The Molecular Characterisation of the *Annexin II* Gene in Pre-eclampsia

by

Jacoba Martina de Jager

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science (MSc) in Genetics at Stellenbosch University

Supervisor: Prof L Warnich
Co-Supervisor: Dr R Hillermann-Rebello
Co-Supervisor: Dr GS Gebhardt

December 2009

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 17 November 2009

Copyright © 2009 Stellenbosch University

All rights reserved

ABSTRACT

The hypertensive conditions of pregnancy (including pre-eclampsia (PE)) is the leading cause of primary obstetric death in South Africa and affects at least five percent of pregnancies in the Western Cape province. Reduced levels of placental protein 13 (PP13) early in pregnancy are associated with a higher incidence of PE in later gestation. PP13 and annexin II have been co-localised to the brush border membrane of syncytiotrophoblasts, and form a complex that is transported to the maternal circulation. It is speculated that genetic variation in the gene encoding annexin II (*ANXA2*) could underlie the reduced PP13 levels.

The aim of this study was to screen the *ANXA2* gene, including the proximal promoter region, in two South African population groups, (Mixed Ancestry and Black) from the Western Cape, to identify whether variants in the *ANXA2* gene confer susceptibility to PE. The study cohort comprised of 120 pre-eclamptic maternal, 94 pre-eclamptic fetal and 54 healthy control individuals. Genomic DNA of patient and control individuals was extracted for PCR amplification of *ANXA2* and Multiphor SSCP/HD analysis was performed for mutation detection. The conformational variants identified were subjected to automated DNA sequencing and subsequently to RFLP analysis, to confirm the genotypes in the remainder of the cohort. Nine previously identified variants (c.-31 T>C, c.292 G>T; p.Val98Leu, c.975 C>T;p.Gly325Gly, c.-12+75 C>A, c.-11-43 G>A, c.-11-13 A>T, c.48+67 C>T, c.449-17 G>A, c.683-56 G>A) and 16 novel variants (c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG, c.-135 C>G, c.-92 A>T, c.222 C>T; p.Ala74Ala, c.600 C>T; p.Asp110Asp, c.934 G>A; p.Gly312Ser, c.244-42 G>C, c.244-76 C>G, c.528+38 C>T, c.589-5 C>T, c.682+49 C>T, c.961-30 A>G, c.961-24 C>G, c.*1057 A>G) were identified upon screening the *ANXA2* gene. Statistical analysis identified significant association at five loci: SNP c.-92 A>T located within the *ANXA2* 5'UTR, exonic SNP c.222 C>T; p.Ala74Ala and three intronic SNPs c.244-76 C>G, c.449-17 G>A and c.589-5 C>T. Three of the five variants (c.-92 A>T, c.244-76 C>G, c.589-5 C>T) were significantly associated with PE (P<0.05) and could contribute to PE susceptibility in these two SA

populations, whereas the other two variants (c.222 C>T; p.Ala74Ala, c.449-17 G>A) revealed a possible protective effect, suggesting a reduced risk of developing PE. *In silico* analysis predicted the disruption and creation of several putative transcription factor binding sites by three SNPs in the *ANXA2* gene, which could subsequently affect *ANXA2* functioning.

This study provides evidence for genetic variation in the *ANXA2* gene, which warrants functional experimental validation in an attempt to investigate the function of these SNPs in molecular, cellular and physiological processes underlying PE. Identifying an association between variants in the *ANXA2* gene and PE could contribute to the development of an additional early biomarker. The early identification of PE would promote the South African health system by providing the appropriate health care support and monitoring of high risk pregnancies, which could ultimately result in improved pregnancy outcome.

OPSOMMING

Die hipertensiewe siektes van swangerskap (insluitende pre-eklampsie (PE)) is die belangrikste direkte oorsaak van moedersterftes in Suid-Afrika en dit kom voor by ongeveer 5% van swangerskape in die Wes-Kaap provinsie. Verlaagde plasentale proteïen 13 (PP13) vlakke tydens vroeë swangerskap word verbind met ‘n hoër voorkoms van PE in latere swangerskap. PP13 en anneksin II kom albei op die borselgrens membraan van synsitiotrofoblaste voor waar hulle ‘n kompleks vorm wat na die moederlike sirkulasie vervoer word. Daar word gespekuleer dat die onderliggende oorsaak vir laer PP13 vlakke as gevolg van genetiese variasie in die geen wat anneksin II kodeer (*ANXA2)* kan wees.

Die doel van hierdie studie was om die *ANXA2* geen, insluitende die proksimale promoter area, in twee Suid-Afrikaanse populasie groepe, (Kleurling en Swart) van die Wes-Kaap, te skandeer met die doel om variante in die *ANXA2* geen te identifiseer en ’n moontlike assosiasie met die vatbaarheid vir PE te bepaal. Hierdie studie populasie het bestaan uit 120 pre-eklamptiese vroue, 94 neonate van pre-eklamptiese ma’s en 54 gesonde kontrole individue. Genomiese DNS van die pasiënte en kontrole individue is geëkstraeer vir polimerase kettingreaksie amplifikasie van die *ANXA2* geen, waarna Multiphor enkelstring konformasie polimorfisme heterodupleks analise uitgevoer is met die doel om DNS variante te identifiseer. Die verskillende konformasies waargeneem is onderwerp aan semi-geoutomatiseerde DNS volgorde bepalingsanalise en gevolglik restriksie fragment lengte polimorfisme analise om genotipes in die res van die studiegroep te bevestig.

Vyf-en-twintig variante is geïdentifiseer met die skandering van die *ANXA2* geen, waarvan nege voorheen geïdentifiseer is (c.-31 T>C, c.292 G>T; p.Val98Leu, c.975 C>T;p.Gly325Gly, c.-12+75 C>A, c.-11-43 G>A, c.-11-13 A>T, c.48+67 C>T, c.449-17 G>A, c.683-56 G>A) en 16 nuwe variante is (c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG, c.-135 C>G, c.-92 A>T, c.222 C>T; p.Ala74Ala, c.600 C>T; p.Asp110Asp, c.934 G>A; p.Gly312Ser, c.244-42 G>C, c.244-76 C>G, c.528+38 C>T, c.589-5 C>T,

c.682+49 C>T, c.961-30 A>G, c.961-24 C>G, c.*1057 A>G). Statistiese analise het ‘n statisties beduidende assosiasie met vyf SNPs geïdentifiseer: SNP c.-92 A>T geleë in die *ANXA2* 5’UTR, die koderende SNP c.222 C>T; p.Ala74Ala en drie SNPs c.244-76 C>G, c.449-17 G>A and c.589-5 C>T geleë in die nie-koderende areas. Drie van hierdie vyf SNPs (c.-92 A>T, c.244-76 C>G, c.589-5 C>T) het statisties beduidende assosiasie met PE (P<0.05) getoon en kan bedra tot die vatbaarheid vir PE in hierdie twee Suid-Afrikaanse populasies, terwyl die ander twee SNPs (c.222 C>T; p.Ala74Ala, c.449-17 G>A) ‘n moontlike beskermende effek gedui het, wat ‘n verlaagde risiko vir die ontwikkeling van PE voorstel. *In silico* analise het voorspel dat verskeie voorgestelde transkripsiefaktor bindingsetels onderbreek of geskep sal word in die teenwoordigheid van drie SNPs in die *ANXA2* geen, wat gevolglik die funksionering van *ANXA2* kan affekteer.

Hierdie studie verskaf bewyse vir genetiese variasie in die *ANXA2* geen, wat verdere funksionele eksperimentele ondersoeke vereis om die funksie van hierdie SNPs in molekulêre, sellulêre en fisiologiese prosesse onderliggend aan PE te bepaal. Die identifisering van ’n assosiasie tussen variante in die *ANXA2* geen en PE kan bydra tot die ontwikkeling van ’n addisionele vroeë genetiese merker. Die vroeë identifisering van PE kan die Suid-Afrikaanse gesondheidsisteem geweldig baat deurdat die geskikte gesondheidsorg en ondersteuning asook deurgaanse monitering van hoë risiko swangerskappe verskaf sal kan word. Dit kan uiteindelik lei tot ’n verbeterde uitkoms vir swangerskappe in Suid-Afrika.

ACKNOWLEDGEMENTS

I would like to express my gratitude towards the following individuals and institutions:

NRF Thuthuka – for funding this project and thereby making it possible.

Stellenbosch University, Department of Genetics, for providing the infra structure and facilities for conducting this study.

My supervisor Prof Warnich, for taking me on as a student in her lab, proofreading the many copies of this thesis and for her support, encouragement and advice during this study.

My co-supervisors Dr Hillermann-Rebello and Dr Gebhardt, for proofreading this thesis and for their valuable input on this subject and study in this interesting field of genetics.

