, M. Kehl, C. Southan, Covalent structure of fibrinogen, Ann. N. Y. Acad. Sci. 408 (1983) 28-43.

[2] C. Wolfenstein-Todel, M.W. Mosesson, Human plasma fibrinogen heterogeneity: evidence for an extended carboxyl-terminal sequence in a normal gamma chain variant (gamma'), Proc. Natl. Acad. Sci. U.S.A. 77 (1980) 5069-5073.

[3] D.W. Chung, E.W. Davie, gamma and gamma' chains of human fibrinogen are produced by alternative mRNA processing, Biochemistry 23 (1984) 4232-4236.

[4] Z. Yang, I. Mochalkin, R.F. Doolittle, A model of fibrin formation based on crystal structures of fibrinogen and fibrin fragments complexed with synthetic peptides, Proc. Natl. Acad. Sci. U.S.A. 97 (2000) 14156-14161.

[5] P. Allan, S. Uitte de Willige, R.H. Abou-Saleh, S.D. Connell, R.A. Ariens, Evidence that fibrinogen gamma' directly interferes with protofibril growth: implications for fibrin structure and clot stiffness, J. Thromb. Haemost. 10 (2012) 1072-1080.

[6] M. Pieters, R.C. Kotze, J.C. Jerling, A. Kruger, R.A. Ariëns, Evidence that fibrinogen γ′ regulates plasma clot structure and lysis and relationship to cardiovascular risk factors in black Africans, Blood 121 (2013) 3254-3260.

[7] K.R. Machlus, J.C. Cardenas, F.C. Church, A.S. Wolberg, Causal relationship between hyperfibrinogenemia, thrombosis, and resistance to thrombolysis in mice, Blood 117 (2011) 4953-4963.

[8] A. Undas, R.A. Ariens, Fibrin clot structure and function: a role in the pathophysiology of arterial and venous thromboembolic diseases, Arterioscler. Thromb. Vasc. Biol. 31 (2011) e88-99.

[9] J. Danesh, S. Lewington, S.G. Thompson, G.D. Lowe, R. Collins, J.B. Kostis, A.C. Wilson, A.R. Folsom, K. Wu, M. Benderly, U. Goldbourt, J. Willeit, S. Kiechl, J.W. Yarnell, P.M. Sweetnam, P.C. Elwood, M. Cushman, B.M. Psaty, R.P. Tracy, A. Tybjaerg-Hansen, F. Haverkate, M.P. de Maat, F.G. Fowkes, A.J. Lee, F.B. Smith, V. Salomaa, K. Harald, R. Rasi, E. Vahtera, P. Jousilahti, J. Pekkanen, R. D'Agostino, W.B. Kannel, P.W. Wilson, G. Tofler, C.L. Arocha-Pinango, A. Rodriguez-Larralde, E. Nagy, M. Mijares, R. Espinosa, E. Rodriquez-Roa, E. Ryder, M.P. Diez-Ewald, G. Campos, V. Fernandez, E. Torres, R. Marchioli, F. Valagussa, A.

Rosengren, L. Wilhelmsen, G. Lappas, H. Eriksson, P. Cremer, D. Nagel, J.D. Curb, B. Rodriguez, K. Yano, J.T. Salonen, K. Nyyssonen, T.P. Tuomainen, B. Hedblad, P. Lind, H. Loewel, W. Koenig, T.W. Meade, J.A. Cooper, B. De Stavola, C. Knottenbelt, G.J. Miller, J.A. Cooper, K.A. Bauer, R.D. Rosenberg, S. Sato, A. Kitamura, Y. Naito, T. Palosuo, P. Ducimetiere, P. Amouyel, D. Arveiler, A.E. Evans, J. Ferrieres, I. Juhan-Vague, A. Bingham, H. Schulte, G. Assmann, B. Cantin, B. Lamarche, J.P. Despres, G.R. Dagenais, H. Tunstall-Pedoe, M. Woodward, Y. Ben-Shlomo, G. Davey Smith, V. Palmieri, J.L. Yeh, A. Rudnicka, P. Ridker, F. Rodeghiero, A. Tosetto, J. Shepherd, I. Ford, M. Robertson, E. Brunner, M. Shipley, E.J. Feskens, D. Kromhout, A. Dickinson, B. Ireland, K. Juzwishin, S. Kaptoge, S. Lewington, A. Memon, N. Sarwar, M. Walker, J. Wheeler, I. White, A. Wood, Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis, JAMA 294 (2005) 1799-1809.

[10] A. Hamsten, U. De Faire, L. Iselius, M. Blombäck, Genetic and cultural inheritance of plasma fibrinogen concentration, Lancet 330 (1987) 988-991.

[11] T. Reed, R.P. Tracy, R.R. Fabsitz, Minimal genetic influences on plasma fibrinogen level in adult males in the NHLBI twin study, Clin. Genet. 45 (1994) 71-77.

[12] Y. Friedlander, Y. Elkana, R. Sinnreich, J.D. Kark, Genetic and environmental sources of fibrinogen variability in Israeli families: the Kibbutzim Family Study, Am. J. Hum. Genet. 56 (1995) 1194.

[13] G. Livshits, G. Schettler, E. Graff, M. Blettner, J. Wahrendorf, D. Brunner, Tel Aviv-Heidelberg three-generation offspring study: genetic determinants of plasma fibrinogen level, Am. J. Med. Genet. 63 (1996) 509-517.

[14] J.S. Pankow, A.R. Folsom, M.A. Province, D. Rao, R.R. Williams, J. Eckfeldt, T.A. Sellers, Segregation analysis of plasminogen activator inhibitor-1 and fibrinogen levels in the NHLBI family heart study, Arterioscler. Thromb. Vasc. Biol. 18 (1998) 1559-1567.

[15] M. de Lange, H. Snieder, R.A.S. Ariëns, T.D. Spector, P.J. Grant, The genetics of haemostasis: a twin study, Lancet 357 (2001) 101-105.

[16] M. Neijts, J. van Dongen, C. Kluft, D.I. Boomsma, G. Willemsen, E.J. de Geus, Genetic architecture of the pro-inflammatory state in an extended twin-family design, Twin Res. Hum. genet. 16 (2013) 931-940.

[17] M. Sabater-Lleal, J. Huang, D. Chasman, S. Naitza, A. Dehghan, A.D. Johnson, A. Teumer, A.P. Reiner, L. Folkersen, S. Basu, A.R. Rudnicka, S. Trompet, A. Malarstig, J. Baumert, J.C. Bis, X. Guo, J.J. Hottenga, S.Y. Shin, L.M. Lopez, J. Lahti, T. Tanaka, L.R. Yanek, T. Oudot-Mellakh, J.F. Wilson, P. Navarro, J.E. Huffman, T. Zemunik, S. Redline, R. Mehra, D. Pulanic, I. Rudan, A.F. Wright, I. Kolcic, O. Polasek, S.H. Wild, H. Campbell, J.D. Curb, R. Wallace, S. Liu, C.B. Eaton, D.M. Becker, L.C. Becker, S. Bandinelli, K. Raikkonen, E. Widen, A. Palotie, M. Fornage, D. Green, M. Gross, G. Davies, S.E. Harris, D.C. Liewald, J.M. Starr, F.M. Williams, P.J. Grant, T.D. Spector, R.J. Strawbridge, A. Silveira, B. Sennblad, F. Rivadeneira, A.G. Uitterlinden, O.H. Franco, A. Hofman, J. van Dongen, G. Willemsen, D.I. Boomsma, J. Yao, N. Swords Jenny, T. Haritunians, B. McKnight, T. Lumley, K.D. Taylor, J.I. Rotter, B.M. Psaty, A. Peters, C. Gieger, T. Illig, A. Grotevendt, G. Homuth, H. Volzke, T. Kocher, A. Goel, M.G. Franzosi, U. Seedorf, R. Clarke, M. Steri, K.V. Tarasov, S. Sanna, D. Schlessinger, D.J. Stott, N. Sattar, B.M. Buckley, A. Rumley, G.D. Lowe, W.L. McArdle, M.H. Chen, G.H. Tofler, J. Song, E. Boerwinkle, A.R. Folsom, L.M. Rose, A. Franco-Cereceda, M. Teichert, M.A. Ikram, T.H. Mosley, S. Bevan, M. Dichgans, P.M. Rothwell, C.L. Sudlow, J.C. Hopewell, J.C. Chambers, D. Saleheen, J.S. Kooner, J. Danesh, C.P. Nelson, J. Erdmann, M.P. Reilly, S. Kathiresan, H. Schunkert, P.E. Morange, L. Ferrucci, J.G. Eriksson, D. Jacobs, I.J. Deary, N. Soranzo, J.C. Witteman, E.J. de Geus, R.P. Tracy, C. Hayward, W. Koenig, F. Cucca, J.W. Jukema, P. Eriksson, S. Seshadri, H.S. Markus, H. Watkins, N.J. Samani, H. Wallaschofski, N.L. Smith, D. Tregouet, P.M. Ridker, W. Tang, D.P. Strachan, A. Hamsten, C.J. O'Donnell, Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease, Circulation 128 (2013) 1310-1324.

[18] F. Dudbridge, Power and predictive accuracy of polygenic risk scores, PLoS Genet. 9 (2013) e1003348.

[19] Y. Li, P. Agarwal, A pathway-based view of human diseases and disease relationships, PloS one 4 (2009) e4346.

[20] B. Childs, D. Valle, Genetics, biology and disease, Annu. Rev. Genomics Hum. Genet. 1 (2000) 1-19.

[21] J. Vockley, P. Rinaldo, M.J. Bennett, D. Matern, G.D. Vladutiu, Synergistic heterozygosity: disease resulting from multiple partial defects in one or more metabolic pathways, Mol. Genet. Metab. 71 (2000) 10-18.

[22] H. Han, H. Shim, D. Shin, J.E. Shim, Y. Ko, J. Shin, H. Kim, A. Cho, E. Kim, T. Lee, TRRUST: a reference database of human transcriptional regulatory interactions, Sci. Rep. 5 (2015).

[23] S.W. Doniger, J.C. Fay, Frequent gain and loss of functional transcription factor binding sites, PLoS Comput. Biol. 3 (2007) e99.

[24] M. Tuğrul, T. Paixão, N.H. Barton, G. Tkačik, Dynamics of transcription factor binding site evolution, PLoS Genet. 11 (2015) e1005639.

[25] S. Kaptoge, I. White, S. Thompson, A. Wood, S. Lewington, G. Lowe, J. Danesh, Associations of plasma fibrinogen levels with established cardiovascular disease risk factors, inflammatory markers, and other characteristics: individual participant meta-analysis of 154,211 adults in 31 prospective studies: the fibrinogen studies collaboration, Am. J. Epidemiology 166 (2007) 867-879.

[26] A.R. Folsom, K. Wu, M. Conlan, A. Finch, C. Davis, G. Marcucci, P. Sorlie, M. Szklo, Distributions of hemostatic variables in blacks and whites: population reference values from the Atherosclerosis Risk in Communities (ARIC) Study, Ethn. Dis. 2 (1991) 35-46.

[27] A. Greyling, M. Pieters, T. Hoekstra, W. Oosthuizen, A. Schutte, Differences in the association of PAI-1 activity with the metabolic syndrome between African and Caucasian women, Nutr. Metab. Cardiovasc. Dis. 17 (2007) 499-507.

[28] M. Pieters, H.H. Vorster, Nutrition and hemostasis: a focus on urbanization in South Africa, Mol. Nutr. Food. Res. 52 (2008) 164-172.

[29] M. Hamer, R. von Känel, M. Reimann, N.T. Malan, A.E. Schutte, H.W. Huisman, L. Malan, Progression of cardiovascular risk factors in black Africans: 3 year follow up of the SABPA cohort study, Atherosclerosis 238 (2015) 52-54.

[30] Y.-S. Chen, A. Torroni, L. Excoffier, A.S. Santachiara-Benerecetti, D.C. Wallace, Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups, Am. J. Hum. Genet. 57 (1995) 133.

[31] S.C. Schuster, W. Miller, A. Ratan, L.P. Tomsho, B. Giardine, L.R. Kasson, R.S. Harris, D.C. Petersen, F. Zhao, J. Qi, Complete Khoisan and Bantu genomes from southern Africa, Nature 463 (2010) 943-947.

[32] R.C. Kotze, C. Nienaber-Rousseau, Z. De Lange, M.P. De Maat, T. Hoekstra, M. Pieters, Genetic polymorphisms influencing total and gamma' fibrinogen levels and fibrin clot properties in Africans, Br. J. Haematol. 168 (2015) 102-112.

[33] A. Yates, W. Akanni, M.R. Amode, D. Barrell, K. Billis, D. Carvalho-Silva, C. Cummins, P. Clapham, S. Fitzgerald, L. Gil, Ensembl 2016, Nucleic acids Res. 44 (2016) D710-D716.

[34] T. van Zyl, J.C. Jerling, K.R. Conradie, E.J. Feskens, Common and rare single nucleotide polymorphisms in the LDLR gene are present in a black South African population and associate with low-density lipoprotein cholesterol levels, J. Hum. Genet. 59 (2014) 88-94.

[35] M. Abifadel, M. Varret, J.-P. Rabès, D. Allard, K. Ouguerram, M. Devillers, C. Cruaud, S. Benjannet, L. Wickham, D. Erlich, Mutations in PCSK9 cause autosomal dominant hypercholesterolemia, Nature genetics 34 (2003) 154-156.

[36] N.G. Seidah, Z. Awan, M. Chrétien, M. Mbikay, PCSK9 A key modulator of cardiovascular health, Circ. Res. 114 (2014) 1022-1036.

[37] Y. Zhang, C.-G. Zhu, R.-X. Xu, S. Li, Y.-L. Guo, J. Sun, J.-J. Li, Relation of circulating PCSK9 concentration to fibrinogen in patients with stable coronary artery disease, J. Clin. Lipidol. 8 (2014) 494-500.

[38] R.S. Rosenson, G.D. Lowe, Effects of lipids and lipoproteins on thrombosis and rheology, Atherosclerosis 140 (1998) 271-280.

[39] N. Martinelli, D. Girelli, B. Lunghi, M. Pinotti, G. Marchetti, G. Malerba, P.F. Pignatti, R. Corrocher, O. Olivieri, F. Bernardi, Polymorphisms at LDLR locus may be associated with coronary artery disease through modulation of coagulation factor VIII activity and independently from lipid profile, Blood 116 (2010) 5688-5697.

[40] S.J. Schrodi, S. Mukherjee, Y. Shan, G. Tromp, J.J. Sninsky, A.P. Callear, T.C. Carter, Z. Ye, J.L. Haines, M.H. Brilliant, Genetic-based prediction of disease traits: prediction is very difficult, especially about the future, Front. Genet. 5 (2014) 1-18.

[41] L. Wang, Z.Q. Yao, J.P. Moorman, Y. Xu, S. Ning, Gene expression profiling identifies IRF4-associated molecular signatures in hematological malignancies, PloS One 9 (2014) e106788.

[42] D.-S. Jiang, Z.-Y. Bian, Y. Zhang, S.-M. Zhang, Y. Liu, R. Zhang, Y. Chen, Q. Yang, X.-D. Zhang, G.-C. Fan, Role of interferon regulatory factor 4 in the regulation of pathological cardiac hypertrophy, Hypertension 61 (2013) 1193-1202.

[43] D. Vaiman, R. Calicchio, F. Miralles, Landscape of transcriptional deregulations in the preeclamptic placenta, PloS One 8 (2013) e65498.

[44] S.-F. Yan, N. Mackman, W. Kisiel, D.M. Stern, D.J. Pinsky, Hypoxia/hypoxemia-induced activation of the procoagulant pathways and the pathogenesis of ischemia-associated thrombosis, Arterioscler. Thromb. Vasc. Biol. 19 (1999) 2029-2035.

[45] C.M. Rein-Smith, N.W. Anderson, D.H. Farrell, Differential regulation of fibrinogen γ chain splice isoforms by interleukin-6, Thromb. Res. 131 (2013) 89-93.

[46] K.S. Alexander, T.E. Madden, D.H. Farrell, Association between γ′ fibrinogen levels and inflammation, J. Thromb. Haemost. 105 (2011) 605.

[47] K. Teo, C.K. Chow, M. Vaz, S. Rangarajan, S. Yusuf, The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries, Am. Heart J. 158 (2009) 1-7. e1.

[48] M. Pieters, M.P. De Maat, J.C. Jerling, T. Hoekstra, A. Kruger, Fibrinogen concentration and its role in CVD risk in black South Africans–effect of urbanisation, J. Thromb. Haemost. 106 (2011) 448-456.

[49] Z. de Lange, D.C. Rijken, T. Hoekstra, K.R. Conradie, J.C. Jerling, M. Pieters, In black South Africans from rural and urban communities, the 4G/5G PAI-1 polymorphism influences PAI-1 activity, but not plasma clot lysis time, PLoS One 8 (2013) e83151.

[50] C. Nienaber-Rousseau, Z. de Lange, M. Pieters, Homocysteine influences blood clot properties alone and in combination with total fibrinogen but not with fibrinogen gamma' in Africans, Blood Coagul. Fibrinolysis 26 (2015) 389-395.

[51] T. Lisman, P.G. de Groot, J.C. Meijers, F.R. Rosendaal, Reduced plasma fibrinolytic potential is a risk factor for venous thrombosis, Blood 105 (2005) 1102-1105.

[52] A. Jobse, M. Pieters, C. Nienaber-Rousseau, H. Boshuizen, T. Hoekstra, M.P.M. de Maat, The contribution of genetic and environmental factors to changes in total and γ' fibrinogen over 5 years, Thromb. Haemost. 135 (2015) 703-709.

[53] C. Nienaber-Rousseau, B. Swanepoel, R.C. Dolman, M. Pieters, K.R. Conradie, G.W. Towers, Interactions between C-reactive protein genotypes with markers of nutritional status in relation to inflammation, Nutrients 6 (2014) 5034-5050.

[54] C. Nienaber-Rousseau, S.M. Ellis, S.J. Moss, A. Melse-Boonstra, G.W. Towers, Gene-environment and gene-gene interactions of specific MTHFR, MTR and CBS gene variants in relation to homocysteine in black South Africans, Gene 530 (2013) 113-118.

[55] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. De Bakker, M.J. Daly, PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (2007) 559-575.

[56] T.V. Perneger, What's wrong with Bonferroni adjustments, BMJ 316 (1998) 1236-1238.

[57] J.A. Smith, E.B. Ware, P. Middha, L. Beacher, S.L. Kardia, Current applications of genetic risk scores to cardiovascular outcomes and subclinical phenotypes, Curr. Epidemiol. Rep. 2 (2015) 180-190.

[58] M.N. Badsi, S. Mediene-Benchekor, H. Ouhaibi-Djellouli, S.A. Lardjam-Hetraf, H. Boulenouar, D.N. Meroufel, X. Hermant, I. Hamani-Medjaoui, N. Saidi-Mehtar, P. Amouyel, L. Houti, A. Meirhaeghe, L. Goumidi, Combined effect of established BMI loci on obesity-related traits in an Algerian population sample, BMC Genetics 15 (2014).

Table 4.5 Supplementary Table 1: SNPs, their minor allele frequencies and the methods used for genotyping

Chr	Gene	SNP pseudonym#	SNP	Genotype frequencies	Minor allele	MAF	HWE (p)	Genotyping method
1	*MTHFR*	*677 C>T*	rs1801133	13/253/1397	T	8.39	0.633	RFLP
1	*PCSK-9*	*c.-26 G>A*	rs28362202	7/234/1429	A	7.43	0.592	BeadXpress®
1	*PCSK-9*	*c.207+15 G>A*	rs2495482	27/385/1251	A	13.20	0.749	BeadXpress®
1	*PCSK-9*	*c.208-161 T>C*	rs2495480	43/486/1140	C	17.14	0.343	BeadXpress®
1	*PCSK-9*	*c.399+165 T>C*	rs4927193	190/706/779	C	32.42	0.119	BeadXpress®
1	*PCSK-9*	*c.400-24 C>T*	rs200109442	0/54/1620	T	1.61	1.000	BeadXpress®
1	*PCSK-9*	*c.523+230 C>T*	rs499718	394/836/443	T	48.54	1.000	BeadXpress®
1	*PCSK-9*	*c.524-68 C>G*	rs624612	6/195/1476	G	6.17	1.000	BeadXpress®
1	*PCSK-9*	*c.657+114 C>A*	rs7552350	137/587/951	A	25.70	0.001	BeadXpress®
1	*PCSK-9*	*c.658-7 C>T*	rs2483205	365/846/463	T	47.07	0.590	BeadXpress®
1	*PCSK-9*	*c.799+3 G>A*	rs2495477	132/644/900	A	27.09	0.266	BeadXpress®
1	*PCSK-9*	*c.799+64 A>C*	rs494198	307/761/573	C	41.90	0.054	BeadXpress®
1	*PCSK-9*	*c.829 C>G*	rs374941781	6/205/1463	G	6.48	0.841	BeadXpress®
1	*PCSK-9*	c.996+379C>T	rs28362256	6/211/1460	T	6.65	0.697	BeadXpress®
1	*PCSK-9*	*c.1274 A>G*	rs28362261	0/34/1643	G	1.01	1.000	BeadXpress®
1	*PCSK-9*	*c.1327 A>G*	rs28362263	4/110/1436	G	3.81	0.279	TaqMan®
1	*PCSK-9*	*c.1354+102 T>C*	rs584626	190/734/752	C	33.23	0.582	BeadXpress®
1	*PCSK-9*	*c.1504-16 C>T*	rs28362269	5/172/1496	T	5.44	1.000	BeadXpress®
1	*PCSK-9*	*c.1658 A>G*	rs28362270	2/115/1556	G	3.56	1.000	BeadXpress®
1	*PCSK-9*	*c.1696 T>C*	rs111563724	3/107/1567	C	3.37	0.434	BeadXpress®

Chr	Gene	SNP pseudonym#	SNP	Genotype frequencies	Minor allele	MAF	HWE (p)	Genotyping method
1	*PCSK-9*	*c.1781 C>A*	rs372506466	6/141/1514	A	4.61	0.159	BeadXpress®
1	*PCSK-9*	*c.2009 G>A*	rs505151	151/687/805	A	30.10	0.815	BeadXpress®
1	*PCSK-9*	*c.2037C>A*	rs28362286	0/72/1328	A	2.57	1.000	TaqMan®
1	*PCSK-9*	*c.*500 C>T*	rs369066144	4/110/1562	T	3.52	0.147	BeadXpress®
1	*PCSK-9*	*c.*571C>T*	rs662145	118/679/875	T	27.36	0.390	BeadXpress®
1	*CRP*	g.159711078 A>G	rs2808630	31/389/1168	G	14.20	0.918	BeadXpress®
1	*CRP*	g.159711574 C>G	rs3093068	212/767/608	G	37.52	0.239	BeadXpress®
1	*CRP*	g.159713301G>A	rs1130864	24/380/1184	A	13.48	0.334	BeadXpress®
1	*CRP*	c.61+29 T>A	rs1417938	0/82/1502	A	2.59	0.623	BeadXpress®
1	*CRP*	c.-409 G>A	rs3093062	31/453/1103	A	16.23	0.052	BeadXpress®
1	*CRP*	g.159728759 G>A	rs7553007	87/578/923	A	23.68	0.835	BeadXpress®
1	*F13A1*	c.344 G>A	rs6003	200/670/710	A	33.86	0.038	TaqMan®
1	*MTR*	c.2756 A>G	rs1805087	85/514/1060	G	20.61	0.035	RFLP
2	*APOB*	c.12541 G>A	rs1042031	17/278/1266	A	9.99	0.673	BeadXpress®
2	*APOB*	g.21014672 G>A	rs673548	42/410/1110	A	15.81	0.569	BeadXpress®
2	*APOB*	c.2817-1223 C>A	rs10199768	19/300/842	A	14.56	0.237	BeadXpress®
2	*APOB*	c.2816+374 A>C	rs3791980	59/485/1014	C	19.35	0.935	BeadXpress®
2	*APOB*	c.693+410 A>G	rs1469513	15/257/1291	G	9.18	0.544	BeadXpress®
2	*APOB*	g.21047682 C>A	rs585967	63/534/964	A	21.14	0.324	BeadXpress®
4	*FGB*	g.154560389 G>A	rs7439150	12/177/1353	A	6.52	0.033	KASP
4	*FGB*	g.154561159 A>G	rs2227385	0/87/1573	G	2.62	0.626	TaqMan®

Chr	Gene	SNP pseudonym#	SNP	Genotype frequencies	Minor allele	MAF	HWE (p)	Genotyping method
4	*FGB*	g.154561868 G>A	rs2227388	39/451/1175	A	15.89	0.647	TaqMan®
4	*FGB*	g.154562157 G>A	rs1800791	10/276/1389	A	8.84	0.447	TaqMan®
4	*FGB*	g.154562556 G>A	rs1800790	5/89/1449	A	3.21	0.018	KASP
4	*FGB*	g.154562762 C>T	rs1800788	3/163/1511	T	5.04	0.796	TaqMan®
4	*FGB*	c.1433 G>A	rs4220	11/231/1341	A	7.99	0.732	BeadXpress®
4	*FGB*	g.154574381 T>C	rs4463047	7/314/1243	C	10.49	0.004	KASP
4	*FGA*	c.991 A>G	rs6050	158/634/782	G	30.18	0.083	BeadXpress®
4	*FGA*	c.-58 G>A	rs2070011	42/452/1087	A	16.95	0.592	BeadXpress®
4	*FGG*	c.1299+79 T>C	rs1049636	47/399/1137	C	15.57	0.104	BeadXpress®
6	*F13A1*	c.103 G>A	rs5985	30/380/1143	A	14.17	0.917	TaqMan®
7	*SERPINE-1*	g.101125148 G>A	rs36228614	26/297/1349	A	10.44	0.049	TaqMan®
19	*LDL-R*	*c.-314 C>T*	rs369850745	0/35/1640	T	1.05	1.000	BeadXpress®
19	*LDL-R*	c.67+18 C>A	rs17242759	9/281/1386	A	8.92	0.229	BeadXpress®
19	*LDL-R*	*c.81 C>T*	rs2228671	3/94/1580	T	2.98	0.181	BeadXpress®
19	*LDL-R*	*c.314-50 T>C*	rs10423288	31/402/1242	C	13.85	0.918	BeadXpress®
19	*LDL-R*	*c.1171 G>A*	rs11669576	38/457/1173	A	15.98	0.465	BeadXpress®
19	*LDL-R*	*c.1187-25 C>T*	rs72658862	7/169/1479	T	5.53	0.342	BeadXpress®
19	*LDL-R*	*c.1413 G>A*	rs5930	34/385/1253	A	13.55	0.466	BeadXpress®
19	*LDL-R*	*c.1586+53 A>G*	rs1569372	287/806/577	G	41.32	0.840	BeadXpress®
19	*LDL-R*	*c.1617 C>T*	rs5929	22/309/1346	T	10.52	0.364	BeadXpress®
19	*LDL-R*	c.1705+56 C>T	rs4508523	332/800/536	T	43.88	0.296	BeadXpress®

Chr	Gene	SNP pseudonym#	SNP	Genotype frequencies	Minor allele	MAF	HWE (p)	Genotyping method
19	*LDL-R*	*c.1773 C>T*	rs688	3/149/1502	T	4.69	1.000	BeadXpress®
19	*LDL-R*	*c.1959 T>C*	rs5925	66/516/1095	C	19.32	0.584	BeadXpress®
19	*LDL-R*	*c.2232 A>G*	rs5927	104/691/876	G	26.90	0.040	BeadXpress®
19	*LDL-R*	*c.2311+9 T>G*	rs369402076	1/75/1598	G	2.30	0.591	BeadXpress®
19	*LDL-R*	*c.*52 G>A*	rs14158	20/306/1335	A	10.42	0.598	BeadXpress®
19	*LDL-R*	*c.*141 G>A*	rs3826810	19/384/1233	A	12.90	0.078	BeadXpress®
19	*LDL-R*	*c.*315 C>G*	rs2738464	193/702/782	G	32.44	0.066	BeadXpress®
19	*LDL-R*	*c.*504 G>A*	rs2738465	113/619/943	A	25.22	0.400	BeadXpress®
19	*LDL-R*	*c.*666 T>C*	rs1433099	354/870/448	C	47.19	0.077	BeadXpress®
19	*APOE*	g.44904820 C>T	rs1081101	21/295/1126	T	11.69	0.702	BeadXpress®
19	*APOE*	g.44905579 G>A	rs405509	45/430/967	A	18.03	0.790	BeadXpress®
19	*APOE*	c.-24+69 G>C	rs440446	10/216/1216	C	8.18	0.861	BeadXpress®
19	*APOE*	c.388 T>C	rs429358	93/586/763	C	26.77	0.179	BeadXpress®
19	*APOE*	c.526 C>T	rs7412	32/392/1247	T	13.64	0.836	BeadXpress®
21	*CBS*	833 T>C	rs5742905	122/662/884	C	27.16	0.951	RFLP
21	*CBS*	---	G9276A	132/669/866	A	27.98	0.856	RFLP

#Ensembl release 84 - March 2016, retrieved 27 May 2016 [33]; A = adenine; Apo-B = apolipoprotein B; Apo-E = apolipoprotein E; CBS = cystathionine beta synthase; CRP = C-reactive protein; C = cytosine; FGA = fibrinogen alpha chain gene; FGB = fibrinogen beta chain gene; FGG= fibrinogen gamma chain gene; FXIII = factor XIII; G = guanine; HWE = Hardy-Weinberg equilibrium; KASP = Competitive allele-specific polymerase chain reaction; LDL-R = low density lipoprotein receptor; MAF = minor allele frequency; MTHFR = methylenetetrahydrofolate reductase; MTR = methionine synthase; PCSK-9 = proprotein convertase subtilisin/kexin type 9; RFLP = restriction fragment length polymorphism; rs = reference sequence; SNP = single nucleotide polymorphism; T = thymine.

Table 4.6 Supplementary Table 2: Single nucleotide polymorphisms and the gain and loss of transcription factor binding sites

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
CBS	rs5742905	251	C>T	loss	V$ZF02	ZNF300.01	KRAB-containing zinc finger protein 300	241	263	-	1	0.993
CBS	rs5742905	251	C>T	gain	V$CAAT	NFY.04	Nuclear factor Y (Y-box binding factor)	244	258	-	1	0.921
CBS	rs5742905	251	C>T	gain	V$GCMF	GCM1.03	Glial cells missing homolog 1 (secondary DNA binding preference)	247	261	-	1	0.853
CRP	rs3093068	501	C>G	loss	V$PLZF	PLZF.02	Promyelocytic leukemia zink finger (TF with nine Krueppel-like zink fingers)	496	510	+	1	0.873
CRP	rs3093068	501	C>G	gain	V$TAIP	CSRNP1.01	Cysteine-serine-rich nuclear protein 1 (AXUD1, AXIN1 up-regulated 1)	500	506	+	1	1.000
CRP	rs3093062	256	A>G	loss	V$MITF	MIT.01	MIT (Microphthalmia transcription factor) and TFE3	249	263	-	1	0.939
CRP	rs3093062	256	A>G	gain	V$EBOX	USF.01	Upstream stimulating factor	247	263	-	1	0.995
CRP	rs3093062	256	A>G	gain	V$HIFF	HRE.03	Hypoxia response elements, binding sites for HIF1alpha/ARNT heterodimers	247	263	+	1	0.972
CRP	rs3093062	256	A>G	gain	V$EBOX	MNT.01	MAX binding protein	248	264	+	1	0.997
CRP	rs3093062	256	A>G	gain	V$HESF	DEC1.02	Basic helix-loop-helix protein known as Dec1, Stra13, Sharp2 or BHLHE40	248	262	-	1	0.923
CRP	rs3093062	256	A>G	gain	V$HIFF	HRE.02	Hypoxia-response elements	248	264	-	1	0.975
CRP	rs3093062	256	A>G	gain	V$HESF	DEC1.02	Basic helix-loop-helix protein known as Dec1, Stra13, Sharp2 or BHLHE40	249	263	+	1	0.976
CRP	rs2808630	501	C>T	gain	V$MEF2	SL1.01	Member of the RSRF (related to serum response factor) protein family from Xenopus laevis	491	513	+	1	0.938

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
CRP	rs2808630	501	C>T	gain	O$PTBP	PTATA.02	Plant TATA box	496	510	-	1	0.923
FA13	rs5985	401	A>G	gain	V$AHRR	AHRARNT.02	Aryl hydrocarbon / Arnt heterodimers, fixed core	389	413	+	1	0.776
FA13	rs5985	401	A>G	gain	V$KLFS	GKLF.03	Gut-enriched Krueppel-like factor / KLF4	392	410	+	1	0.989
FA13	rs5985	401	A>G	gain	V$HASF	HAS.01	HIF-1 ancillary sequence	396	406	-	1	0.923
FGA	rs6050	501	A>G	loss	V$NKXH	NKX31.02	NK3 homeobox 1, NKX-3 ALPHA, BAPX2	492	510	-	1	0.852
FGA	rs6050	501	A>G	loss	V$RU49	RU49.01	Zinc finger transcription factor RU49 (zinc finger proliferation 1 - Zipro 1). RU49 exhibits a strong preference for binding to tandem repeats of the minimal RU49 consensus binding site.	497	503	+	1	0.988
FGA	rs6050	501	A>G	loss	V$RU49	RU49.01	Zinc finger transcription factor RU49 (zinc finger proliferation 1 - Zipro 1). RU49 exhibits a strong preference for binding to tandem repeats of the minimal RU49 consensus binding site.	498	504	-	1	0.983
FGA	rs2070011	501	A>G	gain	V$E2FF	E2F1_DP1.01	E2F-1/DP-1 heterodimeric complex	494	510	+	1	0.882
FGA	rs2070011	501	A>G	gain	V$THAP	THAP1.01	THAP domain containing, apoptosis associated protein	495	505	+	1	0.911
FGA	rs2070011	501	A>G	gain	V$ZF57	ZFP57.01	Krueppel-associated box-containing zinc-finger protein 57 (KRAB-ZFP 57)	495	507	-	1	0.962
FGA	rs2070011	501	A>G	gain	V$E2FF	E2F1_DP1.01	E2F-1/DP-1 heterodimeric complex	494	510	+	1	0.882
FGA	rs2070011	501	A>G	gain	V$THAP	THAP1.01	THAP domain containing, apoptosis associated protein	495	505	+	1	0.911
FGA	rs2070011	501	A>G	gain	V$ZF57	ZFP57.01	Krueppel-associated box-containing zinc-finger protein 57 (KRAB-ZFP 57)	495	507	-	1	0.962

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
FGB	rs1800788	501	C>T	loss	V$ZF10	PRDM14.01	PR domain zinc finger protein 14	495	509	-	1	0.853
FGB	rs1800788	501	C>T	loss	V$HICF	HIC1.02	Hypermethylated in cancer 1 (secondary DNA binding preference)	496	508	+	1	0.974
FGG	rs1049636	501	C>T	loss	V$AP1R	MAFK.01	V-maf musculoaponeurotic fibrosarcoma oncogene homolog K (half site)	486	508	-	1	0.838
FGG	rs1049636	501	C>T	loss	V$SORY	SOX9.08	SRY (sex-determining region Y) box 9, dimeric binding sites	490	512	-	1	0.697
FGG	rs1049636	501	C>T	gain	V$DLXF	DLX3.02	Distal-less homeobox 3	489	507	+	1	0.958
FGG	rs1049636	501	C>T	gain	V$LHXF	LHX4.01	LIM homeobox 4, Gsh4	489	511	+	1	0.830
FGG	rs1049636	501	C>T	gain	V$HOXF	HOXA5.01	Homeobox A5 / Hox-1.3	490	508	-	1	0.848
FGG	rs1049636	501	C>T	gain	V$OCT1	OCT1.06	Octamer-binding factor 1	491	505	+	1	0.819
FGG	rs1049636	501	C>T	gain	V$HBOX	GSH2.01	Homeodomain transcription factor Gsh-2	492	510	+	1	0.961
FGG	rs1049636	501	C>T	gain	V$IRFF	IRF4.01	Interferon regulatory factor (IRF)-related protein (NF-EM5, PIP, LSIRF, ICSAT)	494	518	-	1	0.950
LDL-R	rs2738465	501	A>G	loss	V$HOXC	PBX_HOXA9.01	PBX - HOXA9 binding site	494	510	-	1	0.797
LDL-R	rs2738465	501	A>G	loss	V$GFI1	GFI1B.01	Growth factor independence 1 zinc finger protein Gfi-1B	499	513	+	1	0.886
LDL-R	rs2228671	401	C>G	loss	V$OCT1	OCT.01	Octamer binding site (OCT1/OCT2)	395	409	+	1	0.780
LDL-R	rs2228671	401	C>G	gain	V$MYOD	TCFE2A.02	Transcription factor E2a (E12/E47) (secondary DNA binding preference)	389	405	+	1	0.947
LDL-R	rs2228671	401	C>G	gain	V$HUB1	ZNF282.01	Zinc finger protein 282 (HTLV-I U5 repressive element-binding protein 1)	391	405	-	1	0.775
LDL-R	rs2228671	401	C>G	gain	V$RBPF	RBPJK.02	Mammalian transcriptional repressor RBP-Jkappa/CBF1	395	407	+	1	0.946
LDL-R	rs2228671	401	C>T	loss	V$OCT1	OCT.01	Octamer binding site (OCT1/OCT2)	395	409	+	1	0.780
LDL-R	rs2228671	401	C>T	gain	V$MYOD	TCFE2A.02	Transcription factor E2a (E12/E47) (secondary DNA binding preference)	389	405	+	1	0.943

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
LDL-R	rs2228671	401	C>T	gain	V$RP58	RP58.01	Zinc finger protein RP58 (ZNF238), associated preferentially with heterochromatin	392	404	-	1	0.899
LDL-R	rs2228671	401	C>T	gain	V$GATA	GATA3.02	GATA-binding factor 3	393	405	+	1	0.925
LDL-R	rs2228671	401	C>T	gain	V$CEBP	CEBPB.02	CCAAT/enhancer binding protein beta	394	408	+	1	0.979
LDL-R	rs2228671	401	C>G	loss	V$OCT1	OCT.01	Octamer binding site (OCT1/OCT2)	395	409	+	1	0.780
LDL-R	rs2228671	401	C>G	gain	V$MYOD	TCFE2A.02	Transcription factor E2a (E12/E47) (secondary DNA binding preference)	389	405	+	1	0.947
LDL-R	rs2228671	401	C>G	gain	V$HUB1	ZNF282.01	Zinc finger protein 282 (HTLV-I U5 repressive element-binding protein 1)	391	405	-	1	0.775
LDL-R	rs2228671	401	C>G	gain	V$RBPF	RBPJK.02	Mammalian transcriptional repressor RBP-Jkappa/CBF1	395	407	+	1	0.946
LDL-R	rs2228671	401	C>T	loss	V$OCT1	OCT.01	Octamer binding site (OCT1/OCT2)	395	409	+	1	0.780
LDL-R	rs2228671	401	C>T	gain	V$MYOD	TCFE2A.02	Transcription factor E2a (E12/E47) (secondary DNA binding preference)	389	405	+	1	0.943
LDL-R	rs2228671	401	C>T	gain	V$RP58	RP58.01	Zinc finger protein RP58 (ZNF238), associated preferentially with heterochromatin	392	404	-	1	0.899
LDL-R	rs2228671	401	C>T	gain	V$GATA	GATA3.02	GATA-binding factor 3	393	405	+	1	0.925
LDL-R	rs2228671	401	C>T	gain	V$CEBP	CEBPB.02	CCAAT/enhancer binding protein beta	394	408	+	1	0.979
LDL-R	rs1433099	501	A>G	gain	V$TALE	TGIF2LX.01	TGFB-induced factor homeobox 2-like, X-linked, dimeric binding site	492	508	-	1	0.833
PCSK-9	rs584626	501	A>G	loss	V$HICF	HIC1.01	Hypermethylated in cancer 1	492	504	-	1	0.918
PCSK-9	rs584626	501	A>G	loss	V$EVI1	MEL1.03	MEL1 (MDS1/EVI1-like gene 1) DNA-binding domain 2	497	513	-	1	0.961
PCSK-9	rs584626	501	A>G	gain	V$BRAC	EOMES.02	Eomesodermin, TBR-2 (secondary DNA binding preference)	485	513	+	1	0.887
PCSK-9	rs584626	501	A>G	gain	V$CREB	ATF1.02	Activating transcription factor 1	490	510	-	1	0.936
PCSK-9	rs584626	501	A>G	gain	V$CREB	CREB.02	cAMP-responsive element binding	493	513	-	1	0.967

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
							protein					
PCSK-9	rs584626	501	A>G	gain	V$WHNF	WHN.01	Winged helix protein, involved in hair keratinization and thymus epithelium differentiation	495	505	-	1	0.964
PCSK-9	rs584626	501	A>G	loss	V$HICF	HIC1.01	Hypermethylated in cancer 1	492	504	-	1	0.918
PCSK-9	rs584626	501	A>G	loss	V$EVI1	MEL1.03	MEL1 (MDS1/EVI1-like gene 1) DNA-binding domain 2	497	513	-	1	0.961
PCSK-9	rs584626	501	A>G	gain	V$BRAC	EOMES.02	Eomesodermin, TBR-2 (secondary DNA binding preference)	485	513	+	1	0.887
PCSK-9	rs584626	501	A>G	gain	V$CREB	ATF1.02	Activating transcription factor 1	490	510	-	1	0.936
PCSK-9	rs584626	501	A>G	gain	V$CREB	CREB.02	cAMP-responsive element binding protein	493	513	-	1	0.967
PCSK-9	rs584626	501	A>G	gain	V$WHNF	WHN.01	Winged helix protein, involved in hair keratinization and thymus epithelium differentiation	495	505	-	1	0.964
PCSK-9	rs505151	501	A>G	gain	V$NRSF	NRSF.02	Neuron-restrictive silencer factor (11 bp spacer between half sites)	477	507	+	1	0.754
PCSK-9	rs505151	501	A>G	gain	V$KLFS	GKLF.01	Gut-enriched Krueppel-like factor	490	508	+	1	0.868
PCSK-9	rs499718	264	C>T	loss	V$NFAT	NFAT.01	Nuclear factor of activated T-cells	250	268	-	1	0.970
PCSK-9	rs499718	264	C>T	loss	V$PRDF	BLIMP1.01	Transcriptional repressor B lymphocyte-induced maturation protein-1 (Blimp-1, prdm1)	259	277	-	1	0.818
PCSK-9	rs494198	501	A>C	loss	V$ZF05	ZFP410.01	Zinc finger protein 410, APA-1	492	506	+	1	0.900
PCSK-9	rs494198	501	A>C	loss	V$NF1F	NF1.04	Nuclear factor 1	493	513	-	1	0.910
PCSK-9	rs4927193	501	C>T	gain	V$SORY	SOX21.03	SRY (sex determining region Y)-box 21, dimeric binding sites	489	511	-	1	0.769
PCSK-9	rs37494178	201	G>T	loss	V$IRFF	IRF4.02	Interferon regulatory factor 4	187	211	+	1	0.700
PCSK-9	rs37494178	201	G>T	loss	V$ZFHX	AREB6.04	AREB6 (Atp1a1 regulatory element binding factor 6)	192	204	-	1	0.988
PCSK-9	rs28362263	301	A>G	loss	V$GLIF	GLI3.02	GLI-Kruppel family member GLI3	295	311	+	1	0.908

Gene	SNP	Pos.	Allele change	TFBS Loss/ gain	Matrix family	Matrix	Further Information	Start pos.	End pos.	Strand	Core sim.	Matrix sim.
PCSK-9	rs28362202	256	A>G	loss	V$E2FF	E2F7.01	E2F transcription factor 7	242	258	-	1	0.714
PCSK-9	rs28362202	256	A>G	loss	V$SMAD	SMAD4.01	Smad4 transcription factor involved in TGF-beta signaling	252	262	-	1	0.943
PCSK-9	rs28362202	256	A>G	gain	V$FXRE	FXRE.01	Farnesoid X - activated receptor (RXR/FXR dimer), IR1 sites	255	267	-	1	0.821
PCSK-9	rs2495482	501	C>T	loss	V$EGRF	EGR2.02	Egr-2/Krox-20 early growth response gene product	485	503	-	1	0.964
PCSK-9	rs2495482	501	C>T	loss	V$GCF2	LRRFIP1.01	Leucine rich repeat (in FLII) interacting protein 1	485	503	+	1	0.834
PCSK-9	rs2495482	501	C>T	loss	V$EGRF	WT1.01	Wilms Tumor Suppressor	487	505	-	1	0.937
PCSK-9	rs2495482	501	C>T	loss	V$KLFS	KLF2.01	Krueppel-like factor 2 (lung) (LKLF)	490	508	-	1	0.995
PCSK-9	rs2495482	501	C>T	loss	V$NDPK	NM23.01	NME/NM23 nucleoside diphosphate kinase1 and 2	490	506	-	1	0.916
PCSK-9	rs2495482	501	C>T	loss	V$KLFS	KKLF.01	Kidney-enriched kruppel-like factor, KLF15	495	513	-	1	0.933
PCSK-9	rs2495482	501	C>T	loss	V$SAL2	SALL2.01	Zinc finger protein Spalt-2, sal-like 2, p150(sal2)	495	505	-	1	0.969
PCSK-9	rs2495482	501	C>T	loss	V$GCMF	GCM1.03	Glial cells missing homolog 1 (secondary DNA binding preference)	496	510	+	1	0.844
PCSK-9	rs2495477	501	C>T	loss	V$AP1R	BACH1.01	BTB/POZ-bZIP transcription factor BACH1 forms heterodimers with the small Maf protein family	488	510	-	1	0.833
PCSK-9	rs2495477	501	C>T	loss	V$AP1F	AP1.02	Activator protein 1	493	505	-	1	0.929
PCSK-9	rs2495477	501	C>T	gain	O$VTBP	ATATA.01	Avian C-type LTR TATA box	493	509	-	1	0.803

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS AND RECOMMENDATIONS

5.1 INTRODUCTION

Although family and twin studies have identified the fibrinogen phenotype as heritable, genetic association studies have not been able to allocate a significant proportion of this heritability to common single nucleotide polymorphisms (SNPs). This dissertation sought to investigate the 'missing heritability' of fibrinogen in a highly recombinant African population using three approaches (reviewed in Chapters 1 and 2): firstly a focused investigation of fibrinogen gene polymorphisms and haplotypes hypothesised to be functional; secondly, fibrinogen genotype-interleukin-6 (IL-6) interactions, and lastly, pleiotropic and polygenic co-regulation of candidate SNPs largely outside the fibrinogen gene cluster. These three approaches were investigated in terms of their effect on the outcome phenotypes, total and γ' fibrinogen, as well as fibrin clot properties measured through turbidimetric analysis. Cross-sectional data from the South African arm of the prospective urban and rural epidemiology (PURE) study population was used, as the high fibrinogen and IL-6 concentrations, together with vast genetic diversity and lack of linkage disequilibrium (LD), offered a unique opportunity to explore the missing heritability of the fibrinogen phenotype.

The primary objectives of this study were:

1. To investigate the association of specific fibrinogen polymorphisms and their haplotypes with the fibrinogen phenotypes;

2. To determine the IL-6-interactive effect of polymorphisms and haplotypes within the fibrinogen gene cluster on total and γ' fibrinogen concentration and clot properties;

3. To identify SNPs beyond those in the fibrinogen genes that are associated with the fibrinogen phenotype based on the principles of pleiotropic and polygenic regulation.

These objectives were formulated in an effort to identify SNPs and haplotypes within the fibrinogen genes that are functional contributors to the fibrinogen phenotype independently and through their interaction with the transcriptional enhancer of fibrinogen, IL-6. Furthermore, an attempt was made to locate polymorphic variance outside the fibrinogen gene cluster that contributes to the heritability of fibrinogen through polygenic co-regulation, using a candidate gene approach.