My lab friends: Mauritz, Marika, Brit, Galen, Nats, Vero, Nathan and Jess for keeping the spirits up in the lab and all the fun and laughter.

Mauritz, for all your advice, guidance, many coffees and most certainly your positive attitude towards science and life in general, which is absolutely contagious.

Marika, for all your advice, support, guidance and enthusiasm, thank you, I will always appreciate what you have done for me.

My parents, brother and sisters for your love, care and understanding, your emotional support and for always believing in me, I love you dearly.

Pierre, my fiancé, for being so incredibly patient throughout this study and for your perpetual love and understanding. Your support carried me through this two years, I love you.

God, my Heavenly Father for blessing me with this opportunity and giving me the strength, courage and endurance to see it through to the best of my capabilities.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: LITERATURE REVIEW

CHAPTER 2: METHODS AND MATERIALS

CHAPTER 3: RESULTS

LIST OF TABLES

CHAPTER 2: METHODS AND MATERIALS

CHAPTER 3: RESULTS

LIST OF ABBREVIATIONS

α	alpha
~	approximately
@	at
β	beta
χ^2	chi-squared
˚C	degrees celcius
=	equal to
5’	5-prime
5’mC	5-methylcytosine
3’	3-prime
>	greater than, also/and to indicate change (from one nucleotide to another)
<	less than
µg	microgram
µl	microlitre
µM	micro molar
µg/l	microgram per litre
%	percentage
%C	percentage crosslinking
®	registered trademark
™	trademark
×	times

A	adenine or adenosine
AA	anaesthetic related
AB	abortion
AC	acute collapse
AGT	angiotensinogen
AIDS	acquired immune deficiency syndrome
Ala	alanine
ANXA2	*annexin II* gene
ANXA5	*annexin V* gene

AP-1	activator protein 1
APH	antepartum haemorrhage
aPL	amniotic phospholipids
APS	ammonium persulphate: $(NH_4)_2S_2O_8$
Asp	aspartic acid
ASSP	Alternative Splice Site Predictor
ATG	translation initiation site
BAA	bis-acrylamide; N,N'-methylene-bis-acrylamide: $C_7H_{10}O_2N_2$
BLAST	Basic Local Alignment of Sequences Tool
bp	base-pair
BRCA1	breast cancer susceptibility gene 1
BTV	bring to volume
C	cytosine
Ca^{2+}	calsium
CRDs	carbohydrate recognition domains
CVD	cardio vascular disease
Cys	cysteine
C-termini	carboxy terminal
D'	Normalised Linkage Disequilibrium Parameter
dATP	2'-deoxyadenosine-5'-triphosphate
dbSNP	database single nucleotide polymorphism
dCTP	2'-deoxycytosine-5'triphosphate
dGTP	2'deoxyguanine-5'-triphosphate
dH_2O	distilled water
dHPLC	denaturing high performance liquid chromatography
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTP	2'-deoxy-nucleotide-5'-triphosphate
DTL	Diagnostic Technologies Limited
dTTP	2'deoxythymine-5'-triphosphate

EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbant assay
Emb	embolism
eNOS	endothelial nitric oxide synthase
Ep	ectopic pregnancy
ESE	Exonic Splicing Enhancer
et. al.	and others
EtBr	ethidium bromide: $C_{21}H_{20}BrN_3$
$EtOH_2$	ethanol: CH_3CH_2OH
F	forward primer
Flt-1	Fms related tyrosine kinase 1
g	gram
G	guanosine
Glu	glutamic acid
Gly	glycine
h	hour
HD	heteroduplex
HIV	human immunodeficiency virus
HT	hypertension
HWE	Hardy-Weinberg Equilibrium
IDT	Integrated DNA Technologies
in silico	refers to research conducted using computers only
Inc.	Incorporated
ins	insertion
ISSHP	International Society for the Study of Hypertension in Pregnancy
IUGR	intrauterine growth restriction
kb	kilobases

kDa	kilodaltons
KDR	kinase insert domain receptor
l	litre
LD	Linkage Disequilibrium
LGALS13	lectin, galactose-binding, soluble 13 gene
LOD	Logarithm of the Odds
Ltd	limited
M	moles per litre/ molar
MD	pre-existing medical disease
MF	mutagenic forward primer
mg	milligram
mg/ml	milligram per millilitre
$MgCl_2$	magnesium chloride
min	minutes
ml	millilitre
mM	millimoles per litre/ millimolar
mm	millimetre
mmHg	millimetre of mercury
mmol/l	milli-moles per litre
mRNA	messenger ribonucleic acid
MTHFR	methylenetetrahydrofolate reductase
n	number of individuals
n/a	not applicable
NCBI	National Centre for Biotechnology Information
NCCEMD	National Committee on Confidential Enquiries into Maternal Deaths in South Africa
ng	nanogram
ng/ml	nanogram per millilitre
NPRI	non-pregnancy related infections
nt	nucleotide

N	amino
NF-1	nuclear factor 1
P	probability
p	short arm of chromosome
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PDA	piperizine diacrylamide
PE	pre-eclampsia
pH	potential of hydrogen
PIGF	proangiogenic placental growth factor
PIH	pregnancy-induced hypertension
pmol	pico mole
PP13	placental protein 13
PP13-B	placental protein 13 purified from human placenta
PP13-R	recombinant placental protein 13
PPH	postpartum haemorrhage
PRS	pregnancy related sepsis
q	long arm of chromosome
R	reverse primer or arginine
r^2	correlation coefficient between two loci
RBC	red blood cells
RE	restriction enzymes
RFLP	restriction fragment length polymorphism
RNA	ribonucleic acid
RPL	recurrent pregnancy loss
rpm	revolutions per minute
rs	reference sequence
s	seconds
SA	South Africa

Ser	serine
sFlt-1	secreted splice variant of FLT-1
sFlt-1	soluble fms-like tyrosine kinase
SNP	single nucleotide polymorphism
SP1	specificity protein 1
SR	serine/arginine
SSCP	single strand conformation polymorphism
SU	Stellenbosch University
3-D	three dimensional
T	thymine or thyrosine
T_a	annealing temperature
Taq	*Thermus aquaticus*
TBE	tris-borate EDTA
TE	tris-EDTA
TEMED	N, N, N' N', -tetramethylethylenediamine
TFIID	transcription factor IID
TFPGA	Tools for Population Genetic Analysis
Thr	threonine
T_m	melting temperature
t-PA	tissue plasminogen activator
Thrp	thryptophane
TRIS	trishydromethylaminomethane
TRIS-HCl	tris hydrochloride
U	unit
UK	United Kingdom
Unk	unknown
USA	United States of America
UTR	untranslated region
UV	ultraviolet
v	version

V	volts
v/v	volume per volume
Val	valine
VEGF	vascular-endothelial growth factor
vs.	versus
w/v	weight per volume
wt	wild type

1 LITERATURE REVIEW

1.1 Pregnancy

1.1.1 Normal Pregnancy

The gestational period “involves an intricate interplay of intracellular and extracellular factors including hormones, adhesion molecules, growth factors and immunomolecules” to allow for a normal and healthy pregnancy to proceed (Shankar *et al.*, 2005).

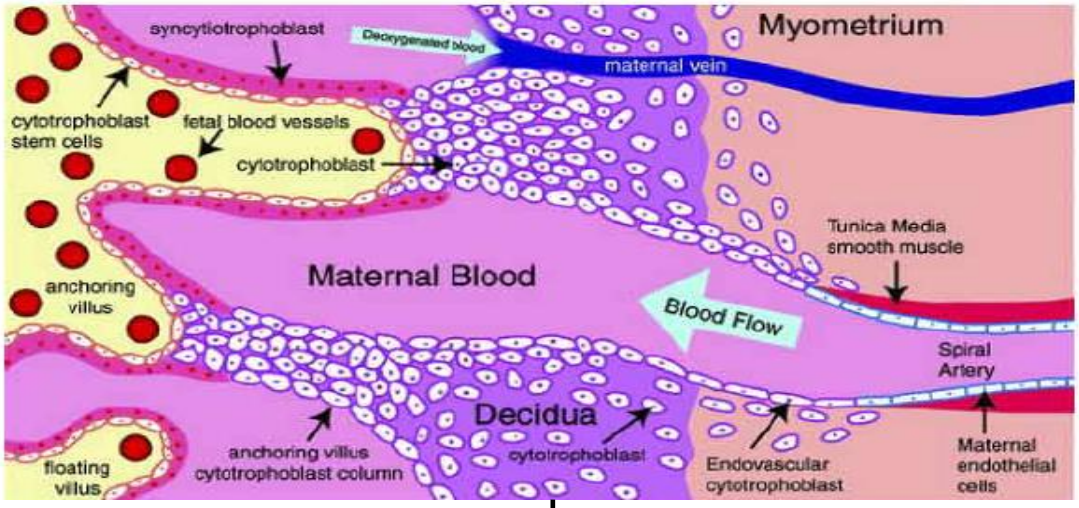
i) Fertilisation, Implantation and Placentation

Following conception, the fertilised ovum undergoes a series of cell divisions, known as cleavage. Three days following fertilisation, further cell divisions result in the formation of a cluster of cells and at this point, the embryo is known as the “morula”. From day five after conception, the morula reaches the uterus and the blastocyst cavity starts to develop. During the following two days, a layer of trophoblasts cells forms an outer layer of cells around the blastocyst, and on the inside of the trophoblast layer of cells, a cluster of cells, known as an inner cell mass (embryoblast) develops. The embryoblast develops into the embryo, while trophoblast cells give rise to extra-embryonic structures such as the placenta and membranes.

The blastocyst sheds the zona pellucida, a glycoprotein membrane enclosing the plasma membrane of the developing oocyte and develops until it is fully exposed to the uterine cavity, after which it adheres and implants into the uterine wall (Martini and Bartholomew, 2003). The attachment of the blastocyst to the endometrial epithelium initiates the differentiation of the trophoblast cells into an inner cytotrophoblast and an outer syncytiotrophoblast cell layer. During the second week of gestation, the blastocyst cavity and implanting embryo have developed into the primary yolk sac and blastocyst, respectively.

Ten days after conception, implantation of the fully developed blastocyst into the uterine lining has taken place (Benirschke *et al.*, 1991). During this time the trophoblast cells divide to form a thicker outer layer. The cell membrane in contact with the endometrium disappears to form an outer layer that consists of multi-nucleated cytoplasm, known as the syncytiotrophoblast layer. As the blastocyst becomes completely imbedded into the uterine wall, further development of the embryo takes place in the endometrium.

The syncytiotrophoblast cells continue to develop and function as the provider and distributor of nutrients to the inner cell mass and play an essential role in the development of the early embryo. Trophoblasts continue to grow and develop, to facilitate a change ("re-modelling") in maternal spiral arteries into low resistance vessels that would cause increased maternal blood to flow through lacunae to ensure sufficient perfusion of the embryo (Figure 1.1a and Figure 1.2) (Caniggia *et al.*, 2000; Norris, 2005).



Normal trophoblast invasion

Uncomplicated pregnancy

Figure 1.1a Trophoblast invasion in a normal pregnancy (Serrano, 2006).

1.1.2 Abnormal Pregnancy

i) Impaired Invasion and Perfusion

In previous studies defective endovascular trophoblast invasion and “shallow” interstitial trophoblast invasion (Redman and Sargent, 2005) have been associated with pregnancy-related disorders such as miscarriage, preterm labour, intra-uterine growth restriction (IUGR) and pre-eclampsia (PE) (Kim *et al.*, 2003; Ozturk *et al*., 2004). Insufficient transformation of maternal spiral arteries leads to a reduction in intervillous blood flow and could lead to a subsequent reduction in the concentration of oxygen and nutrients supplied to the fetus (Caniggia *et al.*, 2000).

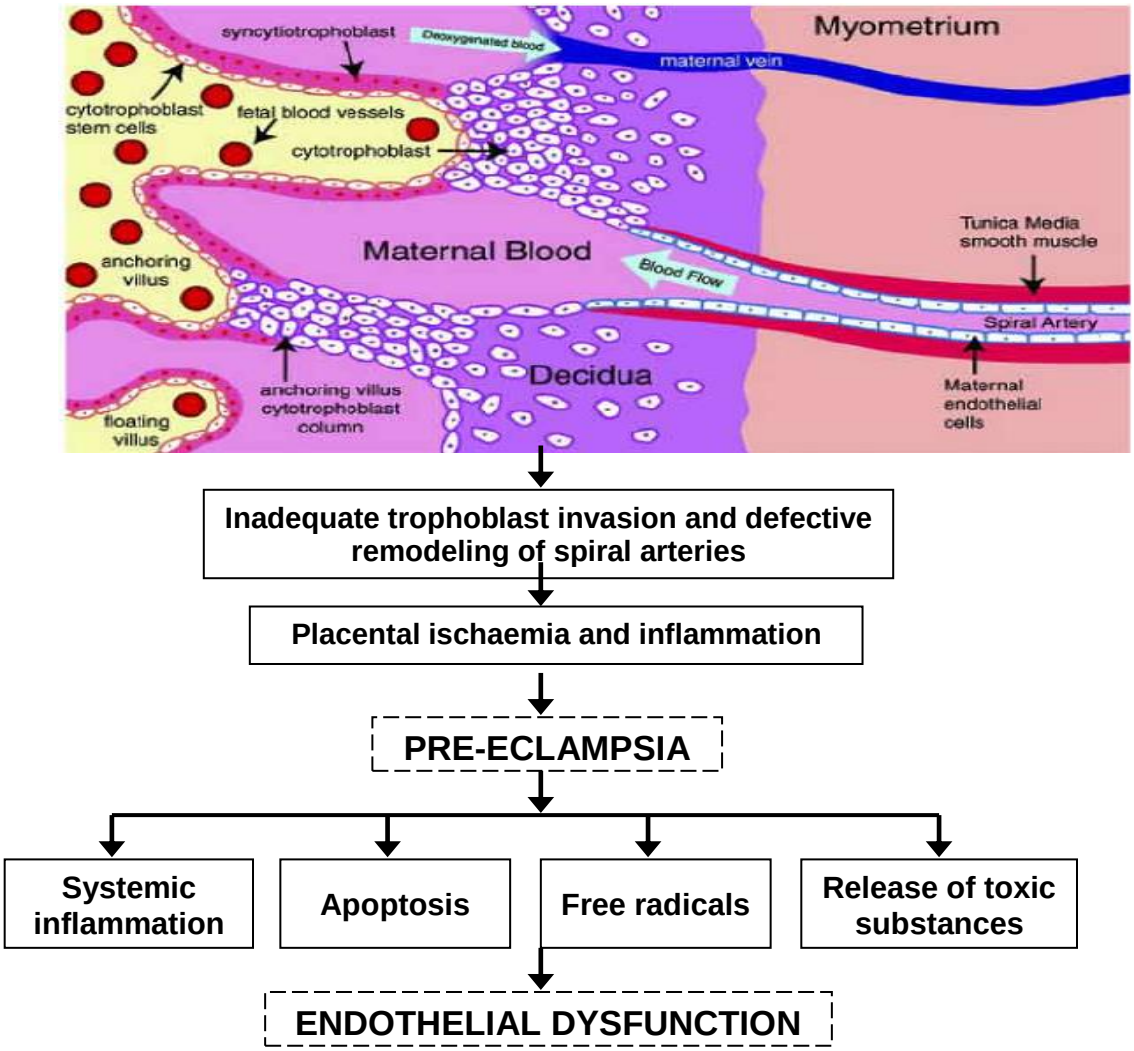


Figure 1.1b Inadequate trophoblast invasion in pre-eclamptic pregnancies (Serrano, 2006).

A certain systemic concentration of the cytokines VEGF (vascular endothelial growth factor) and PIGF (placental growth factor), produced by natural-killer cells, are maintained during a normal pregnancy. These cytokines are angiogenic factors which play an imperative role in maintaining the maternal endothelial state and uteroplacental circulation during pregnancy, by binding to the receptors Flt-1 (Fms related tyrosine kinase 1) and KDR (kinase insert domain receptor) (Maynard *et al.*, 2003). Normal pregnancies have lower sFlt-1 (secreted splice variant of FLT-1) concentrations compared to increased sFlt-1 levels in the placenta and serum of women with PE (Maynard *et al.*, 2003; Buhimschi *et al.*, 2005). Elevated systemic levels of sFlt-1 are associated with PE, since the levels normalise with the delivery of the placenta. During PE, the systemic concentration of free VEGF and PIGF are decreased, most likely because increased levels of the soluble form of placenta-derived sFlt-1, an antagonist of VEGF, binds and inhibits VEGF and PIGF (Figure 1.2) (Burri *et al.*, 2004; Sibai *et al.*, 2005). VEGF factors either bind to trans-membrane receptors or sFlt-1, which in both cases, result in the inhibition of VEGF activity (Clark *et al.*, 1998; Maynard *et al.*, 2003). This subsequently inhibits trophoblast invasion and maternal endothelial and vascular transformation, which result in a similar phenotype that have previously been associated with PE (Figure 1.1b) (Caniggia *et al.*, 2000; Norris, 2005).