This final chapter briefly captures the main findings of the current investigation under the primary goals addressed by the above-mentioned objectives. Objective-specific results were interpreted and discussed in each manuscript (Chapters 3 and 4) in terms of the encompassing literature, and will not be repeated in this chapter. Rather, the larger contribution of the current study will be put in the context of the broader knowledge economy. Possible study limitations and future research recommendations will be discussed under each particular outcome heading.

5.2 GENETIC COMPOSITION OF THE PURE POPULATION

The genetic composition of the PURE population differed extensively from that previously reported of Europeans. None of the frequently occurring European fibrinogen haplotypes (Mannila *et al.*, 2005; Verschuur *et al.*, 2005; Morozumi *et al.*, 2009) was present in the PURE population. In addition, a SNP located in the fibrinogen beta chain gene, *FGB* -993C/T, prominent in European literature for its relevance to the fibrinogen phenotype (Green, 2001; Verschuur *et al.*, 2005; Morozumi *et al.*, 2009), revealed no allelic variance in the current investigation. In contrast, three novel SNPs, *FGB*-rs2227385, *FGB*-rs2227388 and proprotein convertase subtilisin/kexin type 9 (*PCSK-9)*-rs369066144, were identified in the PURE population and are reported on in this dissertation (Yates *et al.*, 2016).

The PURE population generally presented with lower minor allele frequencies (MAFs) than what has been published globally (Yates *et al.*, 2016). As predicted, the genetic recombination in the fibrinogen gene cluster was significantly higher than what has been described in Europeans (Baumann & Henschen, 1994; Behague *et al.*, 1996; van't Hooft *et al.*, 1999). This same phenomenon was noted in other regions such as the low-density lipoprotein receptor (*LDL-R) and PCSK-9* genes where no complete LD was observed between any of the 27 and 26 investigated SNPs, respectively (Yates *et al.*, 2016).

5.3 FUNCTIONALITY OF FIBRINOGEN SNPS AND HAPLOTYPES IN TERMS OF THEIR ASSOCIATION WITH FIBRINOGEN-RELATED PHENOTYPES

FGB-rs7439150, -1420G/A and -148C/T were the SNPs indicative of functionality in the PURE population. These SNPs had the most pronounced independent and IL-6-interactive associations with fibrinogen concentrations and clot properties. It was additionally demonstrated that in the presence of IL-6, risk alleles across the fibrinogen gene cluster harboured concurrently contributed to fibrinogen concentrations in an additive manner, indicating that multiple SNPs functionally influence the fibrinogen phenotype when they are carried collectively.

The value of conducting an investigation into the fibrinogen gene cluster in Africans was the ability to investigate SNPs independently in an effort to locate functionality that has previously been hindered by tight LD observed in this region in Europeans. The high recombination rates of the PURE study population did enable the investigation of SNPs more independently than ever before, but also resulted in a lack of common haplotypes in the study population. The lack of haplotypes in this highly recombinant study population, therefore, led to the inability to investigate fibrinogen gene haplotypes and their associations with the fibrinogen phenotypes in the current study.

The 14 fibrinogen SNPs that were investigated contributed 0.05% to the variance in fibrinogen concentrations, which is less than what would be expected, considering that 12 of these SNPs were selected based on prior reports of their independent contribution to fibrinogen (Cook *et al,* 2001; Green 2001; Jacquemin *et al.,* 2008; Lim *et al.,* 2003; Mannila *et al.,* 2006; Sabater-Lleal *et al.,* 2013; Uitte de Willige *et al.,* 2005; van't Hooft *et al.,* 1999). Consequently, in terms of future research, the following recommendations are made. Firstly, conducting a heritability study for the fibrinogen phenotype in Africans is a necessity, not simply owing to the lack of replication of previous results, but also owing to differences in genotype frequencies, genetic recombination and haplotypes (section 5.2) observed in the fibrinogen genes of Europeans compared to Africans. These distinct differences reveal that European heritability estimates cannot simply be inferred to Africans. Furthermore, it should be considered that within the South African population large genomic differences exist between well-defined ethnic sub-groups, of which the Tswana population (PURE study population) is only one, emphasising the large research gap that still needs to be addressed (Tishkoff & Williams, 2002; Tishkoff *et al.*, 2009; Teo *et al.*, 2010).

Secondly, future research should investigate genotype-IL-6 interaction differences between basal and acute phase associations with the fibrinogen phenotypes. Some of these interactions lost significance in the current study when individuals with high IL-6 were removed from analyses, suggesting that the interactive effects are more pronounced during the acute phase, or in individuals with inflammatory diseases. Carrying out an experimental or case-control study in an African population will further clarify the results reported in this study. Furthermore, a study population with well-defined chronic inflammation should be investigated to determine the haemostatic risk associated with a chronic up-regulation of acute phase proteins such as fibrinogen, which could have different regulatory mechanisms involved than those previously described as part of the basal and short-term exposure to acute phase regulation.

Thirdly, the current study, although large (n = 2010), was unable to investigate three genotype groups for each variant because of the low MAF of some of the investigated SNPs. A future investigation of the independent, IL-6-interactive and additive effect of fibrinogen polymorphisms in a larger study population would be useful, as it will allow differences in the number of risk alleles to be analysed, rather than only the presence thereof.

5.4 PLEIOTROPIC AND POLYGENIC REGULATION OF THE FIBRINOGEN PHENOTYPES BY SNPS BEYOND THOSE IN THE FIBRINOGEN GENES

The candidate gene analysis offered significant direction to future research endeavours by reporting several novel associations of SNPs outside of the fibrinogen gene cluster, including the association of 12 *PCSK-9* and five *LDL-R* polymorphisms, with the fibrinogen phenotypes. In addition, the co-regulatory transcriptional mechanisms that are fundamental to the observed genotype-phenotype associations were elucidated for the first time in the current study. Lastly, the exploration of the genotypic control of fibrinogen-related phenotypes by genetic risk score (GRS) analyses, as with the additive effect reported in section 5.3, confirmed the necessity of investigating a larger array of SNPs to represent individual risk or genetic contribution more holistically. Overall, this study was able to address more of fibrinogen's missing heritability than genome-wide association studies have, due to the investigation of the cumulative contribution of numerous SNPs in an individual (GRS) rather than the investigation of the per SNP contribution of thousands of individual SNPs in a population. The current study explained 29% of fibrinogen's variance using the GRS composed of seven SNPs, which falls within fibrinogen's heritability estimate of 20 to 51%.

Future research should explore the association of polymorphic variance within the *LDL-R* and *PCSK-9* genes, not only with total and γ' fibrinogen concentrations, but also with clot properties, as SNPs located in these genes contributed significantly to all four the turbidity-derived indicators of fibrin clot properties. The novel *PCSK-9*-rs369066144 polymorphism should be further investigated within the PURE study, and possibly other African population groups, as this SNP made the most significant contribution, and appeared to be protective in terms of thrombotic risk. The latter association and the identification of many novel genotype-phenotype associations within this candidate gene analysis reiterated the urgent need for a whole-genome genotyping method that will capture the functional rare and minor variants in Africans that do not occur elsewhere. Furthermore, the value of investigating the underlying co-regulatory processes driven by common transcription factors that are shared by seemingly unrelated sequence divergences is emphasised in this study. Future research should continue to investigate more than mere genotype-phenotype associations; it should consider their

regulatory effects as well, as the identification of more common co-regulatory pathways could enhance the understanding of the genomic and transcriptional control of complex phenotypes, such as fibrinogen.

The search for variants that contribute to fibrinogen-related phenotypes outside the fibrinogen genes themselves was limited in that some of the variants suggested for investigation by international GWAS, particularly related to genes involved in the inflammatory, haemostatic and immunological pathways, were not included. The current investigation only included genes for which a measurable phenotype outcome was quantified in the PURE population, to be used in the pathway analysis. Future research endeavours within PURE will include a larger array of SNPs and measured phenotypes, including firstly those of Factor XIII (*F13A1)* and IL-1 and 6 and their respective receptors (*IL-6R* and *IL1R1).*

5.5 CONCLUSION

The investigation of the missing heritability of fibrinogen in the PURE population reiterated the value of genetic research in Africans. Novel findings that are able to direct future research efforts include the first investigation of fibrinogen SNPs in a highly recombinant manner in which *FGB*-rs7439150, -1420G/A and -148C/T, were identified as possible functional variants; secondly, an additive effect associated with harbouring several minor allele variants within the fibrinogen gene cluster on fibrinogen concentrations, and thirdly, the involvement of *PCSK-9* and *LDL-R* SNPs in the genetic regulation of fibrinogen-related phenotypes. Lastly, this study described more than genotype-phenotype associations by investigating the co-regulatory transcriptional factors that are reflected in these associations. The value generated by the investigation of polygenic co-regulation in the current study advanced the capacity of genetic research to a level beyond genetic coding, *i.e.* to the transcriptional co-regulation thereof. This dissertation also revealed that polymorphisms are able to exert effects beyond what has been observed in terms of fibrinogen concentrations, *i.e.* through intermediate clot-related effects, not only through their association with fibrinogen, but also independently.

BIBLIOGRAPHY

Albert, M.A., Pare, G., Morris, A., Rose, L., Buring, J., Ridker, P.M. & Zee, R.Y. 2009. Candidate genetic variants in the fibrinogen, methylenetetrahydrofolate reductase, and intercellular adhesion molecule-1 genes and plasma levels of fibrinogen, homocysteine, and intercellular adhesion molecule-1 among various race/ethnic groups: data from the women's genome health study. *American heart journal*, 157(4):777-783.

Albrecht, U., Yang, X., Asselta, R., Keitel, V., Tenchini, M.L., Ludwig, S., Heinrich, P.C., Häussinger, D., Schaper, F. & Bode, J.G. 2007. Activation of NF-κB by IL-1β blocks IL-6-induced sustained STAT3 activation and STAT3-dependent gene expression of the human γ-fibrinogen gene. *Cellular signalling*, 19(9):1866-1878.

Alexander, K.S., Madden, T.E. & Farrell, D.H. 2011. Association between γ' fibrinogen levels and inflammation. *Thrombosis and haemostasis*, 105(4):605-609.

Allan, P., Uitte de Willige, S., Abou-Saleh, R.H., Connell, S.D. & Ariens, R.A. 2012. Evidence that fibrinogen γ' directly interferes with protofibril growth: implications for fibrin structure and clot stiffness. *Journal of thrombosis and haemostasis*, 10(6):1072-1080.

Anderson, G.M., Shaw, A.R. & Shafer, J.A. 1993. Functional characterization of promoter elements involved in regulation of human Bβ-fibrinogen expression. Evidence for binding of novel activator and repressor proteins. *Journal of biological chemistry*, 268(30):22650-22655.

Arbustini, E., Narula, N. & D'Armini, A.M. 2013. Fibrinogen: a circulating factor in search of its genetic architecture. *Circulation*, 128(12):1276-1280.

Ariëns, R. 2013. Fibrin(ogen) and thrombotic disease. *Journal of thrombosis and haemostasis*, 11(S1):294-305.

Ariëns, R.A. 2011. Elevated fibrinogen causes thrombosis. *Blood*, 117(18):4687-4688.

Ariëns, R.A., Lai, T.-S., Weisel, J.W., Greenberg, C.S. & Grant, P.J. 2002. Role of factor XIII in fibrin clot formation and effects of genetic polymorphisms. *Blood*, 100(3):743-754.

Asselta, R., Duga, S., Modugno, M., Malcovati, M. & Tenchini, M.L. 1998. Identification of a glucocorticoid response element in the human γ chain fibrinogen promoter. *Thrombosis and haemostasis*, 79(6):1144-1150.

Barrett, J.C., Fry, B., Maller, J. & Daly, M.J. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics*, 21(2):263-265.

Baumann, R.E. & Henschen, A.H. 1994. Linkage disequilibrium relationships among four polymorphisms within the human fibrinogen gene cluster. *Human genetics*, 94(2):165-170.

Baumert, J., Huang, J., McKnight, B., Sabater-Lleal, M., Steri, M., Chu, A.Y., Trompet, S., Lopez, L.M., Fornage, M. & Teumer, A. 2014. No evidence for genome-wide interactions on plasma fibrinogen by smoking, alcohol consumption and body mass index: results from meta-analyses of 80,607 subjects. *PLOS one*, 9(12):e111156.

Behague, I., Poirier, O., Nicaud, V., Evans, A., Arveiler, D., Luc, G., Cambou, J.-P., Scarabin, P.-Y., Bara, L. & Green, F. 1996. β fibrinogen gene polymorphisms are associated with plasma fibrinogen and coronary artery disease in patients with myocardial infarction the ECTIM study. *Circulation*, 93(3):440-449.

Berg, K. & Kierulf, P. 1989. DNA polymorphisms at fibrinogen loci and plasma fibrinogen concentration. *Clinical genetics*, 36(4):229-235.

Bertram, M.Y., Katzenellenbogen, J., Vos, T., Bradshaw, D. & Hofman, K.J. 2013. The disability adjusted life years due to stroke in South Africa in 2008. *International journal of stroke*, 8(3):76-80.

Bini, A., Fenoglio, J., Mesa-Tejada, R., Kudryk, B. & Kaplan, K.L. 1989. Identification and distribution of fibrinogen, fibrin, and fibrin (ogen) degradation products in atherosclerosis. Use of monoclonal antibodies. *Arteriosclerosis, thrombosis, and vascular biology*, 9(1):109-121.

Blombäck, B. 1996. Fibrinogen and fibrin-proteins with complex roles in hemostasis and thrombosis. *Thrombosis research*, 83(1):1-75.

Blombäck, B., Hessel, B. & Hogg, D. 1976. Disulfide bridges in NH 2-terminal part of human fibrinogen. *Thrombosis research*, 8(5):639-658.

Borissoff, J.I., Spronk, H.M. & ten Cate, H. 2011. The hemostatic system as a modulator of atherosclerosis. *New England journal of medicine*, 364(18):1746-1760.

Bremner, W.F., Sothern, R.B., Kanabrocki, E.L., Ryan, M., McCormick, J.B., Dawson, S., Connors, E.S., Rothschild, R., Third, J.L. & Vahed, S. 2000. Relation between circadian patterns in levels of circulating lipoprotein (a), fibrinogen, platelets, and related lipid variables in men. *American heart journal*, 139(1):164-173.

Brennan, S.O., Davis, R.L., Lowen, R. & Ruskova, A. 2009. Deletion of five residues from the coiled coil of fibrinogen (Bβ Asn167_Glu171del) associated with bleeding and hypodysfibrinogenemia. *Haematologica*, 94(4):585-588.

Brull, D., Dhamrait, S., Moulding, R., Rumley, A., Lowe, G., Humphries, S. & Montgomery, H. 2002. The effect of fibrinogen genotype on fibrinogen levels after strenuous physical exercise. *Thrombosis and haemostasis*, 87(1):37-41.

Burger, M., Mensink, G., Brönstrup, A., Thierfelder, W. & Pietrzik, K. 2004. Alcohol consumption and its relation to cardiovascular risk factors in Germany. *European journal of clinical nutrition*, 58(4):605-614.

Cahill, M., Mistry, R. & Barnett, D. 1992. The human platelet fibrinogen receptor: clinical and therapeutic significance. *British journal of clinical pharmacology*, 33(1):3-9.

Calvete, J.J. 1995. On the structure and function of platelet integrin αIIbβ3, the fibrinogen receptor. *Experimental biology and medicine*, 208(4):346-360.

Campbell, M.C. & Tishkoff, S.A. 2008. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annual review of genomics and human genetics*, 9:403.

Cardon, L.R. & Abecasis, G.R. 2003. Using haplotype blocks to map human complex trait loci. *Trends in Genetics*, 19(3):135-140.

Carlson, C.S., Eberle, M.A., Rieder, M.J., Smith, J.D., Kruglyak, L. & Nickerson, D.A. 2003. Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. *Nature genetics*, 33(4):518-521.

Carter, A.M., Catto, A.J. & Grant, P.J. 1999. Association of the α-fibrinogen Thr312Ala polymorphism with poststroke mortality in subjects with atrial fibrillation. *Circulation*, 99(18):2423-2426.

Carter, A.M., Mansfield, M.W., Stickland, M.H. & Grant, P.J. 1996. β-fibrinogen gene− 455 G/A polymorphism and fibrinogen Levels: risk factors for coronary artery disease in subjects with NIDDM. *Diabetes care*, 19(11):1265-1268.

Carty, C., Heagerty, P., Heckbert, S., Jarvik, G., Lange, L., Cushman, M., Tracy, R. & Reiner, A. 2010. Fibrinogen and IL6 gene variants and IL-6 Levels in relation to plasma fibrinogen concentration and cardiovascular disease risk in the cardiovascular health study. *Annals of human genetics*, 74(1):1-10.

Carty, C.L., Cushman, M., Jones, D., Lange, L.A., Hindorff, L.A., Rice, K., Jenny, N.S., Durda, J.P., Walston, J. & Carlson, C.S. 2008. Associations between common fibrinogen gene polymorphisms and cardiovascular disease in older adults-The cardiovascular health study. *Thrombosis and haemostasis*, 99(2):388-395.

Caspary, E. & Kekwick, R. 1957. Some physicochemical properties of human fibrinogen. *Biochemical journal*, 67(1):41-48.

Chambless, L.E., Folsom, A.R., Davis, V., Sharrett, R., Heiss, G., Sorlie, P., Szklo, M., Howard, G. & Evans, G.W. 2002. Risk factors for progression of common carotid atherosclerosis: the atherosclerosis risk in communities study, 1987–1998. *American journal of epidemiology*, 155(1):38-47.

Chen, Y.-S., Torroni, A., Excoffier, L., Santachiara-Benerecetti, A.S. & Wallace, D.C. 1995. Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups. *American journal of human genetics*, 57(1):133-149.

Cheung, E.Y., Uitte de Willige, S., Vos, H.L., Leebeek, F.W., Dippel, D.W., Bertina, R.M. & de Maat, M.P. 2008. Fibrinogen γ' in ischemic stroke: a case-control study. *Stroke*, 39(3):1033-1035.

Cheung, E.Y., Vos, H.L., Kruip, M.J., den Hertog, H.M., Jukema, J.W. & de Maat, M.P. 2009. Elevated fibrinogen γ′ ratio is associated with cardiovascular diseases and acute phase reaction but not with clinical outcome. *Blood*, 114(20):4603-4604.

Chung, D.W. & Davie, E.W. 1984. γ and γ' chains of human fibrinogen are produced by alternative mRNA processing. *Biochemistry*, 23(18):4232-4236.

Clark, R.A., Lanigan, J.M., DellaPelle, P., Manseau, E., Dvorak, H.F. & Colvin, R.B. 1982. Fibronectin and fibrin provide a provisional matrix for epidermal cell migration during wound reepithelialization. *Journal of investigative dermatology*, 79(5):264-269.

Clauss, A. 1957. Gerinnungsphysiologische schnellmethode zur bestimmung des fibrinogens. *Acta haematologica*, 17(4):237-246.

Collen, D. 1999. The plasminogen (fibrinolytic) system. *Thrombosis and haemostasis*, 82(2):259-270.

Collet, J., Allali, Y., Lesty, C., Tanguy, M., Silvain, J., Ankri, A., Blanchet, B., Dumaine, R., Gianetti, J. & Payot, L. 2006. Altered fibrin architecture is associated with hypofibrinolysis and premature coronary atherothrombosis. *Arteriosclerosis, thrombosis, and vascular biology*, 26(11):2567-2573.

Collet, J., Nagaswami, C., Farrell, D., Montalescot, G. & Weisel, J. 2004. Influence of γ′ fibrinogen splice variant on fibrin physical properties and fibrinolysis rate. *Arteriosclerosis, thrombosis, and vascular biology*, 24(2):382-386.

Collet, J., Park, D., Lesty, C., Soria, J., Soria, C., Montalescot, G. & Weisel, J. 2000. Influence of fibrin network conformation and fibrin fiber diameter on fibrinolysis speed dynamic and structural approaches by confocal microscopy. *Arteriosclerosis, thrombosis, and vascular biology*, 20(5):1354-1361.

Collins, F.S., Guyer, M.S. & Chakravarti, A. 1997. Variations on a theme: cataloging human DNA sequence variation. *Science*, 278(5343):1580-1581.

Connor, J., Fowkes, F., Wood, J., Smith, F., Donnan, P. & Lowe, G. 1992. Genetic variation at fibrinogen loci and plasma fibrinogen levels. *Journal of medical genetics*, 29(7):480.

Cook, D.G., Cappuccio, F.P., Atkinson, R.W., Wicks, P.D., Chitolie, A., Nakandakare, E.R., Sagnella, G.A. & Humphries, S.E. 2001. Ethnic differences in fibrinogen levels: the role of environmental factors and the β-fibrinogen gene. *American journal of epidemiology*, 153(8):799-806.

Cotton, J., Webb, K., Mathur, A., Martin, J. & Humphries, S. 2000. Impact of the-455G> A promoter polymorphism in the β fibrinogen gene on stimulated fibrinogen production following bypass surgery. *Thrombosis and haemostasis*, 84(5):926-927.

Courtois, G., Morgan, J.G., Campbell, L.A., Fourel, G. & Crabtree, G.R. 1987. Interaction of a liver-specific nuclear factor with the fibrinogen and alpha 1-antitrypsin promoters. *Science*, 238(4827):688-692.

Crawford, D.C. & Nickerson, D.A. 2005. Definition and clinical importance of haplotypes. *Annual review of medicie*, 56:303-320.

Cronjé, H.T., Nienaber-Rousseau, C., Zandberg, L., Chikowore, T., de Lange, Z., van Zyl, T. & Pieters, M. 2016 Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation. *Matrix biology* (In press).

Cushman, M., Yanez, D., Psaty, B.M., Fried, L.P., Heiss, G., Lee, M., Polak, J.F., Savage, P.J., Tracy, R.P. & Investigators, C.H.S. 1996. Association of fibrinogen and coagulation factors Vll and VIII with cardiovascular risk factors in the elderly: The cardiovascular health study. *American journal of epidemiology*, 143(7):665-676.

Dahlbäck, B. 2012. Coagulation and inflammation—close allies in health and disease. *Seminars in immunopathology*, 34(1):1-3.

Dalmon, J., Laurent, M. & Courtois, G. 1993. The human β fibrinogen promoter contains a hepatocyte nuclear factor 1-dependent interleukin-6-responsive element. *Molecular and cellular biology,* 13(2):1183-1193.

Daly, M.J., Rioux, J.D., Schaffner, S.F., Hudson, T.J. & Lander, E.S. 2001. High-resolution haplotype structure in the human genome. *Nature genetics*, 29(2):229-232.

Danesh, J., Collins, R., Appleby, P. & Peto, R. 1998. Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: Meta-analyses of prospective studies. *Journal of the American medical association*, 279(18):1477-1482.

Danesh, J., Lewington, S., Thompson, S.G., Lowe, G.D., Collins, R., Kostis, J.B., Wilson, A.C., Folsom, A.R., Wu, K., Benderly, M., Goldbourt, U., Willeit, J., Kiechl, S., Yarnell, J.W., Sweetnam, P.M., Elwood, P.C., Cushman, M., Psaty, B.M., Tracy, R.P., Tybjaerg-Hansen, A., Haverkate, F., de Maat, M.P., Fowkes, F.G., Lee, A.J., Smith, F.B., Salomaa, V., Harald, K., Rasi, R., Vahtera, E., Jousilahti, P., Pekkanen, J., D'Agostino, R., Kannel, W.B., Wilson, P.W., Tofler, G., Arocha-Pinango, C.L., Rodriguez-Larralde, A., Nagy, E., Mijares, M., Espinosa, R., Rodriquez-Roa, E., Ryder, E., Diez-Ewald, M.P., Campos, G., Fernandez, V., Torres, E., Marchioli, R., Valagussa, F., Rosengren, A., Wilhelmsen, L., Lappas, G., Eriksson, H., Cremer, P., Nagel, D., Curb, J.D., Rodriguez, B., Yano, K., Salonen, J.T., Nyyssonen, K., Tuomainen, T.P., Hedblad, B., Lind, P., Loewel, H., Koenig, W., Meade, T.W., Cooper, J.A., De Stavola, B., Knottenbelt, C., Miller, G.J., Cooper, J.A., Bauer, K.A., Rosenberg, R.D., Sato, S., Kitamura, A., Naito, Y., Palosuo, T., Ducimetiere, P., Amouyel, P., Arveiler, D., Evans, A.E., Ferrieres, J., Juhan-Vague, I., Bingham, A., Schulte, H., Assmann, G., Cantin, B., Lamarche, B., Despres, J.P., Dagenais, G.R., Tunstall-Pedoe, H., Woodward, M., Ben-Shlomo, Y., Davey Smith, G., Palmieri, V., Yeh, J.L., Rudnicka, A., Ridker, P., Rodeghiero, F., Tosetto, A., Shepherd, J., Ford, I., Robertson, M., Brunner, E., Shipley, M., Feskens, E.J., Kromhout, D., Dickinson, A., Ireland, B., Juzwishin, K., Kaptoge, S., Lewington, S., Memon, A., Sarwar, N., Walker, M., Wheeler, J., White, I. & Wood, A. 2005. Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis. *Journal of the American medical association,* 294(14):1799-1809.