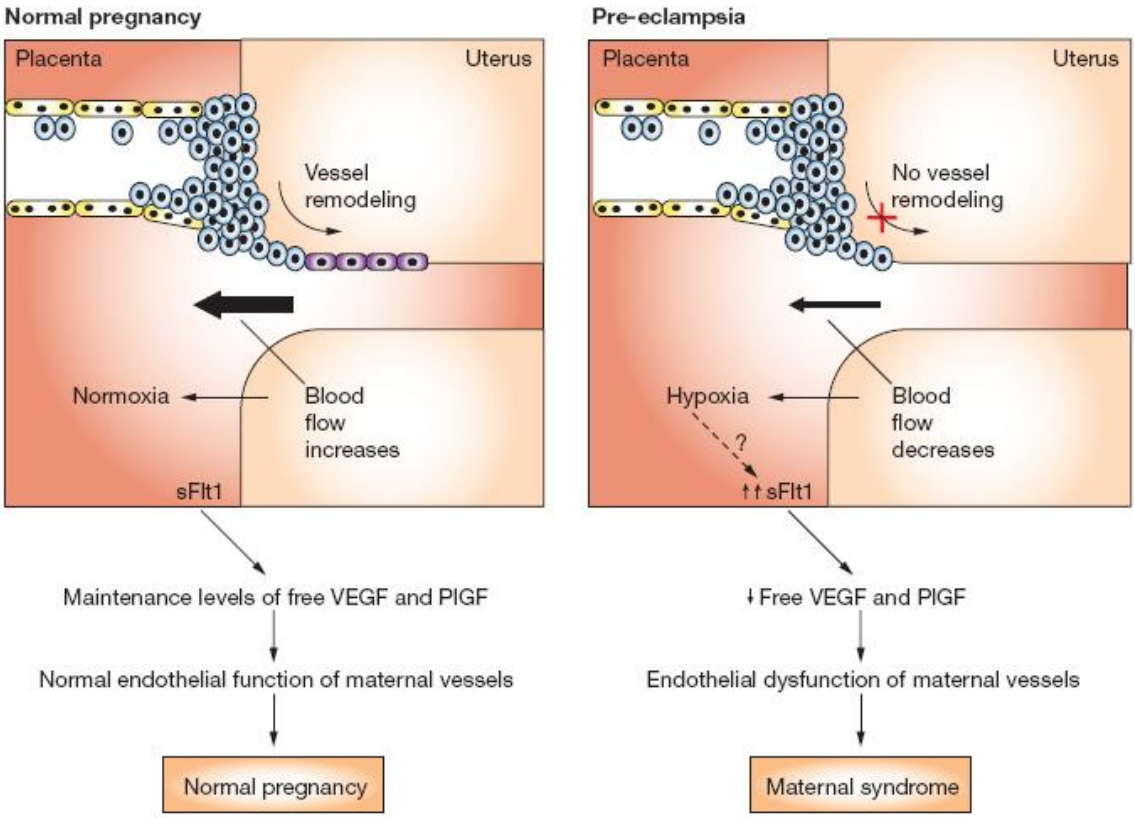


Figure 1.2 Remodeling of the spiral arteries in normal pregnancy versus pre-eclamptic pregnancy (Caniggia *et al.*, 2000).

1.2 PE as a Complex Genetic Disease

Impaired placental implantation and development are associated with the pathophysiology of PE (Roberts and Cooper, 2001), even though the placenta may have normal weight and surface area at birth in a PE patient (Teasedale, 1985). PE occurs only during pregnancy, suggesting that the placenta or fetus may contribute to the disease process (Cross, 2003). Although the etiology, as well as the influence of genetic factors on PE are still unknown, several associations between genetic variants in genes such as the Glu298Asp polymorphism in the endothelial nitric oxide synthase (*eNOS*) gene and the Leiden mutation in the *Factor V* gene and PE have been identified (Dizon-Townson *et al.*, 1996; Serrano *et al.*, 2004; Hillermann *et al.*, 2005). It was found that these variants could possibly affect the function of the genes and consequently be implicated in the pathophysiology of PE. Mutations in the methylenetetrahydrofolate reductase (*MTHFR*) gene have

been associated with PE in several populations (Grandone *et al*., 1997; Sohda *et al*., 1997) but these associations could not be confirmed by subsequent studies (Chikosi *et al*., 1999; Kaiser *et al*., 2000; Rajkovic *et al*., 2000). Since the SNP T235 in the coding region of the angiotensinogen (*AGT*) gene was first identified and found to be associated with PE (Ward *et al*., 1993), it was shown that the T235 SNP was indeed associated with the abnormal development of spiral arteries, a physiological abnormality known to be associated with PE (Morgan *et al*., 1999). These above-mentioned genes and variants implicated in and associated with the pathophysiology of PE are only a glimpse on the complexity of this devastating genetic disease.

1.3 Clinical Symptoms of PE

According to the ISSHP (International Society for the Study of Hypertension in Pregnancy) PE can be identified by hypertension associated with blood pressure levels of 140/90 mmHg or more, measured on two occasions, at least 4 hours apart (Davey and MacGillivray, 1988). These elevated blood pressure levels along with significant proteinuria is diagnostic of PE after 20 weeks of gestation in a previously normotensive, non-proteinuric woman. Headaches, blurred vision, hyperreflexia, epigastric pain and edema are all clinical symptoms or signs of PE. PE could lead to eclampsia and possibly death, if not managed correctly (Cross, 2003).

1.4 Complications – Fetal and Maternal Implications

Universally PE affects five to six percent of pregnancies. In the USA, PE causes approximately 18% of maternal deaths (Chafetz *et al.,* 2007), in contrast to 42% of maternal deaths caused by PE in developing countries such as India and Columbia (Villar *et al.*, 2004). In the Western Cape province of South Africa (SA), at least five percent of pregnancies are affected by PE (Hall *et al.*, 2006; Kenneth *et al.*, 2008) although nationally the figure may be much higher due to pregnancies being managed without appropriate health care support.

There are currently several programs and approaches in place to successfully manage PE in hospitals, but despite these attempts, SA is still losing many mothers to this devastating disease. The NCCEMD (National Committee on Confidential Enquiries into Maternal Deaths in SA) report identified the primary reasons for these deaths in South Africa. The most recent Saving Mothers Report (2005 – 2007), released by the Department of Health, listed Non-Pregnancy Related Infections, which include infections such as pneumonia, HIV-AIDS and Tuberculosis as the primary cause of indirect maternal deaths between 1999 and 2007 (Figure 1.3) (Department of Health, www.doh.gov.za). Hypertension, accountable for 15.7% of maternal deaths during the period from 2005 to 2007 includes all cases with Chronic Hypertension, Pregnancy Induced Hypertension (PIH) and PE.

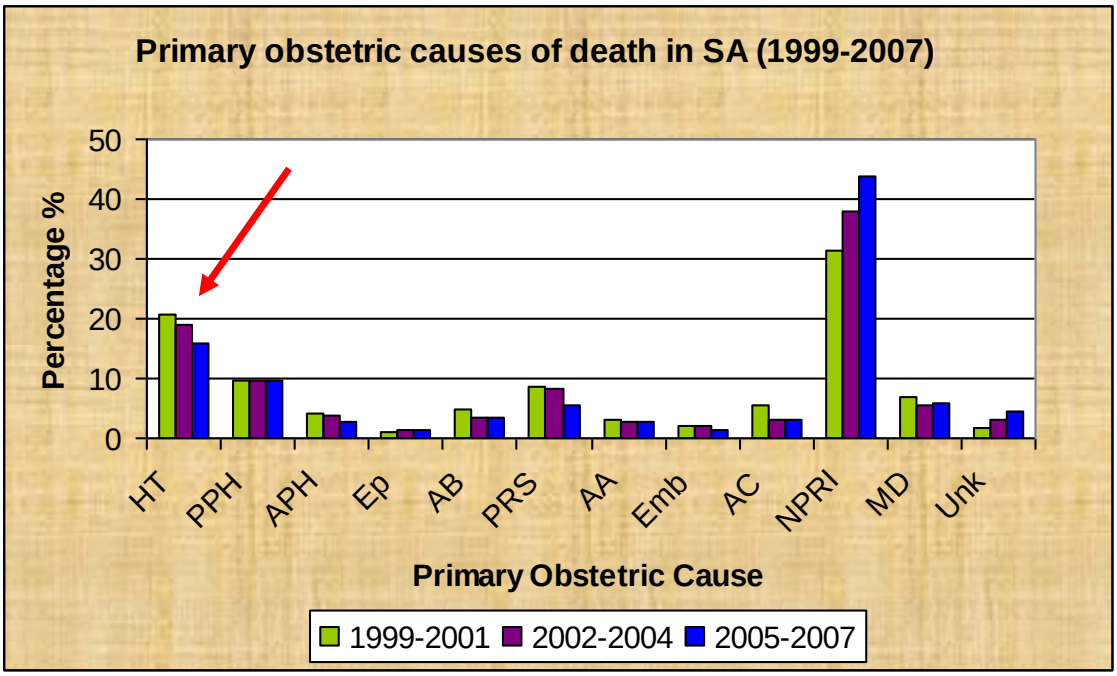


Figure 1.3 Trend of primary obstetric causes of maternal deaths in SA (1999 – 2007) (Saving Mothers Report, Department of Health, www.doh.gov.za) (Abbreviations: AA, Anaesthetic Related; AB, Abortion; AC, Acute Collapse; Emb, Embolism; APH, Antepartum Haemorrhage; Ep, Ectopic Pregnancy; HT, Hypertension; NPRI, Non-Pregnancy Related Infections; MD, Pre-existing Medical Disease; PPH, Postpartum Haemorrhage; PRS, Pregnancy Related Sepsis; Unk , Unknown).

1.5 Risk Factors for Developing PE

i) Maternal, Paternal and Fetal Related Risk Factors

Maternal risk factors for developing PE include aspects such as a family history of PE, previous pregnancy with PE, and maternal age above 40 years (Figure 1.4) (Chappell and Morgan, 2006; Serrano, 2006). PE is known as a "disease of first pregnancy" and a previous normal pregnancy is thus suggested as a protective factor against the development of PE. This, however, is usually only true in the case of long-term sperm exposure to the same partner and if the partner has not previously fathered a PE pregnancy (Tubbergen *et al.*, 1999; Verwoerd *et al.*, 2002; Skjearven *et al*., 2005). Impaired trophoblast invasion or multiple pregnancies (twins, triplets) could lead to the insufficient transformation of maternal spiral arteries, which would subsequently affect placental perfusion, increasing the risk to develop PE (Redman and Sargent, 2005).

i) Risk caused by the Presence of Underlying Disease

Risk factors that may play a role in predisposing women to PE can include underlying complex disorders such as obesity, diabetes, chronic hypertension, renal and autoimmune diseases (Figure 1.4) (Chappel and Morgan, 2006; Serrano, 2006). Obesity and diabetes are complex disorders that are increasingly becoming more prevalent across the world. Since these above-mentioned diseases have a strong link with each other and are strongly associated with insulin resistance, a study by Dekker and Sibai (2001) suggested that a decrease in obesity could subsequently result in a decrease in the frequency of PE.

Hypertension during pregnancy has previously been associated with the future development of cardiovascular disease (CVD) (Mann *et al*., 1976). A study by Jonsdottir *et al*. (1995) found that PE was associated with an increased risk to CVD, since the underlying disease mechanism of PE shares many risk factors with CVD and not hypertension as such. It is therefore

suggested that the factors involved in the disease mechanisms of PE and CVD could be shared (Rodie *et al*., 2004). Although the risk of young women to develop CVD is low, screening for common genotypic and phenotypic factors in these two life threatening diseases could allow for early intervention and therapeutic appraoches for prevention and managing of CVD.

ii) External Related Risk Factors

Cigarette smoking has previously been associated with a reduced risk for the development or incidence of PE (Sibai *et al.*, 1995; Conde-Agudelo and Belizan, 2000). Interestingly, in the study by Sibai *et al.* (1995) women who stopped smoking just before pregnancy had a lower risk of developing PE than women who kept smoking throughout pregnancy. Compared to the harmful effect of smoking on fetal development, birth weight, placental abruption and general health, this advantage seems minor.

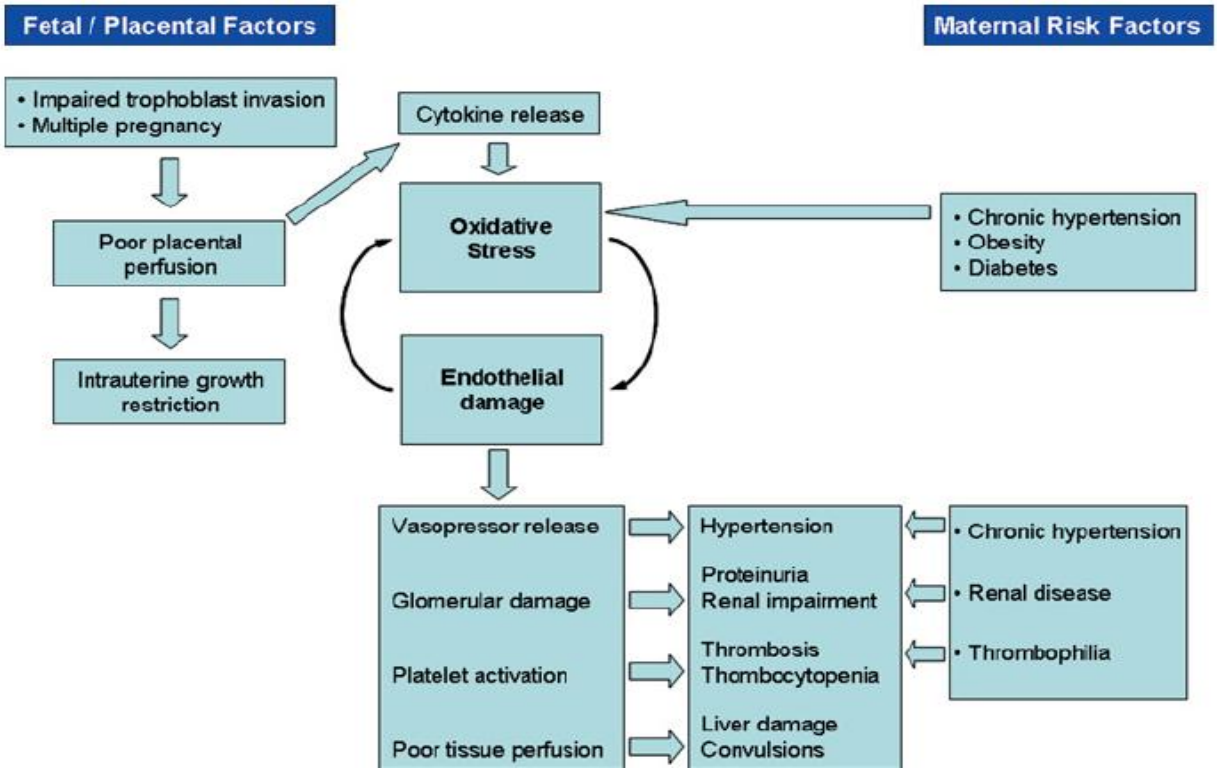


Figure 1.4 A schematic diagram indicating maternal and fetal/placental risk factors for PE (Chappell and Morgan, 2006).

iii) Genetic Risk Factors

Women at-risk of developing PE need to be identified as early as possible, as the primary objective is to facilitate preventative approaches and programmes to prevent the development of PE. A genetic biomarker to assist in the early identification of women at risk for PE would allow for closer observation and immediate medical care (Carty *et al.*, 2008).

1.6 Treatment of - and Prophylaxis to At-Risk Individuals

i) Vitamin C and Vitamin E

Oxidative stress and endothelial damage are implicated in abnormal pregnancies such as PE. The intake of certain antioxidants, such as vitamin C (ascorbic acid) and vitamin E (alpha-tocopheral) are known to inhibit the performance of reactive oxygen species and lipid peroxidation, theoretically resulting in a reduced risk of PE (Chappell *et al.*, 1999). More recently, a study by Poston *et al.* (2006) contradicted this hypothesis, by not finding an association between the intake of vitamin C and vitamin E and a reduction in risk of PE.

ii) Low-Dose Aspirin

Over the years, several studies indicated that low-dose aspirin (60 – 100 mg a day) decreased the risk of PE (Wallenburg *et al.*, 1986; Sibai *et al.*, 1993). Low dose aspirin may favour a prostacyclin preference in the thromboxane A2 – prostacyclin correlation, as thromboxane is responsible for physiological abnormalities such as vasospasms and coagulant problems associated with PE (Sibai *et al.*, 1993).

ii) Magnesium Sulphate

The chemical compound magnesium sulphate has been used as a prophylaxis in women with acute PE in attempt to decrease their risk of

developing eclampsia (convulsions) (Hall *et al*., 2000). Eclampsia is a serious complication of PE with generalized convulsions that can lead to intracranial bleeding if not controlled. A more recent study by Sibai (2005) confirmed that magnesium sulphate has been associated with a reduced risk for developing eclampsia when administered in the acute phase.