Danik, J.S., Paré, G., Chasman, D.I., Zee, R.Y., Kwiatkowski, D.J., Parker, A., Miletich, J.P. & Ridker, P.M. 2009. Novel loci, including those related to crohn disease, psoriasis, and inflammation, identified in a genome-wide association study of fibrinogen in 17 686 women: The women's genome health study. *Circulation: Cardiovascular Genetics*, 2(2):134-141.

Davalos, D. & Akassoglou, K. 2012. Fibrinogen as a key regulator of inflammation in disease. *Seminars in immunopathology*, 34(1):43-62.

Davie, E.W., Fujikawa, K. & Kisiel, W. 1991. The coagulation cascade: Initiation, maintenance, and regulation. *Biochemistry*, 30(43):10363-10370.

Davie, E.W. & Ratnoff, O.D. 1964. Waterfall sequence for intrinsic blood clotting. *Science*, 145(3638):1310-1312.

De Lange, M., Snieder, H., Ariëns, R.A.S., Spector, T.D. & Grant, P.J. 2001. The genetics of haemostasis: A twin study. *Lancet*, 357(9250):101-105.

De Maat, M.P. 2001. Effects of diet, drugs, and genes on plasma fibrinogen levels. *Annals of the New York academy of sciences*, 936(1):509-521.

De Maat, M.P., Kastelein, J.J., Jukema, J.W., Zwinderman, A.H., Jansen, H., Groenemeier, B., Bruschke, A.V., Kluft, C. & Group, R. 1998. − 455G/A polymorphism of the β-fibrinogen gene is associated with the progression of coronary atherosclerosis in symptomatic men: Proposed role for an acute-phase reaction pattern of fibrinogen. *Arteriosclerosis, thrombosis, and vascular biology*, 18(2):265-271.

De Vries, P.S., Chasman, D.I., Sabater-Lleal, M., Chen, M.H., Huffman, J.E., Steri, M., Tang, W., Teumer, A., Marioni, R.E., Grossmann, V., Hottenga, J.J., Trompet, S., Muller-Nurasyid, M., Zhao, J.H., Brody, J.A., Kleber, M.E., Guo, X., Wang, J.J., Auer, P.L., Attia, J.R., Yanek, L.R., Ahluwalia, T.S., Lahti, J., Venturini, C., Tanaka, T., Bielak, L.F., Joshi, P.K., Rocanin-Arjo, A., Kolcic, I., Navarro, P., Rose, L.M., Oldmeadow, C., Riess, H., Mazur, J., Basu, S., Goel, A., Yang, Q., Ghanbari, M., Willemsen, G., Rumley, A., Fiorillo, E., de Craen, A.J., Grotevendt, A., Scott, R., Taylor, K.D., Delgado, G.E., Yao, J., Kifley, A., Kooperberg, C., Qayyum, R., Lopez, L.M., Berentzen, T.L., Raikkonen, K., Mangino, M., Bandinelli, S., Peyser, P.A., Wild, S., Tregouet, D.A., Wright, A.F., Marten, J., Zemunik, T., Morrison, A.C., Sennblad, B., Tofler, G., de Maat, M.P., de Geus, E.J., Lowe, G.D., Zoledziewska, M., Sattar, N., Binder, H., Volker, U., Waldenberger, M., Khaw, K.T., McKnight, B., Huang, J., Jenny, N.S., Holliday, E.G., Qi, L., McEvoy, M.G., Becker, D.M., Starr, J.M., Sarin, A.P., Hysi, P.G., Hernandez, D.G., Jhun, M.A., Campbell, H., Hamsten, A., Rivadeneira, F., McArdle, W.L., Slagboom, P.E., Zeller, T., Koenig, W., Psaty, B.M., Haritunians, T., Liu, J., Palotie, A., Uitterlinden, A.G., Stott, D.J., Hofman, A., Franco, O.H., Polasek, O., Rudan, I., Morange, P.E., Wilson, J.F., Kardia, S.L., Ferrucci, L., Spector, T.D., Eriksson, J.G., Hansen, T., Deary, I.J., Becker, L.C., Scott, R.J., Mitchell, P., Marz, W., Wareham, N.J., Peters, A., Greinacher, A., Wild, P.S., Jukema, J.W., Boomsma, D.I., Hayward, C., Cucca, F., Tracy, R., Watkins, H., Reiner, A.P., Folsom, A.R., Ridker, P.M., O'Donnell, C.J., Smith, N.L., Strachan, D.P. & Dehghan, A. 2016. A meta-analysis of 120 246 individuals identifies 18 new loci for fibrinogen concentration. *Human molecular genetics*, 25(2):358-370.

Dehghan, A., Yang, Q., Peters, A., Basu, S., Bis, J.C., Rudnicka, A.R., Kavousi, M., Chen, M.-H., Baumert, J. & Lowe, G.D. 2009. Association of novel genetic loci with circulating fibrinogen levels a genome-wide association study in 6 population-based cohorts. *Circulation: Cardiovascular Genetics*, 2(2):125-133.

Devlin, B. & Risch, N. 1995. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics*, 29(2):311-322.

Dittrich, A., Khouri, C., Sackett, S.D., Ehlting, C., Böhmer, O., Albrecht, U., Bode, J.G., Trautwein, C. & Schaper, F. 2012. Glucocorticoids increase interleukin-6–dependent gene induction by interfering with the expression of the suppressor of cytokine signaling 3 feedback inhibitor. *Hepatology*, 55(1):256-266.

Domingues, M.M., Macrae, F.L., Duval, C., McPherson, H.R., Bridge, K.I., Ajjan, R.A., Ridger, V.C., Connell, S.D., Philippou, H. & Ariëns, R.A. 2016. Thrombin and fibrinogen γ′ impact clot structure by marked effects on intrafibrillar structure and protofibril packing. *Blood*, 127(4):487-495.

Donaldson, D.J., Mahan, J.T., Amrani, D. & Hawiger, J. 1989. Fibrinogen-mediated epidermal cell migration: structural correlates for fibrinogen function. *Journal of cell science*, 94(1):101-108.

Doniger, S.W. & Fay, J.C. 2007. Frequent gain and loss of functional transcription factor binding sites. *PLOS computational biology*, 3(5):e99.

Dotevall, A., Johansson, S. & Wilhelmsen, L. 1994. Association between fibrinogen and other risk factors for cardiovascular disease in men and women results from the Göteborg MONICA Survey 1985. *Annals of epidemiology*, 4(5):369-374.

Drew, A.F., Liu, H., Davidson, J.M., Daugherty, C.C. & Degen, J.L. 2001. Wound-healing defects in mice lacking fibrinogen. *Blood*, 97(12):3691-3698.

Drouet, L., Paolucci, F., Pasqualini, N., Laprade, M., Ripoll, L., Mazoyer, E., Bal dit Sollier, C. & Vanhove, N. 1999. Plasma γ'/γ fibrinogen ratio, a marker of arterial thrombotic activity: A new potential cardiovascular risk factor? *Blood coagulatuion and fibrinolysis*, 10(S1):35-39.

Dudbridge, F. 2013. Power and predictive accuracy of polygenic risk scores. *PLOS genetics*, 9(3):e1003348.

Duperray, A., Languino, L.R., Plescia, J., McDowall, A., Hogg, N., Craig, A.G., Berendt, A.R. & Altieri, D.C. 1997. Molecular identification of a novel fibrinogen binding site on the first domain of ICAM-1 regulating leukocyte-endothelium bridging. *Journal of biological chemistry*, 272(1):435-441.

Ernst, E. & Resch, K.L. 1993. Fibrinogen as a cardiovascular risk factor: a meta-analysis and review of the literature. *Annals of internal medicine*, 118(12):956-963.

Farrell, D.H. 2004. Pathophysiologic roles of the fibrinogen γ chain. *Current opinion in hematology*, 11(3):151-155.

Fatah, K., Hamsten, A., Blombäck, B. & Blombäck, M. 1992. Fibrin gel network characteristics and coronary heart disease: relations to plasma fibrinogen concentration, acute phase protein, serum lipoproteins and coronary atherosclerosis. *Thrombosis and haemostasis*, 68(2):130-135.

Feinbloom, D. & Bauer, K.A. 2005. Assessment of hemostatic risk factors in predicting arterial thrombotic events. *Arteriosclerosis, thrombosis, and vascular biology*, 25(10):2043-2053.

Fish, R.J. & Neerman-Arbez, M. 2012. Fibrinogen gene regulation. *Thrombosis and haemostasis*, 108(3):419-426.

Flint-Garcia, S.A., Thornsberry, J.M. & IV, B. 2003. Structure of linkage disequilibrium in plants. *Annual review of plant biology*, 54(1):357-374.

Folsom, A. 1995. Epidemiology of fibrinogen. *European heart journal*, 16(Suppl A):21-24.

Folsom, A., Cushman, M., Heckbert, S., Ohira, T., Rasmussen-Torvik, L. & Tsai, M. 2007. Factor VII coagulant activity, factor VII− 670A/C and− 402G/A polymorphisms, and risk of venous thromboembolism. *Journal of thrombosis and haemostasis*, 5(8):1674-1678.

Folsom, A.R. 2001. Hemostatic risk factors for atherothrombotic disease: an epidemiologic view. *Thrombosis and haemostasis*, 86(1):366-373.

Folsom, A.R., Wu, K., Conlan, M., Finch, A., Davis, C., Marcucci, G., Sorlie, P. & Szklo, M. 1991. Distributions of hemostatic variables in blacks and whites: population reference values from the atherosclerosis risk in communities (ARIC) study. *Ethnicity & disease*, 2(1):35-46.

Fornace, A., Cummings, D.E., Comeau, C.M., Kant, J.A. & Crabtree, G.R. 1984. Structure of the human γ-fibrinogen gene. Alternate mRNA splicing near the 3'end of the gene produces γA and γB forms of γ-fibrinogen. *Journal of biological chemistry*, 259(20):12826-12830.

Forsyth, C.B., Solovjov, D.A., Ugarova, T.P. & Plow, E.F. 2001. Integrin αMβ2-mediated cell migration to fibrinogen and its recognition peptides. *Journal of experimental medicine*, 193(10):1123-1134.

Friedlander, Y., Elkana, Y., Sinnreich, R. & Kark, J.D. 1995. Genetic and environmental sources of fibrinogen variability in Israeli families: the Kibbutzim family study. *American journal of human genetics*, 56(5):1194-1206.

Frisse, L., Hudson, R., Bartoszewicz, A., Wall, J., Donfack, J. & Di Rienzo, A. 2001. Gene conversion and different population histories may explain the contrast between polymorphism and linkage disequilibrium levels. *American journal of human genetics*, 69(4):831-843.

Fuller, G.M., Otto, J.M., Woloski, B.M., McGary, C.T. & Adams, M.A. 1985. The effects of hepatocyte stimulating factor on fibrinogen biosynthesis in hepatocyte monolayers. *Journal of cell biology*, 101(4):1481-1486.

Fuller, G.M. & Zhang, Z. 2001. Transcriptional control mechanism of fibrinogen gene expression. *Annals of the New York academy of science*, 936(1):469-479.

Gabay, C. & Kushner, I. 1999. Acute-phase proteins and other systemic responses to inflammation. *New England journal of medicine*, 340(6):448-454.

Gabriel, S.B., Schaffner, S.F., Nguyen, H., Moore, J.M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A. & Faggart, M. 2002. The structure of haplotype blocks in the human genome. *Science*, 296(5576):2225-2229.

Gardemann, A., Schwartz, O., Haberbosch, W., Katz, N., Weiss, T., Tillmanns, H., Hehrlein, F., Waas, W. & Eberbach, A. 1997. Positive association of the beta fibrinogen H1/H2 gene

variation to basal fibrinogen levels and to the increase in fibrinogen concentration during acute phase reaction but not to coronary artery disease and myocardial infarction. *Thrombosis and haemostasis*, 77(6):1120-1126.

Gardiner, E.E. & D'Souza, S.E. 1997. A mitogenic action for fibrinogen mediated through intercellular adhesion molecule-1. *Journal of biological chemistry*, 272(24):15474-15480.

Goldstein, D.B. 2001. Islands of linkage disequilibrium. *Nature genetics*, 29(2):109-112.

Gratten, J. & Visscher, P.M. 2016. Genetic pleiotropy in complex traits and diseases: implications for genomic medicine. *Genome medicine*, 8(1):78-80.

Green, D., Ruth, K.J., Folsom, A.R. & Liu, K. 1994. Hemostatic factors in the coronary artery risk development in young adults (CARDIA) study. *Arteriosclerosis, thrombosis, and vascular biology*, 14(5):686-693.

Green, F., Hamsten, A., Blomback, M. & Humphries, S. 1993. The role of β-fibrinogen genotype in determining plasma fibrinogen levels in young survivors of myocardial infarction and healthy controls from Sweden. *Thrombosis and haemostasis*, 70(6):915-920.

Green, F.R. 2001. Fibrinogen polymorphisms and atherothrombotic disease. *Annals of the New York academy of science*, 936(1):549-559.

Greyling, A., Pieters, M., Hoekstra, T., Oosthuizen, W. & Schutte, A. 2007. Differences in the association of PAI-1 activity with the metabolic syndrome between African and caucasian women. *Nutrition, metabolism and cardiovascular diseases*, 17(7):499-507.

Guadiz, G., Sporn, L.A., Goss, R.A., Lawrence, S.O., Marder, V.J. & Simpson-Haidaris, P.J. 1997. Polarized secretion of fibrinogen by lung epithelial cells. *American journal of respiratory cell and molecular biology*, 17(1):60-69.

Hamer, M., von Känel, R., Reimann, M., Malan, N.T., Schutte, A.E., Huisman, H.W. & Malan, L. 2015. Progression of cardiovascular risk factors in black Africans: 3 year follow up of the SABPA cohort study. *Atherosclerosis*, 238(1):52-54.

Hamsten, A., De Faire, U., Iselius, L. & Blombäck, M. 1987. Genetic and cultural inheritance of plasma fibrinogen concentration. *Lancet*, 330(8566):988-991.

Han, H., Shim, H., Shin, D., Shim, J.E., Ko, Y., Shin, J., Kim, H., Cho, A., Kim, E. & Lee, T. 2015. TRRUST: a reference database of human transcriptional regulatory interactions. *Scientific reports*, 5:11432.

Harley, S.L., Sturge, J. & Powell, J.T. 2000. Regulation by fibrinogen and its products of intercellular adhesion molecule-1 expression in human saphenous vein endothelial cells. *Arteriosclerosis, thrombosis, and vascular biology*, 20(3):652-658.

Hedrick, P.W. 1987. Gametic disequilibrium measures: proceed with caution. *Genetics*, 117(2):331-341.

Heinrich, P.C., Castell, J.V. & Andus, T. 1990. Interleukin-6 and the acute phase response. *Biochemical journal*, 265(3):621-636.

Hemker, H., Kop, J. & Willems, G. 1979. Kinetic aspects of the interaction of blood clotting enzymes. VIII. The relation between clotting time and clotting velocity. *Thrombosis and haemostasis*, 41(2):309-313.

Henry, I., Uzan, G., Weil, D., Nicolas, H., Kaplan, J., Marguerie, C., Kahn, A. & Junien, C. 1984. The genes coding for Aα-, Bβ-, and γ-chains of fibrinogen map to 4q2. *American journal of human genetics*, 36(4):760-768.

Henschen-Edman, A.H. 1995. Human fibrinogen occurs as over 1 million non-identical molecules. (*In* Atassi, M.Z. & Appela E., *ed*. Methods in protein structure analysis. New York: Springer. p. 435-443).

Henschen, A., Lottspeich, F., Kehl, M. & Southan, C. 1983. Covalent structure of fibrinogen. *Annals of the New York academy of science*, 408(1):28-43.

Herrick, S., Blanc-Brude, O., Gray, A. & Laurent, G. 1999. Fibrinogen. *International journal of biochemistry and cell biology*, 31(7):741-746.

Hicks, R.C., Golledge, J., Mir-Hasseine, R. & Powell, J.T. 1996. Vasoactive effects of fibrinogen on saphenous vein. *Nature*, .379(6568):818-820.

Hill, W. & Robertson, A. 1968. Linkage disequilibrium in finite populations. *Theoretical and applied genetics*, 38(6):226-231.

Holm, B. & Godal, H. 1984. Quantitation of the three normally occurring plasma fibrinogens in health and during socalled “acute phase” by SDS electro-phoresis of fibrin obtained from EDTA-plasma. *Thrombosis research*, 35(3):279-290.

Holm, B., Nilsen, D. & Godal, H. 1986. Half life and incorporation into venous thrombi of the normally occurring plasma fibrinogen fractions HMW and LMW. *Thrombosis research*, 41(1):57-66.

Holm, B., Nilsen, D., Kierulf, P. & Godal, H. 1985. Purification and characterization of 3 fibrinogens with different molecular weights obtained from normal human plasma. *Thrombosis research*, 37(1):165-176.

Hu, C.-H., Harris, J.E., Davie, E.W. & Chung, D.W. 1995. Characterization of the 5′-flanking region of the gene for the α chain of human fibrinogen. *Journal of biological chemistry*, 270(47):28342-28349.

Huang, S., Cao, Z. & Davie, E. 1993. The role of amino-terminal disulfide bonds in the structure and assembly of human fibrinogen. *Biochemical and biophysical research communications*, 190(2):488-495.

Huffman, J.E., de Vries, P.S., Morrison, A.C., Sabater-Lleal, M., Kacprowski, T., Auer, P.L., Brody, J.A., Chasman, D.I., Chen, M.H., Guo, X., Lin, L.A., Marioni, R.E., Muller-Nurasyid, M., Yanek, L.R., Pankratz, N., Grove, M.L., de Maat, M.P., Cushman, M., Wiggins, K.L., Qi, L., Sennblad, B., Harris, S.E., Polasek, O., Riess, H., Rivadeneira, F., Rose, L.M., Goel, A., Taylor, K.D., Teumer, A., Uitterlinden, A.G., Vaidya, D., Yao, J., Tang, W., Levy, D., Waldenberger, M., Becker, D.M., Folsom, A.R., Giulianini, F., Greinacher, A., Hofman, A., Huang, C.C., Kooperberg, C., Silveira, A., Starr, J.M., Strauch, K., Strawbridge, R.J., Wright, A.F., McKnight, B., Franco, O.H., Zakai, N., Mathias, R.A., Psaty, B.M., Ridker, P.M., Tofler, G.H., Volker, U., Watkins, H., Fornage, M., Hamsten, A., Deary, I.J., Boerwinkle, E., Koenig, W., Rotter, J.I., Hayward, C., Dehghan, A., Reiner, A.P., O'Donnell, C.J. & Smith, N.L. 2015. Rare and low-

frequency variants and their association with plasma levels of fibrinogen, FVII, FVIII, and vWF. *Blood*, 126(11):e19-29.

Humphries, S., Dubowitz, M., Cook, M., Stirling, Y. & Meade, T. 1987. Role of genetic variation at the fibrinogen locus in determination of plasma fibrinogen concentrations. *Lancet*, 329(8548):1452-1455.

Humphries, S., Thomas, A., Montgomery, H., Green, F., Winder, A. & Miller, G. 1997. Gene-environment interaction in the determination of plasma levels of fibrinogen. *Fibrinolysis and proteolysis*, 11(S1):3-7.

Hunter, D.J. 2005. Gene–environment interactions in human diseases. *Nature reviews genetics*, 6(4):287-298.

International hapmap consortium. 2005. A haplotype map of the human genome. Nature, 437(7063):1299-1320.

Jacquemin, B., Antoniades, C., Nyberg, F., Plana, E., Müller, M., Greven, S., Salomaa, V., Sunyer, J., Bellander, T. & Chalamandaris, A.-G. 2008. Common genetic polymorphisms and haplotypes of fibrinogen alpha, beta, and gamma chains affect fibrinogen levels and the response to proinflammatory stimulation in myocardial infarction survivors: the AIRGENE study. *Journal of the American college of cardiology*, 52(11):941-952.

Jeff, J.M., Brown-Gentry, K. & Crawford, D.C. 2015. Identification of gene-gene and gene-environment interactions within the fibrinogen gene cluster for fibrinogen levels in three ethnically diverse populations. *Pacific symposium on biocomputing*, 2015:219-230.

Jeffreys, A.J., Kauppi, L. & Neumann, R. 2001. Intensely punctate meiotic recombination in the class II region of the major histocompatibility complex. *Nature genetics*, 29(2):217-222.

Jensen, T., Halvorsen, S., Godal, H.C., Sandset, P.M. & Skjønsberg, O.H. 2000. Discrepancy between fibrinogen concentrations determined by clotting rate and clottability assays during the acute-phase reaction. *Thrombosis research*, 100(5):397-403.

Jesty, J. & Nemerson, Y. 1974. Purification of factor VII from bovine plasma: Reaction with tissue factor and activation of factor X. *Journal of biological chemistry*, 249(2):509-515.

Johnson, G.C., Esposito, L., Barratt, B.J., Smith, A.N., Heward, J., Di Genova, G., Ueda, H., Cordell, H.J., Eaves, I.A. & Dudbridge, F. 2001. Haplotype tagging for the identification of common disease genes. *Nature genetics*, 29(2):233-237.

Jorde, L.B. 2000. Linkage disequilibrium and the search for complex disease genes. *Genome research*, 10(10):1435-1444.

Kamath, S. & Lip, G. 2003. Fibrinogen: biochemistry, epidemiology and determinants. *Qjm: An international journal of medicine*, 96(10):711-729.

Kannel, W.B., Wolf, P.A., Castelli, W.P. & D'Agostino, R.B. 1987. Fibrinogen and risk of cardiovascular disease: the Framingham study. *Journal of the American medical association*, 258(9):1183-1186.

Kaplan, I.V., Attaelmannan, M. & Levinson, S.S. 2001. Fibrinogen is an antioxidant that protects β-lipoproteins at physiological concentrations in a cell free system. *Atherosclerosis*, 158(2):455-463.

Kaptoge, S., Di Angelantonio, E., Pennells, L., Wood, A.M., White, I.R., Gao, P., Walker, M., Thompson, A., Sarwar, N., Caslake, M., Butterworth, A.S., Amouyel, P., Assmann, G., Bakker, S.J., Barr, E.L., Barrett-Connor, E., Benjamin, E.J., Bjorkelund, C., Brenner, H., Brunner, E., Clarke, R., Cooper, J.A., Cremer, P., Cushman, M., Dagenais, G.R., D'Agostino, R.B., Sr., Dankner, R., Davey-Smith, G., Deeg, D., Dekker, J.M., Engstrom, G., Folsom, A.R., Fowkes, F.G., Gallacher, J., Gaziano, J.M., Giampaoli, S., Gillum, R.F., Hofman, A., Howard, B.V., Ingelsson, E., Iso, H., Jorgensen, T., Kiechl, S., Kitamura, A., Kiyohara, Y., Koenig, W., Kromhout, D., Kuller, L.H., Lawlor, D.A., Meade, T.W., Nissinen, A., Nordestgaard, B.G., Onat, A., Panagiotakos, D.B., Psaty, B.M., Rodriguez, B., Rosengren, A., Salomaa, V., Kauhanen, J., Salonen, J.T., Shaffer, J.A., Shea, S., Ford, I., Stehouwer, C.D., Strandberg, T.E., Tipping, R.W., Tosetto, A., Wassertheil-Smoller, S., Wennberg, P., Westendorp, R.G., Whincup, P.H., Wilhelmsen, L., Woodward, M., Lowe, G.D., Wareham, N.J., Khaw, K.T., Sattar, N., Packard, C.J., Gudnason, V., Ridker, P.M., Pepys, M.B., Thompson, S.G. & Danesh, J. 2012. C-reactive protein, fibrinogen, and cardiovascular disease prediction. *New England journal of medicine*, 367(14):1310-1320.

Kaptoge, S., Thompson, S.G., Danesh, J. & Collaboration, E.R.F. 2013. C-reactive protein, fibrinogen, and cardiovascular risk. *New England journal of medicine*, 368(1):85-86.

Kaptoge, S., White, I., Thompson, S., Wood, A., Lewington, S., Lowe, G. & Danesh, J. 2007. Associations of plasma fibrinogen levels with established cardiovascular disease risk factors, inflammatory markers, and other characteristics: individual participant meta-analysis of 154,211 adults in 31 prospective studies: the fibrinogen studies collaboration. *American journal of epidemiology*, 166(8):867-879.

Kathiresan, S., Yang, Q., Larson, M.G., Camargo, A.L., Tofler, G.H., Hirschhorn, J.N., Gabriel, S.B. & O'Donnell, C.J. 2006. Common genetic variation in five thrombosis genes and relations to plasma hemostatic protein level and cardiovascular disease risk. *Arteriosclerosis, thrombosis, and vascular biology*, 26(6):1405-1412.

Keavney, B., Danesh, J., Parish, S., Palmer, A., Clark, S., Youngman, L., Delépine, M., Lathrop, M., Peto, R. & Collins, R. 2006. Fibrinogen and coronary heart disease: test of causality by 'mendelian randomization'. *International journal of epidemiology*, 35(4):935-943.

Ken-Dror, G., Humphries, S.E., Kumari, M., Kivimaki, M. & Drenos, F. 2012. A genetic instrument for mendelian randomization of fibrinogen. *European journal of epidemiology*, 27(4):267-279.

Koenig, W. 2003. Fibrin(ogen) in cardiovascular disease: an update. *Thrombosis and haemostasis*, 89(4):601-609.

Koenig, W. & Ernst, E. 1992. The possible role of hemorheology in atherothrombogenesis. *Atherosclerosis*, 94(2-3):93-107.

Koster, T., Rosendaal, F., Reitsma, P., Van Der Velden, P., Briet, E. & Vandenbroucke, J. 1994. Factor VII and fibrinogen levels as risk factors for venous thrombosis. A case-control study of plasma levels and DNA polymorphisms--the leiden thrombophilia study (LETS). *Thrombosis and haemostasis*, 71(6):719-722.

Kotzé, R.C., Ariëns, R.A., de Lange, Z. & Pieters, M. 2014. CVD risk factors are related to plasma fibrin clot properties independent of total and or γ'fibrinogen concentration. *Thrombosis research*, 134(5):963-969.

Kotzé, R.C., Nienaber-Rousseau, C., De Lange, Z., De Maat, M.P., Hoekstra, T. & Pieters, M. 2015. Genetic polymorphisms influencing total and gamma' fibrinogen levels and fibrin clot properties in Africans. *British journal of haematology*, 168(1):102-112.

Krobot, K., Hense, H.W., Cremer, P., Eberle, E. & Keil, U. 1992. Determinants of plasma fibrinogen: relation to body weight, waist-to-hip ratio, smoking, alcohol, age, and sex. Results from the second MONICA Augsburg survey 1989-1990. *Arteriosclerosis, thrombosis, and vascular biology*, 12(7):780-788.

Lammertyn, L., Mels, C.M., Pieters, M., Schutte, A.E. & Schutte, R. 2015. Ethnic-specific relationships between haemostatic and oxidative stress markers in black and white South Africans: The SABPA study. *Clinical and experimental hypertension*, 37(6):511-517.

Languino, L.R., Duperray, A., Joganic, K.J., Fornaro, M., Thornton, G.B. & Altieri, D.C. 1995. Regulation of leukocyte-endothelium interaction and leukocyte transendothelial migration by intercellular adhesion molecule 1-fibrinogen recognition. *Proceedings of the national academy of sciences*, 92(5):1505-1509.

Lawrence, S.O. & Simpson-Haidaris, P.J. 2004. Regulated de novo biosynthesis of fibrinogen in extrahepatic epithelial cells in response to inflammation. *Thrombosis and haemostasis*, 92(2):234-243.

Lee, A.J., Lowe, G., Woodward, M. & Tunstall-Pedoe, H. 1993. Fibrinogen in relation to personal history of prevalent hypertension, diabetes, stroke, intermittent claudication, coronary heart disease, and family history: the Scottish Heart Health Study. *British heart journal*, 69(4):338-342.

Lefkovits, J., Plow, E.F. & Topol, E.J. 1995. Platelet glycoprotein IIb/IIIa receptors in cardiovascular medicine. *New England Journal of Medicine*, 332(23):1553-1559.

Lewontin, R. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*, 49(1):49-67.

Lewontin, R. 1988. On measures of gametic disequilibrium. *Genetics*, 120(3):849-852.

Lim, B.C., Ariëns, R.A., Carter, A.M., Weisel, J.W. & Grant, P.J. 2003. Genetic regulation of fibrin structure and function: complex gene-environment interactions may modulate vascular risk. *Lancet*, 361(9367):1424-1431.

Liu, Y., Berthier-Schaad, Y., Fink, N.E., Fallin, M.D., Tracy, R.P., Klag, M.J., Smith, M.W. & Coresh, J. 2005. β-fibrinogen haplotypes and the risk for cardiovascular disease in a dialysis cohort. *American journal of kidney diseases*, 46(1):78-85.

Livshits, G., Schettler, G., Graff, E., Blettner, M., Wahrendorf, J. & Brunner, D. 1996. Tel Aviv-Heidelberg three-generation offspring study: Genetic determinants of plasma fibrinogen level. *American journal of medical genetics*, 63(4):509-517.

Lou, X.J., Boonmark, N.W., Horrigan, F.T., Degen, J.L. & Lawn, R.M. 1998. Fibrinogen deficiency reduces vascular accumulation of apolipoprotein (a) and development of atherosclerosis in apolipoprotein (a) transgenic mice. *Proceedings of the national academy of sciences*, 95(21):12591-12595.

Lovely, R., Moaddel, M. & Farrell, D. 2003. Fibrinogen γ′ chain binds thrombin exosite II. *Journal of thrombosis and haemostasis*, 1(1):124-131.

Lovely, R.S., Falls, L.A., Al-Mondhiry, H.A., Chambers, C.E., Sexton, G.J., Ni, H. & Farrell, D.H. 2002. Association of γA/γ'fibrinogen levels and coronary artery disease. *Thrombosis and haemostasis*, 88(1):26-31.

Lowe, G. 1992. Blood viscosity and cardiovascular disease. *Thrombosis and haemostasis*, 67(5):494-498.

Lowe, G. & Rumley, A. 2001. Fibrinogen and its degradation products as thrombotic risk factors. *Annals of the New York academy of science*, 936(1):560-565.

Lowe, G.D., Rumley, A. & Mackie, I.J. 2004. Plasma fibrinogen. *Annals of clinical biochemistry*, 41(6):430-440.

Lutsey, P., Cushman, M., Steffen, L., Green, D., Barr, R., Herrington, D., Ouyang, P. & Folsom, A. 2006. Plasma hemostatic factors and endothelial markers in four racial/ethnic groups: the MESA study. *Journal of thrombosis and haemostasis*, 4(12):2629-2635.

Machlus, K.R., Cardenas, J.C., Church, F.C. & Wolberg, A.S. 2011. Causal relationship between hyperfibrinogenemia, thrombosis, and resistance to thrombolysis in mice. *Blood*, 117(18):4953-4963.

Macrae, F.L., Domingues, M.M., Casini, A. & Ariëns, R.A. 2016. The (patho)physiology of fibrinogen γ'. *Seminars in thrombosis and hemostasis,* 42(4):344-355.

Mannila, M.N., Eriksson, P., Ericsson, C.-G., Hamsten, A. & Silveira, A. 2006. Epistatic and pleiotropic effects of polymorphisms in the fibrinogen and coagulation factor XIII genes on plasma fibrinogen concentration, fibrin gel structure and risk of myocardial infarction. *Thrombosis and haemostasis*, 95(3):420-427.

Mannila, M.N., Eriksson, P., Lundman, P., Samnegård, A., Boquist, S., Ericsson, C.-G., Tornvall, P., Hamsten, A. & Silveira, A. 2005. Contribution of haplotypes across the fibrinogen gene cluster to variation in risk of myocardial infarction. *Thrombosis and haemostasis*, 93(3):570-577.

Mannila, M.N., Lovely, R., Kazmierczak, S., Eriksson, P., Samnegård, A., Farrell, D., Hamsten, A. & Silveira, A. 2007. Elevated plasma fibrinogen γ′ concentration is associated with myocardial infarction: effects of variation in fibrinogen genes and environmental factors. *Journal of thrombosis and haemostasis*, 5(4):766-773.

Manolio, T.A., Collins, F.S., Cox, N.J., Goldstein, D.B., Hindorff, L.A., Hunter, D.J., McCarthy, M.I., Ramos, E.M., Cardon, L.R. & Chakravarti, A. 2009. Finding the missing heritability of complex diseases. *Nature*, 461(7265):747-753.

Manuck, S.B. & McCaffery, J.M. 2014. Gene-environment interaction. *Annual review of psychology*, 65(1):41-70.

Maresca, G., Di Blasio, A., Marchioli, R. & Di Minno, G. 1999. Measuring plasma fibrinogen to predict stroke and myocardial infarction an update. *Arteriosclerosis, thrombosis, and vascular biology*, 19(6):1368-1377.

Marlar, R.A., Kleiss, A.J. & Griffin, J.H. 1982. An alternative extrinsic pathway of human blood coagulation. *Blood*, 60(6):1353-1358.

McCarthy, J., Parker, A., Salem, R., Moliterno, D., Wang, Q., Plow, E., Rao, S., Shen, G., Rogers, W. & Newby, L. 2004. Large scale association analysis for identification of genes underlying premature coronary heart disease: cumulative perspective from analysis of 111 candidate genes. *Journal of medical genetics*, 41(5):334-341.

McDonagh, J. & Lee, M. 1997. How does hyperfibrinogenemia lead to thrombosis? *Fibrinolysis and proteolysis*, 11(S1):13-17.

Meade, T., Brozovic, M., Chakrabarti, R., Haines, A., Imeson, J., Mellows, S., Miller, G., North, W., Stirling, Y. & Thompson, S. 1986. Haemostatic function and ischaemic heart disease: principal results of the Northwick Park heart study. *Lancet*, 328(8506):533-537.

Meade, T.W., Humphries, S.E. & De Stavola, B.L. 2006. Commentary: fibrinogen and coronary heart disease - test of causality by 'mendelian' randomization by Keavney et al. *International journal of epidemiology*, 35(4):944-947.

Meh, D.A., Siebenlist, K.R. & Mosesson, M.W. 1996. Identification and characterization of the thrombin binding sites on fibrin. *Journal of biological chemistry*, 271(38):23121-23125.

Mills, J.D., Ariëns, R.A., Mansfield, M.W. & Grant, P.J. 2002. Altered fibrin clot structure in the healthy relatives of patients with premature coronary artery disease. *Circulation*, 106(15):1938-1942.

Mizuguchi, J., Hu, C.-H., Cao, Z., Loeb, K.R., Chung, D.W. & Davie, E.W. 1995. Characterization of the 5′-flanking region of the gene for the γ chain of human fibrinogen. *Journal of biological chemistry*, 270(47):28350-28356.

Morozumi, T., Sharma, A. & De Nardin, E. 2009. The functional effects of the− 455G/A polymorphism on the IL-6-induced expression of the β-fibrinogen gene may be due to linkage disequilibrium with other functional polymorphisms. *Immunological investigations*, 38(3-4):311-323.

Mosesson, M. 2005. Fibrinogen and fibrin structure and functions. *Journal of thrombosis and haemostasis*, 3(8):1894-1904.

Mosesson, M., Finlayson, J. & Umfleet, R. 1972. Human fibrinogen heterogeneities III. Identification of γ chain variants. *Journal of biological chemistry*, 247(16):5223-5227.

Mosesson, M.W. 1998. Fibrinogen structure and fibrin clot assembly. *Seminars in thrombosis and hemostasis*, 24(2):169-174.

Mosesson, M.W., Siebenlist, K.R. & Meh, D.A. 2001. The structure and biological features of fibrinogen and fibrin. *Annals of the New York academy of science*, 936(1):11-30.

Mozaffarian, D., Benjamin, E.J., Go, A.S., Arnett, D.K., Blaha, M.J., Cushman, M., Das, S.R., de Ferranti, S., Després, J.-P. & Fullerton, H.J. 2016. Executive summary: Heart disease and stroke statistics - 2016 ppdate: A report from the American heart association. *Circulation*, 133(4):447-454.

Naito, M., Hayashi, T., Kuzuya, M., Funaki, C., Asai, K. & Kuzuya, F. 1990. Effects of fibrinogen and fibrin on the migration of vascular smooth muscle cells in vitro. *Atherosclerosis*, 83(1):9-14.

Nemerson, Y. 1966. The reaction between bovine brain tissue factor and factors VII and X. *Biochemistry*, 5(2):601-608.

Nienaber, C., Pieters, M., Kruger, S.H., Stonehouse, W. & Vorster, H.H. 2008. Overfatness, stunting and physical inactivity are determinants of plasminogen activator inhibitor-1activity, fibrinogen and thrombin–antithrombin complex in African adolescents. *Blood coagulation and fibrinolysis*, 19(5):361-368.

Nishiuma, S., Kario, K., Yakushijin, K., Maeda, M., Murai, R., Matsuo, T., Ikeda, U., Shimada, K. & Matsuo, M. 1998. Genetic variation in the promoter region of the β-fibrinogen gene is associated with ischemic stroke in a Japanese population. *Blood coagulation and fibrinolysis*, 9(4):373-380.

Olinescu, R.M. & Kummerow, F.A. 2001. Fibrinogen is an efficient antioxidant. *Journal of nutritional biochemistry*, 12(3):162-169.

Osterud, B. & Rapaport, S.I. 1977. Activation of factor IX by the reaction product of tissue factor and factor VII: additional pathway for initiating blood coagulation. *Proceedings of the national academy of sciences*, 74(12):5260-5264.

Otto, J.M., Grenett, H.E. & Fuller, G.M. 1987. The coordinated regulation of fibrinogen gene transcription by hepatocyte-stimulating factor and dexamethasone. *Journal of cell biology*, 105(3):1067-1072.

Palmieri, V., Celentano, A., Roman, M.J., de Simone, G., Best, L., Lewis, M.R., Robbins, D.C., Fabsitz, R.R., Howard, B.V. & Devereux, R.B. 2003. Relation of fibrinogen to cardiovascular events is independent of preclinical cardiovascular disease: the strong heart study. *American heart journal*, 145(3):467-474.

Pankow, J.S., Folsom, A.R., Province, M.A., Rao, D., Williams, R.R., Eckfeldt, J. & Sellers, T.A. 1998. Segregation analysis of plasminogen activator inhibitor-1 and fibrinogen levels in the NHLBI family heart study. *Arteriosclerosis, thrombosis, and vascular biology*, 18(10):1559-1567.

Parkes, M., Cortes, A., van Heel, D.A. & Brown, M.A. 2013. Genetic insights into common pathways and complex relationships among immune-mediated diseases. *Nature reviews genetics*, 14(9):661-673.

Patil, N., Berno, A.J., Hinds, D.A., Barrett, W.A., Doshi, J.M., Hacker, C.R., Kautzer, C.R., Lee, D.H., Marjoribanks, C. & McDonough, D.P. 2001. Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. *Science*, 294(5547):1719-1723.

Peerschke, E., Francis, C.W. & Marder, V.J. 1986. Fibrinogen binding to human blood platelets: effect of gamma chain carboxyterminal structure and length. *Blood*, 67(2):385-390.

Pieters, M., De Maat, M.P., Jerling, J.C., Hoekstra, T. & Kruger, A. 2011. Fibrinogen concentration and its role in CVD risk in black South Africans–effect of urbanisation. *Journal of thrombosis and haemostasis*, 106(3):448-456.

Pieters, M. & Vorster, H.H. 2008. Nutrition and hemostasis: a focus on urbanization in South Africa. *Molecular nutrition and food research,* 52(1):164-172.

Pritchard, J.K. & Przeworski, M. 2001. Linkage disequilibrium in humans: models and data. *American journal of human genetics*, 69(1):1-14.

Pulanić, D. & Rudan, I. 2005. The past decade: fibrinogen. *Collegium antropologicum*, 29(1):341-349.

Qizilbash, N., Jones, L., Warlow, C. & Mann, J. 1991. Fibrinogen and lipid concentrations as risk factors for transient ischaemic attacks and minor ischaemic strokes. *British medical journal*, 303(6803):605-609.

Ratnoff, O. & Menzie, C. 1951. A new method for the determination of fibrinogen in small samples of plasma. *Journal of laboratory and clinical medicine*, 37(2):316-320.

Redman, C. & Xia, H. 2000. A review of the expression, assembly, secretion and intracellular degradation of fibrinogen. *Fibrinolysis and proteolysis*, 14(2):198-205.

Reed, T., Tracy, R.P. & Fabsitz, R.R. 1994. Minimal genetic influences on plasma fibrinogen level in adult males in the NHLBI twin study. *Clinical genetics*, 45(2):71-77.

Reich, D.E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P.C., Richter, D.J., Lavery, T., Kouyoumjian, R., Farhadian, S.F. & Ward, R. 2001. Linkage disequilibrium in the human genome. *Nature*, 411(6834):199-204.

Rein-Smith, C.M., Anderson, N.W. & Farrell, D.H. 2013. Differential regulation of fibrinogen γ chain splice isoforms by interleukin-6. *Thrombosis research,* 131(1):89-93.

Reinhart, W.H. 2003. Fibrinogen-marker or mediator of vascular disease? *Vascular medicine*, 8(3):211-216.

Retzinger, G.S., DeAnglis, A.P. & Patuto, S.J. 1998. Adsorption of fibrinogen to droplets of liquid hydrophobic phases: functionality of the bound protein and biological implications. *Arteriosclerosis, thrombosis, and vascular biology*, 18(12):1948-1957.

Riedel, T., Suttnar, J., Brynda, E., Houska, M., Medved, L. & Dyr, J.E. 2011. Fibrinopeptides A and B release in the process of surface fibrin formation. *Blood*, 117(5):1700-1706.

Risch, N. & Merikangas, K. 1996. The future of genetic studies of complex human diseases. *Science*, 273(5281):1516-1517.

Ritchie, D.G. & Fuller, G.M. 1983. Hepatocyte‐stimulating factor: a monocyte‐derived acute‐phase regulatory protein. *Annals of the New York academy of science*, 408(1):490-502.

Roche, H.M. & Mensink, R.P. 2003. Molecular aspects of nutrition. (In Lanham-New, S.A. *et al*., *ed*. Nutrition and metabolism. 2nd ed. Chichester: Wiley-Blackwell. P.7-29)..

Rosing, J., Tans, G., Govers-Riemslag, J., Zwaal, R. & Hemker, H.C. 1980. The role of phospholipids and factor Va in the prothrombinase complex. *Journal of biological chemistry*, 255(1):274-283.

Roy, S., Overton, O. & Redman, C. 1994. Overexpression of any fibrinogen chain by Hep G2 cells specifically elevates the expression of the other two chains. *Journal of biological chemistry*, 269(1):691-695.

Roy, S.N., Mukhopadhyay, G. & Redman, C. 1990. Regulation of fibrinogen assembly. transfection of Hep G2 cells with Bβ cDNA specifically enhances synthesis of the three component chains of fibrinogen. *Journal of biological chemistry*, 265(11):6389-6393.

Rubel, C., Fernández, G.C., Dran, G., Bompadre, M.B., Isturiz, M.A. & Palermo, M.S. 2001. Fibrinogen promotes neutrophil activation and delays apoptosis. *Journal of immunology*, 166(3):2002-2010.

Sabater-Lleal, M., Huang, J., Chasman, D., Naitza, S., Dehghan, A., Johnson, A.D., Teumer, A., Reiner, A.P., Folkersen, L., Basu, S., Rudnicka, A.R., Trompet, S., Malarstig, A., Baumert, J., Bis, J.C., Guo, X., Hottenga, J.J., Shin, S.Y., Lopez, L.M., Lahti, J., Tanaka, T., Yanek, L.R., Oudot-Mellakh, T., Wilson, J.F., Navarro, P., Huffman, J.E., Zemunik, T., Redline, S., Mehra, R., Pulanic, D., Rudan, I., Wright, A.F., Kolcic, I., Polasek, O., Wild, S.H., Campbell, H., Curb, J.D., Wallace, R., Liu, S., Eaton, C.B., Becker, D.M., Becker, L.C., Bandinelli, S., Raikkonen, K., Widen, E., Palotie, A., Fornage, M., Green, D., Gross, M., Davies, G., Harris, S.E., Liewald, D.C., Starr, J.M., Williams, F.M., Grant, P.J., Spector, T.D., Strawbridge, R.J., Silveira, A., Sennblad, B., Rivadeneira, F., Uitterlinden, A.G., Franco, O.H., Hofman, A., van Dongen, J., Willemsen, G., Boomsma, D.I., Yao, J., Swords Jenny, N., Haritunians, T., McKnight, B., Lumley, T., Taylor, K.D., Rotter, J.I., Psaty, B.M., Peters, A., Gieger, C., Illig, T., Grotevendt, A.,

Homuth, G., Volzke, H., Kocher, T., Goel, A., Franzosi, M.G., Seedorf, U., Clarke, R., Steri, M., Tarasov, K.V., Sanna, S., Schlessinger, D., Stott, D.J., Sattar, N., Buckley, B.M., Rumley, A., Lowe, G.D., McArdle, W.L., Chen, M.H., Tofler, G.H., Song, J., Boerwinkle, E., Folsom, A.R., Rose, L.M., Franco-Cereceda, A., Teichert, M., Ikram, M.A., Mosley, T.H., Bevan, S., Dichgans, M., Rothwell, P.M., Sudlow, C.L., Hopewell, J.C., Chambers, J.C., Saleheen, D., Kooner, J.S., Danesh, J., Nelson, C.P., Erdmann, J., Reilly, M.P., Kathiresan, S., Schunkert, H., Morange, P.E., Ferrucci, L., Eriksson, J.G., Jacobs, D., Deary, I.J., Soranzo, N., Witteman, J.C., de Geus, E.J., Tracy, R.P., Hayward, C., Koenig, W., Cucca, F., Jukema, J.W., Eriksson, P., Seshadri, S., Markus, H.S., Watkins, H., Samani, N.J., Wallaschofski, H., Smith, N.L., Tregouet, D., Ridker, P.M., Tang, W., Strachan, D.P., Hamsten, A. & O'Donnell, C.J. 2013. Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease. *Circulation*, 128(12):1310-1324.

Sahni, A. & Francis, C.W. 2000. Vascular endothelial growth factor binds to fibrinogen and fibrin and stimulates endothelial cell proliferation. *Blood*, 96(12):3772-3778.

Sahni, A., Odrljin, T. & Francis, C.W. 1998. Binding of basic fibroblast growth factor to fibrinogen and fibrin. *Journal of biological chemistry*, 273(13):7554-7559.

Sakao, E., Ishihara, A., Horikawa, K., Akiyama, M., Arai, M., Kato, M., Seki, N., Fukunaga, K., Shimizu-Yabe, A. & Iwase, K. 2003. Two-peaked synchronization in day/night expression rhythms of the fibrinogen gene cluster in the mouse liver. *Journal of biological chemistry*, 278(33):30450-30457.