1.7 Biomarkers for PE

Biomarkers have previously been defined as "biological characteristics that can be objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes or pharmacological responses to a therapeutic intervention" (Naylor, 2003). Over the years PE has been studied extensively with the primary aim of identifying possible risk factors and biomarkers to timeously identify women at-risk of this devastating condition. Recognising clinical, genetic, proteomic and metabolic risk factors in an attempt to successfully identify women at risk of PE is important for the survival and wellbeing of the unborn fetus and the mother.

1.7.1 PP13 as Maternal Biomarker

Pregnancy-related proteins, such as PP13 are only expressed during pregnancy (Than *et al.*, 2004). PP13 levels measured in maternal serum in pre-eclamptic pregnancies were found to be lower in the first trimester and higher in the second and third trimesters, when compared to PP13 levels in normal pregnancies (Figure 1.5) (Burger *et al.*, 2004). Altered levels of PP13 in PE patients could therefore be an indicator or cause of impaired placental function (Burger *et al.*, 2004).

Serum PP13 was isolated in 1983 by Bohn and his colleagues (Bohn *et al*., 1983) and has since been fully characterised by DTL (Diagnostic Technologies Ltd.). PP13 is a member of the galectin superfamily which is known for roles in cellular differentiation and proliferation (Perillo *et al.*, 1998). PP13 is encoded by the *LGALS13* gene (NM_ 013268) which is located on

the long arm of chromosome 19 (19q13. 1) and consists of two equal 16 kDa subunits that are held together by disulphide bonds.

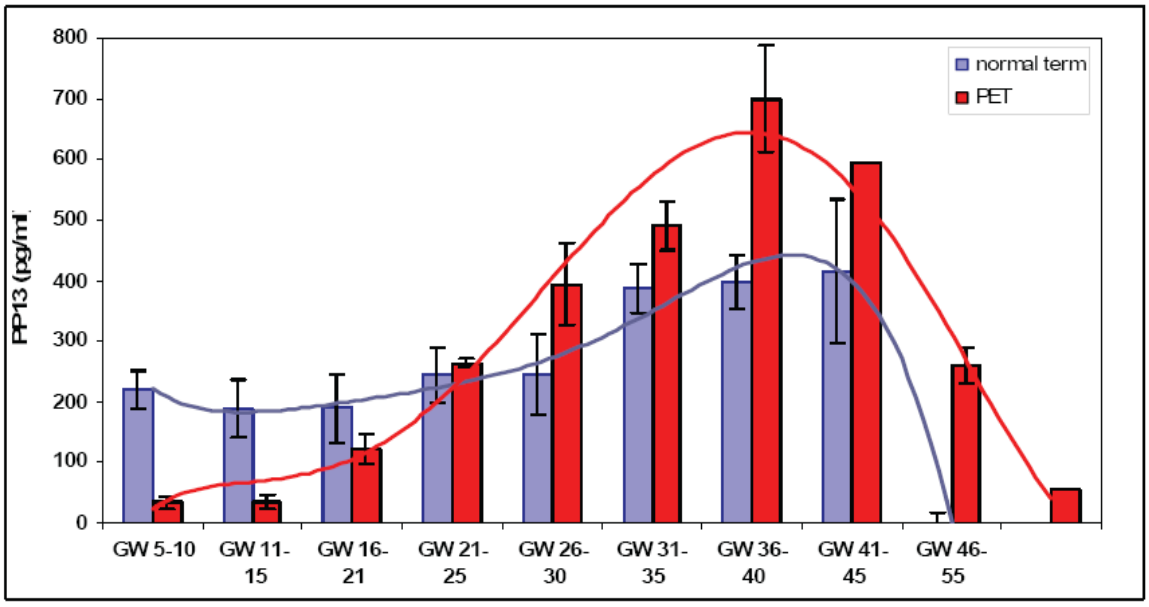


Figure 1.5 PP13 levels measured during normal and pre-eclamptic pregnancies. Abbreviations: GW, gestational week, PET, pre-ecalmpsia. (Image: Dr M Sammar, workshop presentation at the IFPA conference, 2008).

Placental Protein 13 (PP13), as a disease biomarker, could play a pivotal role in early detection of various pregnancy complications due to placental insufficiency (Chafetz *et al.*, 2007; Spencer *et al.*, 2007). PP13 levels can be measured relatively easily, in comparison to Doppler Ultrasonography of the uterine arteries, which is restricted to specialised centres. PP13 levels are measured using ELISA which can be performed during the first trimester of pregnancy (Burger *et al.*, 2003).

Doppler Ultrasonography is routinely used to determine the velocity of uterine-artery blood flow from the 20th to 24th week of pregnancy when results may be provided too late for possible intervention (Nicolaides *et al.*, 2006). In addition, the sensitivity for predicting PE is only 20%-60% and the false positive predictive rate 6% to 40% using this screening technique alone during the first trimester (Yu *et al*., 2003; Nicolaides *et al.*, 2006). PP13 measurement in maternal body fluid, such as maternal serum, would therefore

be a more accurate and reliable screening tool when used in combination with Doppler Ultrasound since a sensitivity of 90% has been reported (Nicolaides *et al.*, 2006). PP13 as biomarker would allow for early intervention and treatment of PE and monitoring of high-risk pregnancies (Burger *et al.*, 2004).

1.7.2 Reasons for Altered PP13 Levels

Reasons for the altered PP13 levels observed in pre-eclamptic pregnancies are currently unknown. DNA mutations in the *LGALS13* gene may have an effect at the level of transcription, influencing the mRNA transcript. Errors during mRNA processing or transport, could affect the stability of RNA, subsequently affecting translation. Alterations at post-translational modification eg, glycosylation and protein folding, could also lead to the production of a defective protein (Lorkowski and Cullen, 2003; Burger *et al.*, 2004). Furtermore, disabled transport of perfectly formed PP13 across the fetal-maternal interface into maternal circulation, could also contribute to decreased/altered PP13 levels. However, to date little is known about the transport of PP13 to the extracellular syncytiotrophoblast apical membrane (Than *et al.*, 2004).

1.8 Galectin Family

Galectins are soluble cytoplasmic proteins that contain either one or two conserved carbohydrate recognition domains (CRDs) that consist of 135 amino acids (Lobsanov *et al.*, 1997; Than *et al.*, 2004). The mammalian galectin family consists of 13 galectins (galectins 1 to 10 and 12 to 14) and 5 proteins similar to galectins namely GRIFIN, HSPC159, PP13, PPL13 AND OvGal11. Members of the galectin family can be subdivided into three classes; prototype, chimera and tandem repeat, according to their individual structures (Figure 6) (Barondes *et al.*, 1994a). Prototypes (galectins 1, 2, 5, 7, 8, 10, 11, 13, 14) represent galectins that contain either one (monomers) or two (dimers) CRDs in a single peptide. Chimeras (galectin 3) are rare galectins, which contain a large CRD, consisting of multiple repeats on a domain that exert numerous functions (Bidon *et al.*, 2001). A small peptide

separates two CRDs on a single protein chain to form a tandem repeat (galectins 4, 6, 8, 9, 12).

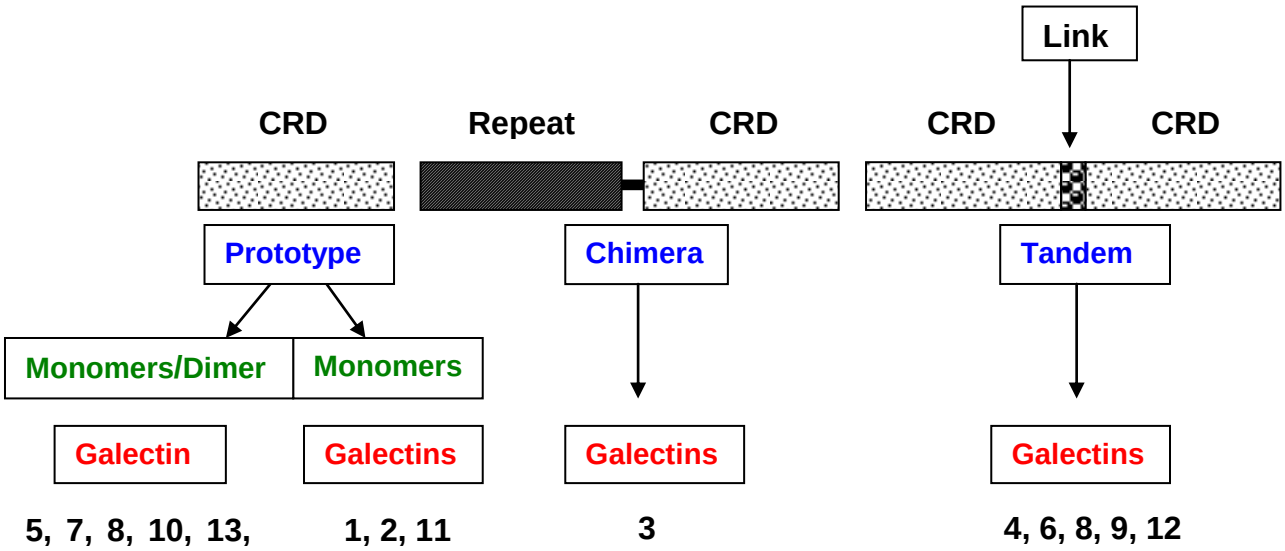


Figure 1.6 Galectin sub-classes (Adapted from: Hughes, 1999).