Schadt, E.E., Molony, C., Chudin, E., Hao, K., Yang, X., Lum, P.Y., Kasarskis, A., Zhang, B., Wang, S. & Suver, C. 2008. Mapping the genetic architecture of gene expression in human liver. *PLOS biology*, 6(5):e107.

Scheraga, H.A. & Laskowski, M. 1957. The fibrinogen-fibrin conversion. *Advances in protein chemistry*, 12:1-131.

Schuster, S.C., Miller, W., Ratan, A., Tomsho, L.P., Giardine, B., Kasson, L.R., Harris, R.S., Petersen, D.C., Zhao, F. & Qi, J. 2010. Complete Khoisan and Bantu genomes from southern Africa. *Nature*, 463(7283):943-947.

Scrutton, M., Ross-Murphy, S., Bennett, G., Stirling, Y. & Meade, T. 1994. Changes in clot deformability-a possible explanation for the epidemiological association between plasma fibrinogen concentration and myocardial infarction. *Blood coagulation & fibrinolysis*, 5(5):719-723.

Siebenlist, K.R., Mosesson, M.W., Hernandez, I., Bush, L.A., Di Cera, E., Shainoff, J.R., Di Orio, J.P. & Stojanovic, L. 2005. Studies on the basis for the properties of fibrin produced from fibrinogen-containing γ' chains. *Blood*, 106(8):2730-2736.

Sinha, S., Luben, R.N., Welch, A., Bingham, S., Wareham, N.J., Day, N.E. & Khaw, K.-T. 2005. Fibrinogen and cigarette smoking in men and women in the European prospective investigation into cancer in Norfolk (EPIC-Norfolk) population. *European journal of cardiovascular prevention & rehabilitation*, 12(2):144-150.

Smith, A., Patterson, C., Yarnell, J., Rumley, A., Ben-Shlomo, Y. & Lowe, G. 2005. Which hemostatic markers add to the predictive value of conventional risk factors for coronary heart disease and ischemic stroke? The caerphilly study. *Circulation*, 112(20):3080-3087.

Smith, G.F. 1980. Fibrinogen–fibrin conversion. The mechanism of fibrin-polymer formation in solution. *Biochemical journal*, 185(1):1-11.

Solovieff, N., Cotsapas, C., Lee, P.H., Purcell, S.M. & Smoller, J.W. 2013. Pleiotropy in complex traits: challenges and strategies. *Nature reviews genetics*, 14(7):483-495.

Souto, J.C., Almasy, L., Borrell, M., Garí, M., Martínez, E., Mateo, J., Stone, W.H., Blangero, J. & Fontcuberta, J. 2000. Genetic determinants of hemostasis phenotypes in Spanish families. *Circulation*, 101(13):1546-1551.

Spiel, A.O., Gilbert, J.C. & Jilma, B. 2008. Von Willebrand factor in cardiovascular disease focus on acute coronary syndromes. *Circulation*, 117(11):1449-1459.

Sporn, L.A., Bunce, L.A. & Francis, C.W. 1995. Cell proliferation on fibrin: modulation by fibrinopeptide cleavage. *Blood*, 86(5):1802-1810.

Standeven, K.F., Ariëns, R.A. & Grant, P.J. 2005. The molecular physiology and pathology of fibrin structure/function. *Blood reviews*, 19(5):275-288.

Stats SA (Statistics South Africa). 2015. Mortality and causes of death in South Africa, 2014: findings from death notification. http://www.statssa.gov.za/publications/P03093/P030932014.pdf Date of access: 31 Oct.2016.

Stec, J.J., Silbershatz, H., Tofler, G.H., Matheney, T.H., Sutherland, P., Lipinska, I., Massaro, J.M., Wilson, P.F., Muller, J.E. & D'Agostino, R.B. 2000. Association of fibrinogen with cardiovascular risk factors and cardiovascular disease in the Framingham offspring population. *Circulation*, 102(14):1634-1638.

Swaim, W.R. & Feders, M.B. 1967. Fibrinogen assay. *Clinical chemistry*, 13(11):1026-1028.

Teo, Y.-Y., Small, K.S. & Kwiatkowski, D.P. 2010. Methodological challenges of genome-wide association analysis in Africa. *Nature reviews genetics*, 11(2):149-160.

The emerging risk factors collaboration. 2012. C-reactive protein, fibrinogen, and cardiovascular disease prediction. *New England journal of medicine*, 2012(367):1310-1320.

Theodoraki, E.V., Nikopensius, T., Suhorutšenko, J., Peppes, V., Fili, P., Kolovou, G., Papamikos, V., Richter, D., Zakopoulos, N. & Krjutškov, K. 2010. Fibrinogen beta variants confer protection against coronary artery disease in a Greek case-control study. *Biomed central medical genetics*, 11(1):1-7.

Thomas, A., Green, F. & Humphries, S. 1996. Association of genetic variation at the β-fibrinogen gene locus and plasma fibrinogen evels; interaction between allele frequency of the G/A-455polymorphism, age and smoking. *Clinical genetics*, 50(4):184-190.

Tishkoff, S.A. & Williams, S.M. 2002. Genetic analysis of African populations: human evolution and complex disease. *Nature reviews genetics*, 3(8):611-621.

Tousoulis, D., Papageorgiou, N., Androulakis, E., Briasoulis, A., Antoniades, C. & Stefanadis, C. 2011. Fibrinogen and cardiovascular disease: genetics and biomarkers. *Blood reviews*, 25(6):239-245.

Tsakadze, N.L., Zhao, Z. & D'Souza, S.E. 2002. Interactions of intercellular adhesion molecule-1 with fibrinogen. *Trends in cardiovascular medicine*, 12(3):101-108.

Tsurupa, G., Hantgan, R.R., Burton, R.A., Pechik, I., Tjandra, N. & Medved, L. 2009. Structure, stability, and interaction of the fibrin (ogen) αC-domains. *Biochemistry*, 48(51):12191-12201.

Tuğrul, M., Paixão, T., Barton, N.H. & Tkačik, G. 2015. Dynamics of transcription factor binding site evolution. *PLOS genetics*, 11(11):e1005639.

Tybjærg-Hansen, A., Agerholm-Larsen, B., Humphries, S.E., Abildgaard, S., Schnohr, P. & Nordestgaard, B.G. 1997. A common mutation (G-455--> A) in the beta-fibrinogen promoter is an independent predictor of plasma fibrinogen, but not of ischemic heart disease. A study of 9,127 individuals based on the Copenhagen city heart study. *Journal of clinical investigation*, 99(12):3034-3039.

Uitte de Willige, S., de Visser, M.C., Houwing-Duistermaat, J.J., Rosendaal, F.R., Vos, H.L. & Bertina, R.M. 2005. Genetic variation in the fibrinogen gamma gene increases the risk for deep venous thrombosis by reducing plasma fibrinogen γ′ levels. *Blood*, 106(13):4176-4183.

Uitte de Willige, S., Standeven, K.F., Philippou, H. & Ariëns, R.A. 2009. The pleiotropic role of the fibrinogen γ′ chain in hemostasis. *Blood*, 114(19):3994-4001.

Undas, A. & Ariëns, R.A. 2011. Fibrin clot structure and function: a role in the pathophysiology of arterial and venous thromboembolic diseases. *Arteriosclerosis, thrombosis, and vascular biology*, 31(12):e88-99.

Van de Stolpe, A., Jacobs, N., Hage, W., Tertoolen, L., Van Kooyk, Y., Novakova, I. & De Witte, T. 1996. Fibrinogen binding to ICAM-1 on EA.hy 926 endothelial cells is dependent on an intact cytoskeleton. *Thrombosis and haemostasis*, 75(1):182-189.

Van Der Bom, J.G., De Maat, M.P., Bots, M.L., Haverkate, F., De Jong, P., Hofman, A., Kluft, C. & Grobbee, D.E. 1998. Elevated plasma fibrinogen cause or consequence of cardiovascular disease? *Arteriosclerosis, thrombosis, and vascular biology*, 18(4):621-625.

van't Hooft, F.M., von Bahr, S.J., Silveira, A., Iliadou, A., Eriksson, P. & Hamsten, A. 1999. Two common, functional polymorphisms in the promoter region of the β-fibrinogen gene contribute to regulation of plasma fibrinogen concentration. *Arteriosclerosis, thrombosis, and vascular biology*, 19(12):3063-3070.

Vasse, M., Paysant, J., Soria, J., Collet, J., Vannier, J. & Soria, C. 1996. Regulation of fibrinogen biosynthesis by cytokines, consequences on the vascular risk. *Pathophysiology of haemostasis and thrombosis*, 26(S4):331-339.

Vermylen, C., De Vreker, R. & Verstraete, M. 1963. A rapid enzymatic method for assay of fibrinogen fibrin polymerization time (FPT test). *Clinica chimica acta*, 8(3):418-424.

Verschuur, M., de Jong, M., Felida, L., de Maat, M.P. & Vos, H.L. 2005. A hepatocyte nuclear factor-3 site in the fibrinogen β promoter is important for interleukin 6-induced expression, and its activity is influenced by the adjacent–148C/T polymorphism. *Journal of biological chemistry*, 280(17):16763-16771.

Wall, J.D. 2001. Insights from linked single nucleotide polymorphisms: what we can learn from linkage disequilibrium. *Current opinion in genetics & development*, 11(6):647-651.

Wall, J.D. & Pritchard, J.K. 2003. Haplotype blocks and linkage disequilibrium in the human genome. *Nature reviews genetics*, 4(8):587-597.

Wang, N., Akey, J.M., Zhang, K., Chakraborty, R. & Jin, L. 2002. Distribution of recombination crossovers and the origin of haplotype blocks: the interplay of population history, recombination, and mutation. *American journal of human genetics*, 71(5):1227-1234.

Wang, X., Wang, J., McCredie, R. & Wilcken, D. 1997. Polymorphisms of Factor V, Factor VII, and Fibrinogen Genes Relevance to Severity of Coronary Artery Disease. *Arteriosclerosis, thrombosis, and vascular biology*, 17(2):246-251.

Wannamethee, S.G., Lowe, G.D., Shaper, A.G., Rumley, A., Lennon, L. & Whincup, P.H. 2005. Associations between cigarette smoking, pipe/cigar smoking, and smoking cessation, and haemostatic and inflammatory markers for cardiovascular disease. *European heart journal*, 26(17):1765-1773.

Wassel, C.L., Lange, L.A., Keating, B.J., Taylor, K.C., Johnson, A.D., Palmer, C., Ho, L.A., Smith, N.L., Lange, E.M. & Li, Y. 2011. Association of genomic loci from a cardiovascular gene SNP array with fibrinogen levels in European Americans and African-Americans from six cohort studies: the candidate gene association resource (CARe). *Blood*, 117(1):268-275.

Weijers, E.M. 2011. Fibrin matrices for tissue engineering. Amsterdam: Vrije Universiteit. (Thesis – PhD).

Weinstock, N. & Ntefidou, M. 2006. SSC international collaborative study to establish the first high fibrinogen plasma reference material for use with different fibrinogen assay techniques. *Journal of thrombosis and haemostasis*, 4(8):1825-1827.

Weisel, J.W. & Nagaswami, C. 1992. Computer modeling of fibrin polymerization kinetics correlated with electron microscope and turbidity observations: clot structure and assembly are kinetically controlled. *Biophysical journal*, 63(1):111-128.

Widmaier, E.P., Raff, H., Strang, K.T. & Vander, A.J. 2011. Vander's human physiology : the mechanisms of body function. 12th ed. New York: McGraw-Hill.

Wilhelmsen, L., Svärdsudd, K., Korsan-Bengtsen, K., Larsson, B., Welin, L. & Tibblin, G. 1984. Fibrinogen as a risk factor for stroke and myocardial infarction. *New England journal of medicine*, 311(8):501-505.

Wilner, G., Nossel, H. & LeRoy, E. 1968. Activation of hageman factor by collagen. *Journal of clinical investigation*, 47(12):2608-2615.

Wolfenstein-Todel, C. & Mosesson, M.W. 1980. Human plasma fibrinogen heterogeneity: evidence for an extended carboxyl-terminal sequence in a normal gamma chain variant (gamma'). *Proceedings of the national academy of sciences of the United States of America*, 77(9):5069-5073.

Yang, Q., Tofler, G.H., Cupples, L.A., Larson, M.G., Feng, D., Lindpaintner, K., Levy, D., D'Agostino, R.B. & O'Donnell, C.J. 2003. A genome-wide search for genes affecting circulating fibrinogen levels in the Framingham Heart Study. *Thrombosis research*, 110(1):57-64.

Yang, Z., Mochalkin, I. & Doolittle, R.F. 2000. A model of fibrin formation based on crystal structures of fibrinogen and fibrin fragments complexed with synthetic peptides. *Proceedings of the national academy of sciences of the United States of America*, 97(26):14156-14161.

Yates, A., Akanni, W., Amode, M.R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S. & Gil, L. 2016. Ensembl 2016. *Nucleic acids research*, 44(D1):710-716.

Yu, S., Sher, B., Kudryk, B. & Redman, C. 1983. Intracellular assembly of human fibrinogen. *Journal of biological chemistry*, 258(22):13407-13410.

Yu, S., Sher, B., Kudryk, B. & Redman, C.M. 1984. Fibrinogen precursors. Order of assembly of fibrinogen chains. *Journal of biological chemistry*, 259(16):10574-10581.

Zhang, J.-Z. & Redman, C. 1992. Identification of Bβ chain domains involved in human fibrinogen assembly. *Journal of biological chemistry*, 267(30):21727-21732.

Zhang, K., Calabrese, P., Nordborg, M. & Sun, F. 2002. Haplotype block structure and its applications to association studies: power and study designs. *American journal of human genetics*, 71(6):1386-1394.

ANNEXURE A

This annexure includes:

- Published article: “Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation” (Chapter 4).

Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation

H. Toinét Cronjé†, Cornelie Nienaber-Rousseau†, Lizelle Zandberg, Tinashe Chikowore, Zelda de Lange, Tertia van Zyl and Marlien Pieters

Centre of Excellence for Nutrition, *North-West University, Potchefstroom, South Africa*

Correspondence to Cornelie Nienaber-Rousseau: *Private bag x6001, Nutrition, Box 594, Potchefstroom 2520, South Africa.*
cornelie.nienaber@nwu.ac.za.
http://dx.doi.org/10.1016/j.matbio.2016.10.005

Abstract

Fibrinogen and its functional aspects have been linked to cardiovascular disease. There is vast discrepancy between the heritability of fibrinogen concentrations observed in twin studies and the heritability uncovered by genome wide association studies. We postulate that some of the missing heritability might be explained by the pleiotropic and polygenic co-regulation of fibrinogen through multiple targeted genes, apart from the fibrinogen genes themselves. To this end we investigated single nucleotide polymorphisms (SNPs) in genes coding for phenotypes associated with total and γ′ fibrinogen concentrations and clot properties. Their individual and accumulative associations with the fibrinogen variables were explored together with possible co-regulatory processes as a result of the gain and loss of transcription factor binding sites (TFBS). Seventy-eight SNPs spanning the *APOB*, *APOE*, *CBS*, *CRP*, *F13A1*, *FGA*, *FGB*, *FGG*, *LDL-R*, *MTHFR*, *MTR, PCSK-9* and *SERPINE-1* genes were included in the final analysis. A novel *PCSK-9* SNP (rs369066144) was identified in this population, which associated significantly ($p = 0.04$) with clot lysis time (CLT). Apart from SNPs in the fibrinogen (*FGA*, *FGB* and *FGG*) and FXIII (*F13A1*) genes, the fibrinogen phenotypes were also associated with SNPs in genes playing a role in lipid homeostasis (*LDL-R, PCSK-9*) together with *CBS* and *CRP* polymorphisms (particularly, *CRP*-rs3093068). The genetic risk scores, presenting accumulative genetic risk, were significantly associated ($p \leq 0.007$) with total and γ′ fibrinogen concentrations, lag time, slope and CLT, highlighting the importance of a polygenetic approach in determining complex phenotypes. SNPs significantly associated with the fibrinogen phenotypes, resulted in a total of 75 TFBS changes, of which 35 resulted in a loss and 40 in a gain of TFBS. In terms of co-regulation, V$IRF4.02, V$E2FF and V$HIFF were of particular importance. The investigation into TFBS provided valuable insight as to how sequence divergences in seemingly unrelated genes can result in transcriptional co-regulation of the fibrinogen phenotypes. The observed associations between the identified SNPs and the fibrinogen phenotypes therefore do not imply direct effects on cardiovascular disease outcomes, but may prove useful in explaining more of the genetic regulation of the investigated fibrinogen phenotypes.

Introduction

Fibrinogen is a hexameric glycoprotein consisting of two alpha (α), two beta (β) and two gamma (γ) polypeptide chains, and is encoded for by the fibrinogen α, β and γ chain genes (*FGA*, *FGB* and *FGG*) on the q-arm of chromosome four [1]. Several variants of the fibrinogen molecule exist, of which fibrinogen gamma prime (γ′) is considered to be a common splice variant contributing to between 8 and 11% of total plasma fibrinogen concentrations [2,3]. Fibrinogen γ′ arises in response to alternative splicing of the carboxyl-terminal region of γ-mRNA, resulting in a higher molecular weight γ′ than the γA chain [2,3]. Fibrinogen is an essential haemostatic protein, as the precursor of fibrin in the final stages of

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

Table 1. Descriptive characteristics of the study participants.

Variable	Median (25–75th percentiles)
Age (years)	48 (41–56)
Gender [n (%)]	
Male	628 (37.4)
Female	1049 (62.6)
HIV status [n (%)]	
Positive	269 (16.1)
Negative	1401 (83.9)
Total fibrinogen (g/L)	2.9 (2.2–5.0)
Fibrinogen γ′ (%)	10.2 (1.2–14.7)
Lag time (min)	6.5 (5.1–7.8)
Slope ($\times 10^{-3}$ au/s)	9.0 (6.5–12.0)
Maximum absorbance (nm)	0.4 (0.3–0.5)
CLT (min)	57.1 (50.9–63.9)

CLT = clot lysis time; HIV = human immunodeficiency virus

blood coagulation [4]. Both total and γ′ fibrinogen influence clot structure [5,6], and increased levels have been related to denser blood clots resistant to clot lysis [7]. Denser clots that remain in the vasculature for a longer time are associated with cardiovascular disease (CVD) outcomes such as myocardial infarction, stroke and coronary artery disease [8]. Increased fibrinogen concentration, irrespective of its functional properties, is also associated with CVD [9].

Family and twin studies investigating the genetics of fibrinogen have reported the heritability of fibrinogen concentrations to be between 30 and 51% [10–16]. Genome wide association studies (GWAS) have only allocated 3.7% of this possible 51% fibrinogen heritability to common single nucleotide polymorphisms (SNPs; single base pair variations occurring at frequencies of >1% in the population) to date [17]. As these GWAS included millions of SNPs without significant results, investigating pleiotropic and polygenic sequence divergences and the possible co-regulation thereof among candidate genes chosen based on association analyses of phenotypes related to fibrinogen, might prove valuable. There is growing evidence that a significant proportion of the heritability of complex phenotypes, such as fibrinogen, may be explained by a combination of genetic variants, and their combined effects can be calculated in the form of polygenic risk scores [18], making use of a biological filtering approach by taking into account the mechanistic pathways associated with the complex phenotypes. Furthermore, the combination of genetics and molecular biology has greatly facilitated the identification of candidate genes [19,20]. Multifactorial phenotypes can now be represented as complex interactive networks, which consist of a combination of genetic and non-genetic factors. Therefore, genetic variations in multiple genes in one particular pathway or disease network could lead to synergistic heterozygosity [21]. We have incorporated the concept of synergistic heterozygosity into our hypothesis, through genetic risk score (GRS) analyses, to observe the polygenic effects of harbouring several risk alleles concurrently. In addition, gene expression is generally controlled by transcriptional enhancers, which consist of a cluster of transcription factor binding sites (TFBS) spaced by spacer sequences and enhancers. SNPs in these regions have potential functional significance, which is not necessarily obvious when observed independently from a functional context [22–24]. Being in regulatory regions, these SNPs may, therefore, associate with the outcome phenotype through polygenic co-regulation rather than being in the causal pathway.

In agreement with international research [25,26], ethnic differences in fibrinogen concentrations have been reported in the South African population. Black South Africans present with higher total fibrinogen concentrations than their counterparts from European ancestry [27,28]. A pre-disposition to hypercoagulability has also been observed in the black South African population [29]. Studies on the genetic variation in African individuals have revealed vast genetic diversity in Africa [30,31]. Furthermore, research conducted in the South African Prospective Urban and Rural Epidemiology (PURE) study population revealed less linkage disequilibrium (LD; neighbouring polymorphisms that are inherited together) in the fibrinogen genes than what has been observed in European populations [32]. The higher variance in fibrinogen concentrations, together with a unique genetic diversity and low LD, presents a promising opportunity to unravel the missing heritability observed for fibrinogen.

To this end, we conducted a candidate gene association study by investigating SNPs of variables associated with fibrinogen-related phenotypes. Candidate genes to be included were further verified through an *in silico* pathway and network analysis to increase our understanding of the polygenic regulation of the fibrinogen phenotype. Polygenetic and pleiotropic co-regulation were explored *via* GRS and TFBS analyses. Protein concentration (both total and γ′ fibrinogen) and functionality (plasma clot properties) were used as outcome variables.

Results and discussion

The total study population consisted of 1677 participants included after quality control (QC). Table 1 presents the descriptive characteristics of the study population. Total and γ′ fibrinogen concentrations, together with lag time, slope, maximum absorbance and clot lysis time (CLT), were used as phenotype outcomes for all further analyses.

This study is the first exploration of the pleiotropic and polygenetic co-regulation of candidate SNPs associated with the fibrinogen protein and its functional phenotypes in a large black South African

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

Table 2. Phenotypes correlating significantly with total and γ′ fibrinogen concentration and clot properties.

Total fibrinogen* (g/L)		Fibrinogen γ′* (%)		Clot lysis time* (min)	
Variable	r	Variable	r	Variable	r
Age (years)	0.15	Age (years)	−0.11	BMI (kg/m^2)[a]	0.28
BMI (kg/m^2)[a]	0.16	Hcy (μmol/L)	−0.11	Hip C (cm)[a]	0.27
Hip C (cm)[a]	0.17	HDL-c (mmol/L)	−0.12	HbA1c (%)	0.22
HbA1c (%)	0.20	PAI-1$_{act}$ (U/mL)	−0.17	TG (mmol/L)[b]	0.20
LDL-c (mmol/L)	0.13			LDL-c (mmol/L)[b]	0.17
CRP (mg/L)	0.42			HDL-c (mmol/L)[b]	0.20
IL-6 (pg/mL)	0.26			PAI-1$_{act}$ (U/mL)	0.31

Lag time* (min)		Slope* (×10^{-3} au/s)		Maximum absorbance* (nm)	
Variable	r	Variable	r	Variable	r
Age (years)	0.14	PA Index	0.18	Age (years)	0.14
PA index	−0.18	Hcy (μmol/L)	0.10	BMI (kg/m^2)[a]	0.10
SBP (mm Hg)	0.13	PAI-1$_{act}$ (U/mL)	−0.17	Hip C (cm)[a]	0.10
		CRP (mg/L)	0.23	CRP (mg/L)	0.26

BMI = body mass index; CRP = C-reactive protein; HbA1c = glycated haemoglobin; Hcy = homocysteine; HDL-c = high-density lipoprotein cholesterol; Hip C = hip circumference; IL-6 = interleukin-6; LDL-c = low-density lipoprotein cholesterol; PAI-1$_{act}$ = plasminogen activator inhibitor type 1 activity; PA index- = physical activity index; SPB = systolic blood pressure; TG = triglycerides

* $p \leq 0.001$ for all variables.