1.8.1 PP13 as a Member of the Galectin Family

In a study by Than *et al.* (1999), a CRD identified in PP13 was shown to be identical to the structure of other members of the galectin family. PP13 has demonstrated conservation to other members of the galectin family on various levels, such as DNA sequence, gene structure and general functioning (Barondes *et al.*, 1994a). A three dimensional (3-D) model of PP13, constructed to substantiate homology to the galectin family, confirmed PP13 as new member of the galectin family, and was named galectin-13 (Figure 1.7) (Visegrady *et al.*, 2001).

Galectins show an affinity for β-galactosides (Barondes *et al.*, 1994b), but have various additional functions, depending on their location and structure. Most galectins have previously been implicated in cell - cell interactions, cell adhesion and cell signaling (Hughes, 2001). In addition, members of the galectin family are well characterised in the roles they play in inflammation and immune responses. Therefore, PP13/galectin 13 was assigned similar

functions and was thought to play an imperative role in the exchange of blood between the mother and fetus (Than *et al.*, 2004).

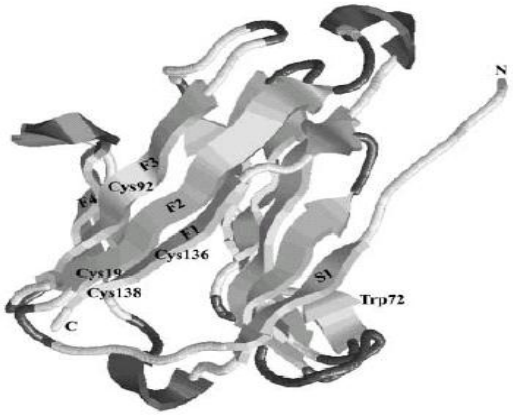


Figure 1.7 Proposed PP13/galectin-13 complex structure Than *et al*. (2004). RASMOL identified cysteine residues on beta-sheets F1 (Cys136, Cys138), F2 (Cys19) and F3 (Cys92), along with a highly conserved tryptophan (Trp72) residue on beta-sheet S6a. Beta-sheet F4 and C- and N-termini are indicated (Than *et al*., 2004).

Individual galectins are expressed in various healthy and disease-affected tissues and are differentially regulated. A study by Than *et al.* (2004) confirmed that PP13/galectin-13 was not only strongly expressed in the placenta, but also at markedly lower levels in healthy spleen, kidney and bladder tissues along with liver adenocarcinoma, neurogen tumour and malignant melanoma (Figure 1.8) (Visegrady *et al.*, 2001).

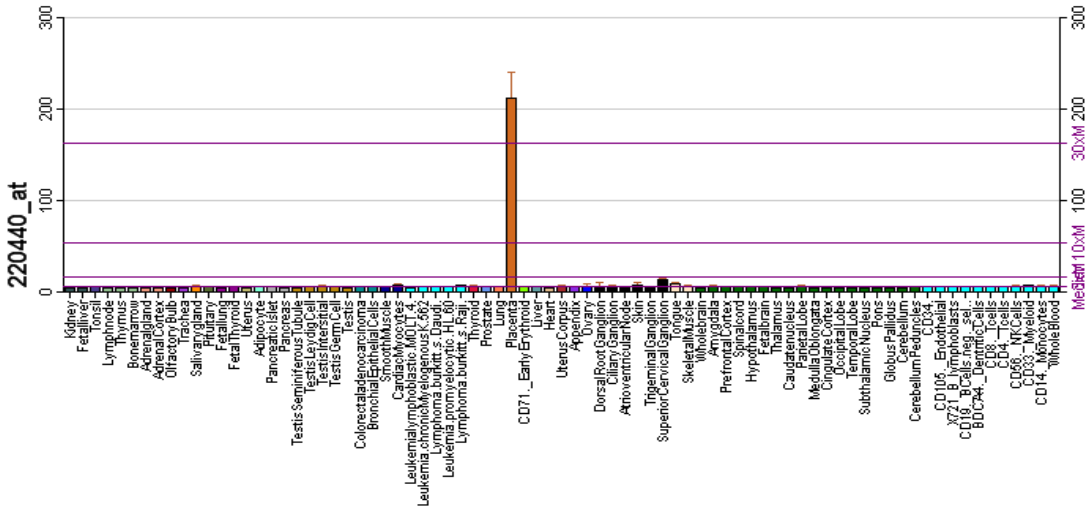


Figure 1.8 Expression profile of the *LGALS13* gene in various human tissues. (Retrieved from: GNF SymAtlas, **http://biogps.gnf.org**).

1.8.2 Secretion of Galectins

When galectins are bound to cell surface receptors they exert essential extracellular functions (Hughes, 1999). Their secretion is regulated by several factors, including cytokines and membrane fusion events, as they move from cells either in free form or contained within vesicles. Cytosolic proteins, clustered below the plasma membrane in the cytoplasm, are excreted as extracellular vesicles that protrude from the plasma membrane as “blebs” (Figure 1.9) (Hughes, 1999). Some galectins (galectins 1and 3) are secreted via ectocytosis from fibroblasts within microvesicles, in conjunction with high amounts of actin and annexin II (Than *et al.,* 2004).

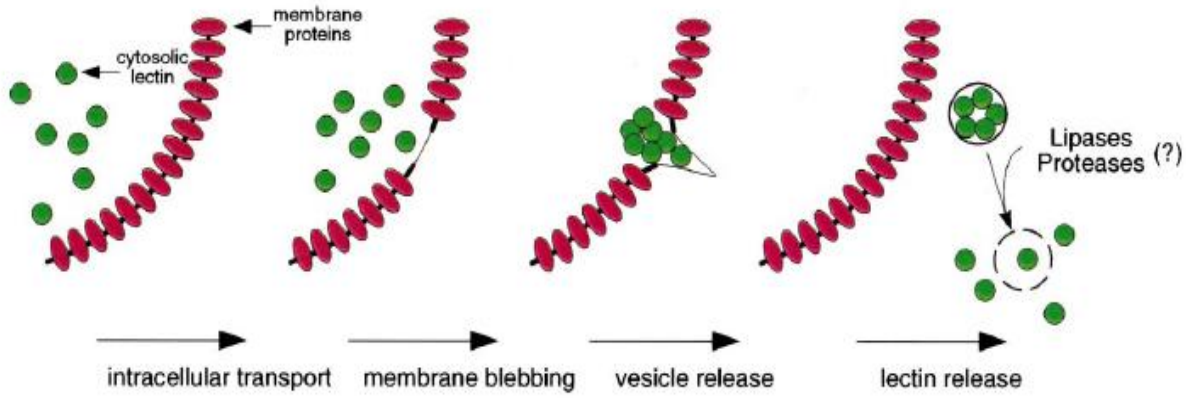


Figure 1.9 Secretion of galectins eg, galectin 3 (Hughes, 1999).

1.8.3 Proteins Associated with PP13

In a study conducted by Than *et al.* (2004), proteins PP13-B (purified from human placenta) and PP13-R (recombinant PP13 protein) were immobilised to determine which proteins were bound to PP13. MALDI-TOF MS peptide mapping and sequencing revealed that PP13-B and PP13-R were specifically bound to annexin II and beta/gamma actin in placental and fetal hepatic cells. This was confirmed by co-localisation of PP13/galectin-13 and annexin II (as A2-S100A10 heterotetramer) on the outer cellular membrane of syncytiotrophoblasts. Staining of placental tissue with anti-annexin II antibodies revealed the presence of annexin II in syncytiotrophoblasts and the brush border membrane (Figure 1.10) (Than *et al.,* 2004).