[a] Adjustments were made only for the highest correlating variable when both BMI and Hip C were correlated to the outcome variable owing to co-linearity

[b] Only TG and HDL-c adjusted for owing to co-linearity.

population. Association analyses were conducted to identify appropriate SNPs to include. Table 2 presents the phenotypes that correlated significantly ($r > 0.1$; $p \leq 0.001$) with the fibrinogen variables in this study population. Genes were identified for inclusion in genetic association analyses based on associations of their respective phenotypes with the outcome variables. In addition, the continuous variables presented in Table 2 were treated as covariates and adjusted for in all further analyses. Categorical variables, gender and human immunodeficiency virus (HIV) status were identified as covariates when investigating total and γ′ fibrinogen concentrations; however, for CLT we co-varied for gender alone, and for slope and maximum absorbance we only adjusted for HIV status, based on the significance of these associations.

The 13 genes of interest, coding for apolipoprotein-B (Apo-B), apolipoprotein-E (Apo-E), cystathionine beta synthase (CBS), C-reactive protein (CRP), factor XIII (FXIII), fibrinogen, low-density lipoprotein receptor (LDL-R), methylenetetrahydrofolate reductase (MTHFR), methionine synthase (MTR), proprotein convertase subtilisin/kexin type 9 (PCSK 9)

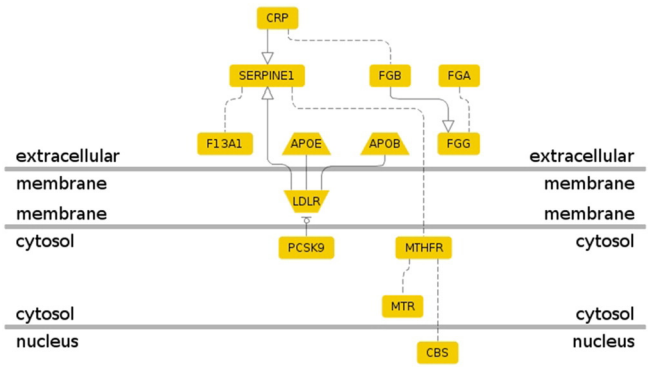


Fig. 1. The *in silico* portrayal of the 13 genes included in a CVD network. ------ Indirect relationship —— Direct relationship ⊤ Suppressive relationship ⟶ Directional relationship.

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

Table 3. SNPs significantly associated with the fibrinogen phenotypes.

Chr	Gene	SNP	SNP pseudonyms[a]	Location	MAF (%)	β	p
Total fibrinogen (g/L)							
1	*PCSK-9*	rs28362202	c.-26 G>A	Promoter exon	7.4	0.41	0.004
1	*PCSK-9*	rs2483205	c.658-7 C>T	Intron	47.1	0.17	0.021
1	*PCSK-9*	rs2495477	c.799+3 G>A	Intron	27.1	−0.19	0.022
1	*PCSK-9*	rs374941781	c.829 G>C	Intron	6.5	0.43	0.003
1	*PCSK-9*	rs584626[b]	c.1354+102 T>C	Intron	33.2	0.19	0.013
4	*FGA*	rs2070011	c.-58 G>A	Promoter exon	17.0	0.22	0.032
4	*FGG*	rs1049636	c.1299+79 T>C	Exon intron	15.6	0.32	0.002
Fibrinogen γ′ (%)							
1	*CRP*	rs2808630	g.159711078 A>G	Intergenic	14.2	1.00	0.024
1	*CRP*	rs3093068[c]	g.159711574 C>G	Intergenic	37.5	−1.28	0.001
1	*CRP*	rs3093062[b]	c.-409 G>A	Promoter intergenic	16.2	−1.11	0.011
1	*CRP*	rs7553007[b]	g.159728759 G>A	Intergenic	23.7	1.19	0.001
Lag time (min)							
1	*PCSK-9*	rs499718	c.523+230 C>T	Intron	48.5	−1.60	0.025
1	*PCSK-9*	rs505151	c.2009 G>A	Exon	30.1	0.18	0.024
6	*F13A1*	rs5985	c.344 G>A	Exon	14.2	0.29	0.007
Slope ($\times 10^{-3}$ au/s)							
1	*PCSK-9*	rs2495482	c.207+15 G>A	Intron	13.2	−0.59	0.012
1	*PCSK-9*	rs4927193	c.399+165 T>C	Intron	32.4	−0.37	0.020
1	*PCSK-9*	rs28362263	c.1327 A>G	Exon intron	3.8	−0.86	0.034
1	*PCSK-9*	rs584626[b]	c.1354+102 T>C	Intron	33.2	0.33	0.042
4	*FGA*	rs6050	c.991 A>G	Exon intron	30.2	0.41	0.016
19	*LDL-R*	rs2228671	c.81 C>T	Exon	3.0	1.15	0.010
19	*LDL-R*	rs3826810	c.*141 G>A	Exon	12.9	0.52	0.030
19	*LDL-R*	rs2738465	c.*504 G>A	3′ UTR	25.2	0.37	0.033
19	*LDL-R*	rs1433099	c.*666 T>C	3′ UTR	47.2	0.31	0.049
Maximum absorbance (nm)							
4	*FGB*	rs4463047	g.154574381 T>C	Intergenic	10.5	0.02	0.027
4	*FGA*	rs2070011	c.-58 G>A	Promoter exon	17.0	0.02	0.014
19	*LDL-R*	rs17242759	c.67+18 C>A	Intron	8.9	0.02	0.045
Clot lysis time (min)							
1	*PCSK-9*	rs494198	c.799+64 A>C	Intron	41.9	−0.76	0.043
1	*PCSK-9*	rs369066144[c]	c.*500 C>T	3′ UTR	3.5	−3.61	0.001
21	*CBS*	rs5742905[b]	c.833 T>C	Exon	27.2	−0.97	0.022

CBS = cystathionine beta synthase; Chr = chromosome; CRP = C-reactive protein; FGA = fibrinogen alpha chain gene; FGB = fibrinogen beta chain gene; FGG = fibrinogen gamma chain gene; F13 A1 = factor XIII; LDL-R = low-density lipoprotein receptor; MAF = minor allele frequency; PCSK-9 = proprotein convertase subtilisin/kexin type 9; rs = reference sequence; SNP = single nucleotide polymorphism; UTR = untranslated region.

[a] Ensembl release 84 - March 2016, retrieved 27 May 2016 [33].

[b] rs3093062 is in linkage with rs3093058; rs7553007 is in LD with rs1341665, rs2027471, rs1205 and rs2794520; rs584626 is in LD with rs533375 and rs585131; rs5742905 is in LD with *CBS* 844ins68.

[c] Remained significant after adjustment for multiple testing.

and plasminogen activator inhibitor type 1 (PAI-1), identified *via* association analyses (Table 2), were uploaded into the Genomatix Pathways System application of the Genomatix Software Suite package. The connectivity between these genes was explored to verify whether they would meaningfully contribute to further statistical models. All the genes included were significantly (13/892, $p = 9.38 \times 10^{-5}$) related to the CVD pathway. The direct and indirect relationships indicative of the shortest route between the genes are depicted in Fig. 1. The interconnectivity between these genes demonstrates the rationale for their inclusion to explore the polygenic co-regulation as contributing factors to the fibrinogen phenotypes observed in this study population.

Upon evaluation of the results presented in Fig. 1 and Table 2, SNP selection commenced (as detailed in the Methods section). One hundred and eight polymorphisms analysed in previous PURE sub-studies were included prior to QC. Four polymorphisms were removed on account of having a minor allele frequency (MAF) < 1%, three SNPs did not adhere to the Hardy–Weinberg equilibrium (HWE), and 15 SNPs were pruned because of linkage with other SNPs in the model ($r^2 > 0.5$; $p < 0.05$). A further eight polymorphisms were removed as a result of insufficient call rates. Final analyses consisted of 78 SNPs with a genotype call rate of 97%, spanning the *APOB*, *APOE*, *CBS*, *CRP*, *F13A1*, *FGA*, *FGB*, *FGG*, *LDL-R*, *MTHFR*, *MTR*, *PCSK-9* and *SERPINE-1*

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

Table 4. Contribution of genetic risk scores to phenotypes.

Phenotype	GRS β (95% CI)	GRS p-value	GRS total model r^2
Total fibrinogen (g/L)	0.54 (0.34–0.75)	<0.001	0.29
Fibrinogen γ′ (%)	0.55 (0.32–0.79)	<0.001	0.04
Lag time (min)	0.12 (0.03–0.21)	0.007	0.05
Slope ($\times 10^{-3}$ au/s)	0.79 (0.54–1.05)	<0.001	0.15
Maximum absorbance (nm)	0.71 (−0.67–2.10)	0.31	0.01
CLT (min)	1.07 (0.64–1.50)	<0.001	0.18

CI = confidence interval; CLT = clot lysis time; GRS = genetic risk score.

genes. A complete list of polymorphisms included in this study is presented in a supplementary table (Supplementary Table 1).

Individual association of SNPs with fibrinogen variables

Regression analyses were performed for individual phenotype outcomes, with all 78 SNPs and covariates as determined for each phenotype (Table 2) included. The 27 SNPs that presented with significant results, after adjusting for covariates, prior to adjustment for multiple testing, are reported in Table 3. Apart from SNPs directly related to the outcome variables (*i.e. FGA*, *FGB*, *FGG* and *F13A1*), loci in the *LDL-R*, *PCSK-9*, *CRP* and *CBS* genes also contributed significantly to the fibrinogen phenotypes. After adjusting for multiple testing *PCSK-9*-rs369066144 (which was not in LD with any of the other *PCSK-9* SNPs) and *CRP*-rs3093068 (also no LD with other *CRP* SNPs) remained significantly associated with CLT ($p = 0.04$) and γ′ fibrinogen ($p = 0.02$), respectively.

PCSK-9-rs369066144 (c.500 C>T) was a novel variant identified in the PURE population with a MAF of 3.5%. It is situated in the 3′ untranslated region (3′ UTR), generally known to be involved in the regulation of gene expression. The functional relevance of this particular variant is still to be determined. Twelve of the 24 *PCSK-9* polymorphisms, together with five of the 19 *LDL-R* polymorphisms under investigation, presented with significant associations with one or more of the outcome phenotypes. Both of these genes encode for proteins involved in blood lipid homeostasis and have been associated with low-density lipoprotein cholesterol (LDL-c) concentrations in the PURE population in which statin use was negligible [34]. *PCSK-9* is the gene coding for the PCSK-9 enzyme that is involved in the degradation of the LDL-R, thereby inhibiting the ability of the LDL-R to remove LDL-c from the bloodstream [35,36]. The number of significant associations seen between these genes and all fibrinogen phenotypes (apart from γ′ fibrinogen) indicates that these lipid mediating genes might provide some insight into the missing heritability observed for fibrinogen and the genetic regulation of clot properties. Literature on the LDL-R and PCSK-9 and the current phenotypes is limited, with only a single study revealing PCSK-9 protein concentrations positively correlated to fibrinogen concentrations [37], and some evidence of an association between LDL-c concentrations and the enhanced onset of clot formation *via* enhanced platelet activation and tissue factor expression [38]. One *LDL-R* polymorphism, rs2228671, has specifically been observed to be an independent predictor of Factor VIII (FVIII), a component of the coagulation cascade, with the T-allele being associated with increased FVIII concentrations [39]. Enhanced functioning of the coagulation cascade could explain the accelerated clot formation indicated by an increase in slope. The current investigation is, to our knowledge, the first exploration of the association of the genetic determinants of PCSK-9 and LDL-R with fibrinogen and clot properties. Future studies, investigating PCSK-9 and LDL-R concentrations and how they are associated with fibrinogen and its functional aspects, will shed more light on our observations.

As mentioned above, *CRP*-rs3093068 also remained significantly associated with γ′ following multiple testing ($p = 0.02$). The fibrinogen γ′ phenotype was associated exclusively with four of the included six SNPs of the *CRP* gene. These associations may be explained by co-regulatory processes, rather than being in the fibrinogen causal pathway, as detailed in Section 2.3.

Polygenic association of SNPs with fibrinogen variables

The idea of combining multiple genetic markers into a single score for predicting disease risk has recently become a popular avenue to explore [40]. Generally, individual markers show no detected effect on their own; however, when combined in a score of several genetic factors it has a stronger predictive power for a particular phenotype. With this in mind we made use of a GRS model as a way of measuring the accumulative effect these SNPs have on each phenotype outcome. The model uses a basic weighted polygenic risk score approach where a combination of genetic variants, in this case SNPs that associated significantly with the respective fibrinogen phenotypes (Table 3) prior to adjustment for multiple testing, are evaluated as a combined score allocated to each outcome. The GRS model takes the direction of the β-values obtained above into account, thereby allocating higher scores to the risk allele, not necessarily the minor allele (see Section 3.5.1). As the African population has been observed to have a predisposition to hypercoagulability, any allele adding to the risk of hypercoagulability was identified as a risk

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

allele in this population. Being a weighted score, it also considers the magnitude of the contribution of the allele to the phenotype.

Adjustments were made for the specific covariates identified previously (Table 2) for each outcome variable. Table 4 presents the individual contribution of each GRS to the variance of the respective fibrinogen phenotypes, as well as the contribution of the total model. Each GRS was positively associated with its respective fibrinogen phenotype ($p \leq 0.007$) except for maximum absorbance ($p = 0.31$), highlighting the value of the polygenic approach. The use of a GRS is more valuable than investigating SNPs individually, as it considers the accumulative genetic variation present in an individual and calculates the combined effect of a specific set of sequence variations across the genome on the outcome phenotype. Though significant, the fibrinogen phenotypes' variance explained by the GRS, ranged from 4 to 29% only, placing the SNPs' contributions into perspective in terms of other contributing factors such as environmental factors or SNPs not genotyped in this study.

Polygenic co-regulation

In an effort to examine the mechanism behind the cumulative effect observed in the GRS results, we investigated possible co-regulation of the genes included in this study through the gain and loss of TFBS resulting from the SNPs identified. Considering the majority of these variations were not necessarily coding for molecules in the fibrinogen causal pathway, and were found in the non–protein coding regions, the functionality of the particular variation is not as obvious as in the case of the coding region variants. The possible effects of the variation of the non-coding region become evident only when viewed in a particular functional context, such as when investigating TFBS, which the included genes have in common. The gain and loss of TFBS due to SNPs create the possibility for seemingly unrelated genes to be co-regulated *via* the recruitment of a common transcription factor (TF). The importance of the gain and loss of TFBS lies within the resultant secondary signalling response, which follows as a consequence of transcriptional demand. The 27 SNPs in Table 3 gave rise to a total of 75 TFBS changes, of which 35 sequence divergences resulted in a loss of TFBS and 40 formed a new TFBS, hence an overall gain (Supplementary Table 2).

One noticeable feature in terms of fibrinogen concentrations was the loss and gain of the V$IRF4.02 matrix that binds the TF interferon regulatory factor 4 (IRF4). The loss of this TFBS occurs at the *PCSK-9* loci as a result of *PCSK-9*-rs37494178 and the gain of the same V$IRF4.02 matrix at the *FGG* loci was created as a result of *FGG*-rs1049636. The gain and loss of the same TFBS might be indicative of a possible TF interaction between *PCSK-9* and *FGG.* However, the exact mechanism between the TFBS gain and loss and TF recruitment still needs to be elucidated. It is known that the interferon regulatory factor family of proteins has a particularly important function in the regulatory effect of interferons on interferon-inducible genes. IRF4 is a family member specific for lymphocytes and negatively regulates toll-like-receptor (TLR) signalling that is central to the activation of innate and adaptive immune systems [41,42]. Recently, several TFs, of which IRF4 is one, have been observed to contribute to cardiac hypertrophy [42]. The role of IRF4 in the context of fibrinogen concentrations and its regulation needs further investigation. Similarly, the V$E2FF family matrix was lost at the *PCSK-9* loci as a result of rs28362202 and gained for *FGA*-rs2070011. The significance of the gain and loss of E2FF in the context of fibrinogen remains unclear. However, E2F TFs such as E2F1 regulate specific biological pathways such as cell responses to inflammation, hypoxia, DNA damage and proliferation [43] and, therefore, might be important in the fibrinogen phenotype context.

In terms of the *CRP* SNPs associated with γ′ fibrinogen, the gain of TFBS, V$HIFF, a binding site for hypoxia inducible factor 1 alpha (*HIF1A*) was noted. The induction of *HIF1A* expression is thought to regulate the expression of hepatocyte nuclear factor 4 alpha (*HNF4A*), which activates *HNF1* expression and is followed by an increased expression of γ chain fibrinogen. Therefore, the gain of the V$HIFF site on the *CRP* gene allows the binding of a regulatory component (HIFIA), suggesting a possible co-regulatory pathway for CRP and γ′ fibrinogen. HIF1 also plays an essential role in vasculogenesis [44] and has an effect on other acute phase proteins such as interleukin-6 (IL-6). The relationship between *CRP* SNPs and γ′ fibrinogen might be modulated through IL-6 and not through CRP directly, since IL-6 stimulates the hepatic production of both fibrinogen and CRP concentrations [45,46]. As for the phenotypes related to clot properties, no clear mechanism in terms of TFBS was observed. Supplementary Table 2 reports all TFBS gains and losses present for the respective loci.

Although our association results uncovered SNPs that significantly associated with fibrinogen phenotypes, it remains possible that our findings were attributable to other SNPs that were in strong LD with these polymorphisms, which we did not genotype. In addition, the development of a GWAS chip for black South African populations will contribute significantly to the identification of SNPs related to the fibrinogen phenotype, beyond what we were able to determine in this study, including a more extensive array of SNPs involved in the inflammatory pathway being critical in fibrinogen regulation. The concept of the polygenic regulation of fibrinogen phenotypes proved valuable in our study, since it gave a more integrative approach to groupings of SNPs that contributed alone and as a

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

collective to our fibrinogen phenotypes. Future studies gathering transcriptomic data will provide stronger evidence of the possible effects of the gain and/or loss of TFBS explored in this study.

In conclusion, our results revealed not only the regulation of the fibrinogen phenotypes through fibrinogen (*FGA*, *FGB* and *FGG*) and FXIII (*F13A1*) SNPs, but also through SNPs in the *LDL-R*, *PCSK-9*, *CBS* and *CRP* genes. The last-mentioned genes have not previously been related to the fibrinogen phenotypes. The accumulating associations of these SNPs as determined through the GRSs proved the importance of a polygenetic approach in determining complex phenotypes and might be of practical importance to an individual's risk in terms of hypercoagulability. Further investigation of TFBS indicated that polygenic transcriptional co-regulation may form the basis of the SNPs' association with the fibrinogen phenotypes. The associations between the SNPs identified here and the fibrinogen phenotypes do not imply any causal link to CVD outcomes, but adds to our current knowledge of the genetic regulation of fibrinogen and may prove useful in explaining more of the variance of the investigated fibrinogen phenotypes.

Methods

Study population

This is a cross-sectional study embedded in the international multi-centred PURE study [47]. Baseline measurements of the South African arm of the PURE study, consisting of 2010 self-identified Tswana individuals of the North West province, are reported. Apparently healthy men and women who were between the ages of 35 and 70 and not suffering from any acute or chronic illness were recruited for this study. Data collection was performed in accordance with the Declaration of Helsinki and ethical approval was granted by the ethics committee of the North-West University, South Africa (NWU-00016-10-A1).

Blood sampling and biochemical analyses

Blood samples were collected from the antecubital vein by a registered nurse between 7:00 am and 11:00 am after an overnight fast. Samples were centrifuged within 30 min of collection at 2000 × *g* for 15 min. A detailed description of the analyses of the biochemical markers associated with the fibrinogen phenotype, namely IL-6, homocysteine, high-sensitivity CRP, plasminogen activator inhibitor type 1 activity and serum lipids has been published previously [48–50]. The modified Clauss method on the Dade Behring BCS coagulation analyser was used to quantify total fibrinogen (Multifibrin U-test Dade Behring, Deerfield, IL, USA). Fibrinogen γ′ concentration was measured with an enzyme-linked immunosorbent assay making use of a 2.G2.H9 mouse monoclonal coating antibody against human γ′ fibrinogen (Santa Cruz Biotechnology, Santa Cruz, USA) and a goat polyclonal horseradish peroxidase-conjugated antibody against human fibrinogen (Abcam Cambridge, USA) [6].

Plasma fibrinolytic potential was determined by adding tissue plasminogen activator (tPA) to tissue-factor-induced plasma clots and measuring turbidity with a spectrophotometer [51]. Tissue factor and tPA concentrations were modified slightly for the purpose of obtaining comparable CLTs of about 60–100 min. Final concentrations in the clots were: tissue factor diluted 125 times (Dade Innovin, Siemens Healthcare Diagnostics Inc., Marburg, Germany), 17 mmol/L $CaCl_2$, 100 ng/mL tPA (Actilyse, Boehringer Ingelheim, Ingelheim, Germany) and 10 mmol/L phospholipid vesicles (Rossix, Mölndal, Sweden). CLT was calculated as the difference between the time at the midpoint of clear and maximum turbidity (clot formation) and the midpoint between maximum and clear turbidity (clot lysis). In addition to CLT, the following variables were calculated from the turbidity curves: lag time, an indicator of the time required for the activation of the coagulation cascade and for protofibrils to reach sufficient length to allow lateral aggregation; slope, as a representation of the rate of lateral aggregation of fibrin protofibrils; and maximum absorbance, as an indicator of fibre diameter.

Selection of genetic variants

Correlation analyses were used for the identification of phenotypes significantly associated ($r > 0.1$; $p < 0.05$) with one or more of our outcome variables, *i.e.* total and γ′ fibrinogen concentrations, lag time, slope, maximum absorbance and CLT. Subsequently polymorphisms previously genotyped in the PURE population, that were located in genes coding for Apo-B, Apo-E, CBS, CRP, FXIII, LDL-R, MTHFR, MTR, PCSK-9 and PAI-1 were included based on these phenotypic associations. Included polymorphisms were selected based on sequencing of the respective genes in 28 individuals in the PURE population. The sequencing data was used to determine genetic variation at the loci involved to establish the viability of genotyping candidate SNPs within this specific population, as well as to identify novel SNPs. In addition, SNPs known from the literature to be associated with the respective phenotypes were also included.

Genotyping

Isolation of genomic deoxyribonucleic acid (DNA) from the leukocyte layer was performed using QIAGEN® Flexigene® DNA extraction kits

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

(QIAGEN® Valencia, CA, USA; catalogue number 51,206), and concentrations were determined by the NanoDrop™ spectrophotometer (ND-1000, Wilmington, DE, USA).

Four major genotyping methods were used for the genotyping of the polymorphisms included. The BeadXpress® platform was used to genotype selected loci in the *APOB, APOE, CRP, F13A1, FGA*, *FGB*, *FGG*, *LDL-R* and *PCSK-9* [34,52,53] genes. PCR-based restriction fragment length polymorphism (RFLP) analysis was used for the genotyping of the *CBS, MTHFR* and *MTR* polymorphisms [54]. TaqMan-based assays (Thermo Fischer Scientific®, Walthem MA, USA), using the MyIQ5 Bio-Rad® real-time polymerase chain reaction (RT-PCR) thermal cycler (Bio-Rad® Laboratories Inc., Hercules, CA, USA) were used for the genotyping of selected loci within the *FGB*, *SERPINE-1* and *PCSK-9* genes [49,52]. Competitive allele-specific polymerase chain reaction (KASP) methods with supplies obtained from LGC Limited were used to genotype four additional *FGB* polymorphisms. Supplementary Table 1 presents the genotyping method used for each SNP included.

Statistical analysis

The statistical package for the social sciences (SPSS®) version 23 (IBM® Corp, 2015) and PLINK version 1.9 [55] were used for statistical analyses. Significance was set at a p-value of ≤ 0.05. Kolmogorov–Smirnov testing together with histograms and Q–Q plots were used to evaluate the normality of our data. Most of the variables were not normally distributed and, therefore, non-parametric analyses were performed. Correlations between continuous variables were determined by Spearman rank order correlations. Between-group comparisons for gender and HIV status were performed by conducting Mann–Whitney U tests. The results of these analyses were used to identify the phenotypes associated with the fibrinogen variables.

Quality control and statistical analysis of genetic data

QC included the removal of loci with a MAF below 1%, and failure to meet the requirements of the HWE. Correlated SNPs with LD indicated by r^2 of 0.5 or more were pruned and only one of these genetic variants were used for further statistical analysis. These pruned SNPs were, however, taken into consideration again when determining possible functional effects.

Association tests for the genetic variants that had passed QC were performed through PLINK software. One thousand permutation tests were used to adjust for multiple testing. The variants that presented with significant results prior to multiple testing were then grouped and used to compute a GRS [56]. A weighted method was used to calculate the GRS, which assumes the additive genetic model for each SNP. The additive genetic model performs well when the true genetic model is unknown or wrongly indicated [57]. The genotypes of the selected SNPs were assigned a weight of 0, 1 and 2 depending on the number of minor alleles. This score was then multiplied with the β-value obtained for each variant. The GRS was determined by adding the scores of the selected SNPs [57,58]. Linear regression models were used to evaluate the association of the GRS with total and γ′ fibrinogen concentrations, lag time, slope, maximum absorbance and CLT, respectively. Both the association and linear regression models for each phenotype outcome were performed without (results not shown) and after inclusion of covariates as determined by the Spearman correlations and Mann–Whitney U tests discussed previously.

In silico network analyses

The Genomatix Software Suite, version 3.6 (Genomatix Software GmbH, Munich, Germany) was applied for the *in silico* network analyses. Genomatix Pathways System was used to explore the network connectivity between the genes of interest and the literature mining extracted from the Genomatic Literature Mining database (Release 02-2016). SNPinspector imbedded in the Genomatix Software Suite package was used to inspect the gain and loss of TFBS and the TFBS information was extracted from the Matrix Library version 9.4 (December 2015) and ElDorado genome database (Release 02-2016). The *in silico* analyses were performed in the default settings as recommended for Homo sapiens data mining.

Conflict of interest

The authors declared no conflict of interest.

Author contribution

HTC: isolated the DNA (second round), performed KASP analyses, preparation of files for statistical analyses, statistical analyses using SPSS, wrote the manuscript with CN-R, reviewed and approved the final manuscript. CN-R: isolated the DNA (first round), performed RFLP analyses, conceptualised the paper, wrote the manuscript with HTC and was involved in the critical review thereof. LZ: undertook interpretation of the results, gene functionality analyses and interpretation, reviewed and approved the final manuscript. TC: did statistical analyses using PLINK and SPSS, reviewed and approved the final manuscript. ZdL: performed the global fibrinolytic assay, critical review of the interpretation of

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

results with regards to clot properties, reviewed and approved the final manuscript. TvZ: dealt with quality control of the BeadXpress® data, critical review of the interpretation of results with regard to lipid mediators, review and approval of the final manuscript. MP: supervised the laboratory analyses of all the haemostatic variables, involved in writing the manuscript and critical review thereof. All authors gave permission for submission to Matrix Biology.

Acknowledgements

This work was supported by grants from the South African National Research Foundation, North-West University (NWU), Population Health Research Institute (PHRI), Medical Research Council, the North West Province Health Department, DNAbiotec Pty Ltd., South African Netherlands Partnerships in Development and the Newton Fund. Funding sources had no involvement in the writing of the report, the collection or interpretation of the data.

We thank all subjects, all supporting staff, the PURE-SA research team, especially Prof Annamarie Kruger, the field workers involved in recruiting the participants, the office staff of the Africa Unit for Transdisciplinary Health Research, the Faculty of Health Sciences, NWU, South Africa, as well as the PURE-International research team, especially Dr. Yusuf, and the PURE-study office staff at the PHRI, Hamilton Health Sciences and McMaster University, ON, Canada. We would also like to thank the staff of the Profiles of Resistance to Insulin in Multiple Ethnicities and Regions study from the Centre for Genome Research and DNAbiotec Pty Ltd.

Appendix A. Supplementary data

Supplementary data to this article can be found online at doi:10.1016/j.matbio.2016.10.005.

Received 7 June 2016;
Received in revised form 20 September 2016;
Accepted 13 October 2016
Available online xxxx

Keywords:
Fibrinolysis;
Gamma prime fibrinogen;
Transcription factor binding sites;
Lipids;
Turbidimetry

†H.T.C. and C.N-R. contributed equally to this manuscript.

References

[1] A. Henschen, F. Lottspeich, M. Kehl, C. Southan, Covalent structure of fibrinogen, Ann. N. Y. Acad. Sci. 408 (1983) 28–43.
[2] C. Wolfenstein-Todel, M.W. Mosesson, Human plasma fibrinogen heterogeneity: evidence for an extended carboxyl-terminal sequence in a normal gamma chain variant (gamma′), Proc. Natl. Acad. Sci. U. S. A. 77 (1980) 5069–5073.
[3] D.W. Chung, E.W. Davie, gamma and gamma′ chains of human fibrinogen are produced by alternative mRNA processing, Biochemistry 23 (1984) 4232–4236.
[4] Z. Yang, I. Mochalkin, R.F. Doolittle, A model of fibrin formation based on crystal structures of fibrinogen and fibrin fragments complexed with synthetic peptides, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 14156–14161.
[5] P. Allan, S. Uitte de Willige, R.H. Abou-Saleh, S.D. Connell, R.A. Ariens, Evidence that fibrinogen gamma′ directly interferes with protofibril growth: implications for fibrin structure and clot stiffness, J. Thromb. Haemost. 10 (2012) 1072–1080.
[6] M. Pieters, R.C. Kotze, J.C. Jerling, A. Kruger, R.A. Ariëns, Evidence that fibrinogen γ′ regulates plasma clot structure and lysis and relationship to cardiovascular risk factors in black Africans, Blood 121 (2013) 3254–3260.
[7] K.R. Machlus, J.C. Cardenas, F.C. Church, A.S. Wolberg, Causal relationship between hyperfibrinogenemia, thrombosis, and resistance to thrombolysis in mice, Blood 117 (2011) 4953–4963.
[8] A. Undas, R.A. Ariens, Fibrin clot structure and function: a role in the pathophysiology of arterial and venous thromboembolic diseases, Arterioscler. Thromb. Vasc. Biol. 31 (2011) e88–e99.
[9] J. Danesh, S. Lewington, S.G. Thompson, G.D. Lowe, R. Collins, J.B. Kostis, A.C. Wilson, A.R. Folsom, K. Wu, M. Benderly, U. Goldbourt, J. Willeit, S. Kiechl, J.W. Yarnell, P.M. Sweetnam, P.C. Elwood, M. Cushman, B.M. Psaty, R.P. Tracy, A. Tybjaerg-Hansen, F. Haverkate, M.P. de Maat, F.G. Fowkes, A.J. Lee, F.B. Smith, V. Salomaa, K. Harald, R. Rasi, E. Vahtera, P. Jousilahti, J. Pekkanen, R. D'Agostino, W.B. Kannel, P.W. Wilson, G. Tofler, C.L. Arocha-Pinango, A. Rodriguez-Larralde, E. Nagy, M. Mijares, R. Espinosa, E. Rodriquez-Roa, E. Ryder, M.P. Diez-Ewald, G. Campos, V. Fernandez, E. Torres, R. Marchioli, F. Valagussa, A. Rosengren, L. Wilhelmsen, G. Lappas, H. Eriksson, P. Cremer, D. Nagel, J.D. Curb, B. Rodriguez, K. Yano, J.T. Salonen, K. Nyyssonen, T.P. Tuomainen, B. Hedblad, P. Lind, H. Loewel, W. Koenig, T.W. Meade, J.A. Cooper, B. De Stavola, C. Knottenbelt, G.J. Miller, J.A. Cooper, K.A. Bauer, R.D. Rosenberg, S. Sato, A. Kitamura, Y. Naito, T. Palosuo, P. Ducimetiere, P. Amouyel, D. Arveiler, A.E. Evans, J. Ferrieres, I. Juhan-Vague, A. Bingham, H. Schulte, G. Assmann, B. Cantin, B. Lamarche, J.P. Despres, G.R. Dagenais, H. Tunstall-Pedoe, M. Woodward, Y. Ben-Shlomo, G.D. Smith, V. Palmieri, J.L. Yeh, A. Rudnicka, P. Ridker, F. Rodeghiero, A. Tosetto, J. Shepherd, I. Ford, M. Robertson, E. Brunner, M. Shipley, E.J. Feskens, D. Kromhout, A. Dickinson, B. Ireland, K. Juzwishin, S. Kaptoge, S. Lewington, A. Memon, N. Sarwar, M. Walker, J. Wheeler, I. White, A. Wood, Plasma fibrinogen level and the risk of major cardiovascular diseases and nonvascular mortality: an individual participant meta-analysis, JAMA 294 (2005) 1799–1809.

[10] A. Hamsten, U. De Faire, L. Iselius, M. Blombäck, Genetic and cultural inheritance of plasma fibrinogen concentration, Lancet 330 (1987) 988–991.
[11] T. Reed, R.P. Tracy, R.R. Fabsitz, Minimal genetic influences on plasma fibrinogen level in adult males in the NHLBI twin study, Clin. Genet. 45 (1994) 71–77.
[12] Y. Friedlander, Y. Elkana, R. Sinnreich, J.D. Kark, Genetic and environmental sources of fibrinogen variability in Israeli families: the Kibbutzim Family Study, Am. J. Hum. Genet. 56 (1995) 1194.
[13] G. Livshits, G. Schettler, E. Graff, M. Blettner, J. Wahrendorf, D. Brunner, Tel Aviv-Heidelberg three-generation offspring study: genetic determinants of plasma fibrinogen level, Am. J. Med. Genet. 63 (1996) 509–517.
[14] J.S. Pankow, A.R. Folsom, M.A. Province, D. Rao, R.R. Williams, J. Eckfeldt, T.A. Sellers, Segregation analysis of plasminogen activator inhibitor-1 and fibrinogen levels in the NHLBI family heart study, Arterioscler. Thromb. Vasc. Biol. 18 (1998) 1559–1567.
[15] M. de Lange, H. Snieder, R.A.S. Ariëns, T.D. Spector, P.J. Grant, The genetics of haemostasis: a twin study, Lancet 357 (2001) 101–105.
[16] M. Neijts, J. van Dongen, C. Kluft, D.I. Boomsma, G. Willemsen, E.J. de Geus, Genetic architecture of the pro-inflammatory state in an extended twin-family design, Twin Res. Hum. genet. 16 (2013) 931–940.
[17] M. Sabater-Lleal, J. Huang, D. Chasman, S. Naitza, A. Dehghan, A.D. Johnson, A. Teumer, A.P. Reiner, L. Folkersen, S. Basu, A.R. Rudnicka, S. Trompet, A. Malarstig, J. Baumert, J.C. Bis, X. Guo, J.J. Hottenga, S.Y. Shin, L.M. Lopez, J. Lahti, T. Tanaka, L.R. Yanek, T. Oudot-Mellakh, J.F. Wilson, P. Navarro, J.E. Huffman, T. Zemunik, S. Redline, R. Mehra, D. Pulanic, I. Rudan, A.F. Wright, I. Kolcic, O. Polasek, S.H. Wild, H. Campbell, J.D. Curb, R. Wallace, S. Liu, C.B. Eaton, D.M. Becker, L.C. Becker, S. Bandinelli, K. Raikkonen, E. Widen, A. Palotie, M. Fornage, D. Green, M. Gross, G. Davies, S.E. Harris, D.C. Liewald, J.M. Starr, F.M. Williams, P.J. Grant, T.D. Spector, R.J. Strawbridge, A. Silveira, B. Sennblad, F. Rivadeneira, A.G. Uitterlinden, O.H. Franco, A. Hofman, J. van Dongen, G. Willemsen, D.I. Boomsma, J. Yao, N.S. Jenny, T. Haritunians, B. McKnight, T. Lumley, K.D. Taylor, J.I. Rotter, B.M. Psaty, A. Peters, C. Gieger, T. Illig, A. Grotevendt, G. Homuth, H. Volzke, T. Kocher, A. Goel, M.G. Franzosi, U. Seedorf, R. Clarke, M. Steri, K.V. Tarasov, S. Sanna, D. Schlessinger, D.J. Stott, N. Sattar, B.M. Buckley, A. Rumley, G.D. Lowe, W.L. McArdle, M.H. Chen, G.H. Tofler, J. Song, E. Boerwinkle, A.R. Folsom, L.M. Rose, A. Franco-Cereceda, M. Teichert, M.A. Ikram, T.H. Mosley, S. Bevan, M. Dichgans, P.M. Rothwell, C.L. Sudlow, J.C. Hopewell, J.C. Chambers, D. Saleheen, J.S. Kooner, J. Danesh, C.P. Nelson, J. Erdmann, M.P. Reilly, S. Kathiresan, H. Schunkert, P.E. Morange, L. Ferrucci, J.G. Eriksson, D. Jacobs, I.J. Deary, N. Soranzo, J.C. Witteman, E.J. de Geus, R.P. Tracy, C. Hayward, W. Koenig, F. Cucca, J.W. Jukema, P. Eriksson, S. Seshadri, H.S. Markus, H. Watkins, N.J. Samani, H. Wallaschofski, N.L. Smith, D. Tregouet, P.M. Ridker, W. Tang, D.P. Strachan, A. Hamsten, C.J. O'Donnell, Multiethnic meta-analysis of genome-wide association studies in >100 000 subjects identifies 23 fibrinogen-associated Loci but no strong evidence of a causal association between circulating fibrinogen and cardiovascular disease, Circulation 128 (2013) 1310–1324.
[18] F. Dudbridge, Power and predictive accuracy of polygenic risk scores, PLoS Genet. 9 (2013), e1003348.
[19] Y. Li, P. Agarwal, A pathway-based view of human diseases and disease relationships, PLoS One 4 (2009), e4346.
[20] B. Childs, D. Valle, Genetics, biology and disease, Annu. Rev. Genomics Hum. Genet. 1 (2000) 1–19.
[21] J. Vockley, P. Rinaldo, M.J. Bennett, D. Matern, G.D. Vladutiu, Synergistic heterozygosity: disease resulting from multiple partial defects in one or more metabolic pathways, Mol. Genet. Metab. 71 (2000) 10–18.
[22] H. Han, H. Shim, D. Shin, J.E. Shim, Y. Ko, J. Shin, H. Kim, A. Cho, E. Kim, T. Lee, TRRUST: a reference database of human transcriptional regulatory interactions, Sci. Rep. 5 (2015).
[23] S.W. Doniger, J.C. Fay, Frequent gain and loss of functional transcription factor binding sites, PLoS Comput. Biol. 3 (2007), e99.
[24] M. Tuğrul, T. Paixão, N.H. Barton, G. Tkačik, Dynamics of transcription factor binding site evolution, PLoS Genet. 11 (2015), e1005639.
[25] S. Kaptoge, I. White, S. Thompson, A. Wood, S. Lewington, G. Lowe, J. Danesh, Associations of plasma fibrinogen levels with established cardiovascular disease risk factors, inflammatory markers, and other characteristics: individual participant meta-analysis of 154,211 adults in 31 prospective studies: the fibrinogen studies collaboration, Am. J. Epidemiol. 166 (2007) 867–879.
[26] A.R. Folsom, K. Wu, M. Conlan, A. Finch, C. Davis, G. Marcucci, P. Sorlie, M. Szklo, Distributions of hemostatic variables in blacks and whites: population reference values from the Atherosclerosis Risk in Communities (ARIC) Study, Ethn. Dis. 2 (1991) 35–46.
[27] A. Greyling, M. Pieters, T. Hoekstra, W. Oosthuizen, A. Schutte, Differences in the association of PAI-1 activity with the metabolic syndrome between African and Caucasian women, Nutr. Metab. Cardiovasc. Dis. 17 (2007) 499–507.
[28] M. Pieters, H.H. Vorster, Nutrition and hemostasis: a focus on urbanization in South Africa, Mol. Nutr. Food Res. 52 (2008) 164–172.
[29] M. Hamer, R. von Känel, M. Reimann, N.T. Malan, A.E. Schutte, H.W. Huisman, L. Malan, Progression of cardiovascular risk factors in black Africans: 3 year follow up of the SABPA cohort study, Atherosclerosis 238 (2015) 52–54.
[30] Y.-S. Chen, A. Torroni, L. Excoffier, A.S. Santachiara-Benerecetti, D.C. Wallace, Analysis of mtDNA variation in African populations reveals the most ancient of all human continent-specific haplogroups, Am. J. Hum. Genet. 57 (1995) 133.
[31] S.C. Schuster, W. Miller, A. Ratan, L.P. Tomsho, B. Giardine, L.R. Kasson, R.S. Harris, D.C. Petersen, F. Zhao, J. Qi, Complete Khoisan and bantu genomes from southern Africa, Nature 463 (2010) 943–947.
[32] R.C. Kotze, C. Nienaber-Rousseau, Z. De Lange, M.P. De Maat, T. Hoekstra, M. Pieters, Genetic polymorphisms influencing total and gamma′ fibrinogen levels and fibrin clot properties in Africans, Br. J. Haematol. 168 (2015) 102–112.
[33] A. Yates, W. Akanni, M.R. Amode, D. Barrell, K. Billis, D. Carvalho-Silva, C. Cummins, P. Clapham, S. Fitzgerald, L. Gil, Ensembl 2016, Nucleic Acids Res. 44 (2016) D710–D716.
[34] T. van Zyl, J.C. Jerling, K.R. Conradie, E.J. Feskens, Common and rare single nucleotide polymorphisms in the LDLR gene are present in a black South African population and associate with low-density lipoprotein cholesterol levels, J. Hum. Genet. 59 (2014) 88–94.
[35] M. Abifadel, M. Varret, J.-P. Rabès, D. Allard, K. Ouguerram, M. Devillers, C. Cruaud, S. Benjannet, L. Wickham, D. Erlich, Mutations in PCSK9 cause autosomal dominant hypercholesterolemia, Nat. Genet. 34 (2003) 154–156.

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005

[36] N.G. Seidah, Z. Awan, M. Chrétien, M. Mbikay, PCSK9 a key modulator of cardiovascular health, Circ. Res. 114 (2014) 1022–1036.
[37] Y. Zhang, C.-G. Zhu, R.-X. Xu, S. Li, Y.-L. Guo, J. Sun, J.-J. Li, Relation of circulating PCSK9 concentration to fibrinogen in patients with stable coronary artery disease, J. Clin. Lipidol. 8 (2014) 494–500.
[38] R.S. Rosenson, G.D. Lowe, Effects of lipids and lipoproteins on thrombosis and rheology, Atherosclerosis 140 (1998) 271–280.
[39] N. Martinelli, D. Girelli, B. Lunghi, M. Pinotti, G. Marchetti, G. Malerba, P.F. Pignatti, R. Corrocher, O. Olivieri, F. Bernardi, Polymorphisms at LDLR locus may be associated with coronary artery disease through modulation of coagulation factor VIII activity and independently from lipid profile, Blood 116 (2010) 5688–5697.
[40] S.J. Schrodi, S. Mukherjee, Y. Shan, G. Tromp, J.J. Sninsky, A.P. Callear, T.C. Carter, Z. Ye, J.L. Haines, M.H. Brilliant, Genetic-based prediction of disease traits: prediction is very difficult, especially about the future, Front. Genet. 5 (2014) 1–18.
[41] L. Wang, Z.Q. Yao, J.P. Moorman, Y. Xu, S. Ning, Gene expression profiling identifies IRF4-associated molecular signatures in hematological malignancies, PLoS One 9 (2014), e106788.
[42] D.-S. Jiang, Z.-Y. Bian, Y. Zhang, S.-M. Zhang, Y. Liu, R. Zhang, Y. Chen, Q. Yang, X.-D. Zhang, G.-C. Fan, Role of interferon regulatory factor 4 in the regulation of pathological cardiac hypertrophy, Hypertension 61 (2013) 1193–1202.
[43] D. Vaiman, R. Calicchio, F. Miralles, Landscape of transcriptional deregulations in the preeclamptic placenta, PLoS One 8 (2013), e65498.
[44] S.-F. Yan, N. Mackman, W. Kisiel, D.M. Stern, D.J. Pinsky, Hypoxia/hypoxemia-induced activation of the procoagulant pathways and the pathogenesis of ischemia-associated thrombosis, Arterioscler. Thromb. Vasc. Biol. 19 (1999) 2029–2035.
[45] C.M. Rein-Smith, N.W. Anderson, D.H. Farrell, Differential regulation of fibrinogen γ chain splice isoforms by interleukin-6, Thromb. Res. 131 (2013) 89–93.
[46] K.S. Alexander, T.E. Madden, D.H. Farrell, Association between γ′ fibrinogen levels and inflammation, J. Thromb. Haemost. 105 (2011) 605.
[47] K. Teo, C.K. Chow, M. Vaz, S. Rangarajan, S. Yusuf, The Prospective Urban Rural Epidemiology (PURE) study: examining the impact of societal influences on chronic noncommunicable diseases in low-, middle-, and high-income countries, Am. Heart J. 158 (2009) 1–7, e1.
[48] M. Pieters, M.P. De Maat, J.C. Jerling, T. Hoekstra, A. Kruger, Fibrinogen concentration and its role in CVD risk in black South Africans–effect of urbanisation, J. Thromb. Haemost. 106 (2011) 448–456.
[49] Z. de Lange, D.C. Rijken, T. Hoekstra, K.R. Conradie, J.C. Jerling, M. Pieters, In black South Africans from rural and urban communities, the 4G/5G PAI-1 polymorphism influences PAI-1 activity, but not plasma clot lysis time, PLoS One 8 (2013), e83151.
[50] C. Nienaber-Rousseau, Z. de Lange, M. Pieters, Homocysteine influences blood clot properties alone and in combination with total fibrinogen but not with fibrinogen gamma′ in Africans, Blood Coagul. Fibrinolysis 26 (2015) 389–395.
[51] T. Lisman, P.G. de Groot, J.C. Meijers, F.R. Rosendaal, Reduced plasma fibrinolytic potential is a risk factor for venous thrombosis, Blood 105 (2005) 1102–1105.
[52] A. Jobse, M. Pieters, C. Nienaber-Rousseau, H. Boshuizen, T. Hoekstra, M.P.M. de Maat, The contribution of genetic and environmental factors to changes in total and γ′ fibrinogen over 5 years, Thromb. Haemost. 135 (2015) 703–709.
[53] C. Nienaber-Rousseau, B. Swanepoel, R.C. Dolman, M. Pieters, K.R. Conradie, G.W. Towers, Interactions between C-reactive protein genotypes with markers of nutritional status in relation to inflammation, Nutrients 6 (2014) 5034–5050.
[54] C. Nienaber-Rousseau, S.M. Ellis, S.J. Moss, A. Melse-Boonstra, G.W. Towers, Gene–environment and gene–gene interactions of specific MTHFR, MTR and CBS gene variants in relation to homocysteine in black South Africans, Gene 530 (2013) 113–118.
[55] S. Purcell, B. Neale, K. Todd-Brown, L. Thomas, M.A. Ferreira, D. Bender, J. Maller, P. Sklar, P.I. De Bakker, M.J. Daly, PLINK: a tool set for whole-genome association and population-based linkage analyses, Am. J. Hum. Genet. 81 (2007) 559–575.
[56] T.V. Perneger, What's wrong with Bonferroni adjustments, BMJ 316 (1998) 1236–1238.
[57] J.A. Smith, E.B. Ware, P. Middha, L. Beacher, S.L. Kardia, Current applications of genetic risk scores to cardiovascular outcomes and subclinical phenotypes, Curr. Epidemiol. Rep. 2 (2015) 180–190.
[58] M.N. Badsi, S. Mediene-Benchekor, H. Ouhaibi-Djellouli, S.A. Lardjam-Hetraf, H. Boulenouar, D.N. Meroufel, X. Hermant, I. Hamani-Medjaoui, N. Saidi-Mehtar, P. Amouyel, L. Houti, A. Meirhaeghe, L. Goumidi, Combined effect of established BMI loci on obesity-related traits in an Algerian population sample, BMC Genet. 15 (2014).

Please cite this article as: H. T. Cronjé, et al., Candidate gene analysis of the fibrinogen phenotype reveals the importance of polygenic co-regulation, Matrix Biol (2016), http://dx.doi.org/10.1016/j.matbio.2016.10.005