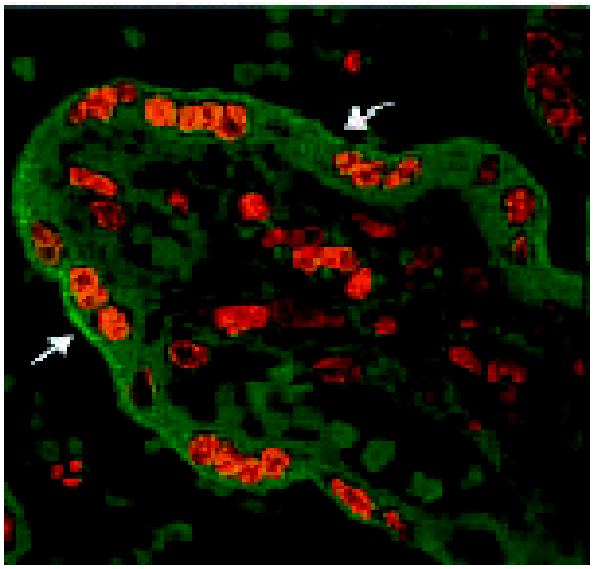


Figure 1.10 Localisation of PP13 with annexin II on the brush border membrane (Than *et al.*, 2004). Green is an indication of annexin II staining in the syncytiotrophoblasts. Arrows indicate more intense staining observed in the brush border membrane.

1.9 Annexin Gene Family

In Greek, annexin refers to the word *annex* that means to “bring or hold together” (Gerke and Moss, 2002). The annexin multigene family is classified according to a conserved region, consisting of a 70 amino acid repeat, of which a sequence of 17 amino acids has been found to be particularly conserved. This area occurs four times within the various annexin genes (Spano *et al*., 1990). The conserved **core** of the annexin sequence consists of 310 residues that contain type-2 motifs for the binding of Ca^{2+} ions (Aukrust *et al*., 2006). While each annexin has particular properties because of its individual protease-sensitive N-terminal tails of variable length (Spano *et al*., 1990), collectively the conserved domains, and the distinctive properties of each annexin most likely play a vital role in their general functioning (Moss and Morgan, 2004).

The annexin family thus consists of globular proteins that are structurally related. In the presence of Ca^{2+}, annexins associate with amniotic phospholipids (aPL) (Aukrust *et al*., 2006). Annexins can be categorized into

three groups according to their structure; ANXA7, ANXA11 and ANXA13 are the earliest vertebrate annexins, ANXA4, ANXA5 and ANXA8 form the second group and ANXA1, ANXA2 and ANXA3 are part of the third group (Gerke and Moss, 2002). Annexins ANXA9 and ANXA10 are seen as distant-related members of the annexin family and ANXA6 is difficult to assign because of its similarity to the annexins in both the second and third groups.

1.9.1 *Annexin V* Gene

Annexins interact with each other in such a way, that one annexin regulates the function of another. A direct interaction between annexin II and annexin V was observed in a study by Brooks *et al*. (2002), suggesting that annexin II may possibly bind directly to annexin V.

Genetic defects often associated with recurrent pregnancy loss (RPL, another association with poor placentation) are the Factor V Leiden mutation and the Factor II 20210 G/A mutation, along with the presence of maternal antiphospholipid antibodies (aPL) (Empson *et al.*, 2002). Reduced levels of the placental anticoagulant protein, annexin V, expressed on the placental trophoblasts, have not only been documented in the presence of aPL (Rand *et al.*, 1994) but have been implicated in the pathophysiology of PE (Shu *et al.*, 2000). It was found that DNA sequence variants in the *ANXA5* gene possibly influenced annexin V expression (Bogdanova *et al.*, 2007). *ANXA5* is localised to chromosome 4q27, has a genomic span of approximately 9 kb and consists of 12 exons (Rand and Wu, 1999).

In the study by Bogdanova *et al.* (2007) the over-expression of a haplotype (-19 G/A, 1 A/C, 27 T/C, 76 G/A), named M2, located within the *ANXA5* promoter, was associated with RPL. The patients in the study cohort were pre-screened for thrombophilic genetic factors to ensure the reason for the reduced *ANXA5* levels most likely caused by the seqeunce variants identified in the promoter region (Bogdanova *et al.*, 2007). Functional reporter gene assays in heterozygous individuals revealed a significant decrease of 37% -

42% in *ANXA5* expression, compared to normal expression levels in wild-type individuals. Both *ANXA2* and *ANXA5* have been implicated in pregnancy-related disorders such as PE, which would suggest that studying the promoter area of the *ANXA2* gene parallel to the study by Bogdavona *et al.* (2007) could potentially reveal a similar haplotype in the *ANXA2* promoter area.

1.10 The *Annexin II* Gene

Human annexin II was first isolated from the placenta as a phospholipase A2 inhibitor, known as lipocortin II (Huang *et al*., 1986). *Annexin II* is therefore part of the human lipocortin-2-encoding multigene family.

The *Annexin II* (*ANXA2*) gene (NM_001002858) is located on the long arm of chromosome 15 (15q21) (Spano *et al.,* 1990), while its three intronless pseudogenes *ANX2P1*, *ANX2P2* and *ANX2P3* (NR_001562; NR_003573.1; NR_001446) are located on chromosomes 4q21-q31, 9p13 and 10q21-q22, respectively (Huebner *et al.*, 1988; Spano *et al.*, 1990).

The *ANXA2* gene is represented by four alternative mRNA transcripts. These transcripts differ in **length** (ie, the length of the N-terminus is shorter in transcripts 2, 3 and 4 as opposed to transcript 1) and **structure** (ie, the position of the translational start site is downstream in isoforms 2, 3 and 4 as opposed to that of transcript 1), but three of the four transcripts encode the same isoform 2. Figure 1.11 represents a comparison of the four different annexin II transcripts as retrieved from the Entrez Gene database (http://www.ncbi.nlm.nih.gov/gene/, 1 September 2009).

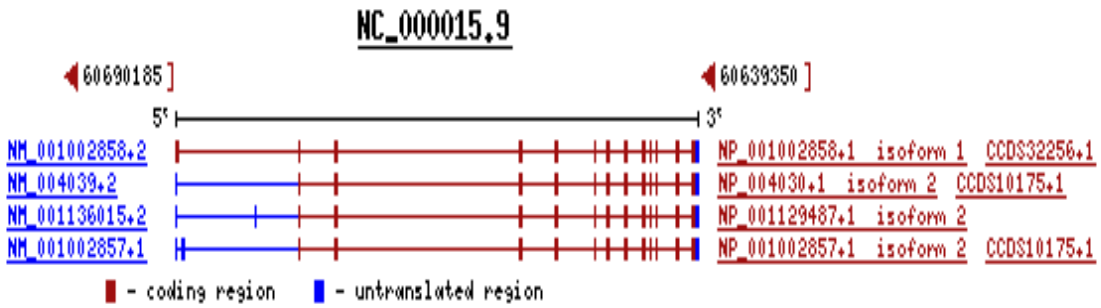


Figure 1.11 Comparison between the four transcript variants for the *ANXA2* gene. (Retrieved from: Entrez Gene database, http://www.ncbi.nlm.nih.gov/gene/)

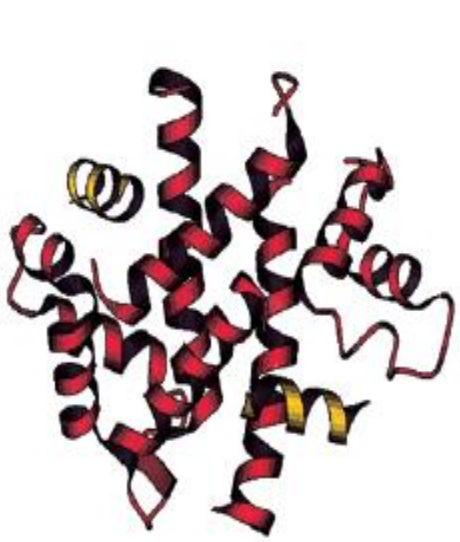
The gene structure for *ANXA2* has been documented. *ANXA2* consists of 13 exons and has a genomic span of 40 kb. In the 5' untranslated region of the *ANXA2* gene, a canonical TATA box, three binding sites for transcription factor Sp1 and a consensus sequence for transcription factor AP-1 have been reported (Spano *et al.*, 1990).

1.10.1 Association of S100A10 with Annexin II (A2-S100A10 Heterotetramer)

Interactions between S100A10 subunits (also known as p11 subunits) which bind to the N-terminal region of annexin II (also known as p36 molecule), result in the formation of the annexin A2-S100A10 heterotetramer (AIIt) (Kaczan-Burgouis *et al.*, 1996; Gerke and Moss, 1997). The S100A10 S100 Ca-binding protein A10, encoded by the *S100A10* gene, is located in the cytoplasm and nucleus of various cells. The S100 protein family consists of at least 13 members, all localised to chromosome 1q21, of which S100A10 is thought to play a role in exocytosis and endocytosis. Amino acid replacements in the Ca^{2+} binding loops result in inactive Ca^{2+} binding sites on S100A10 causing this protein to be in a permanently active state (Rety *et al.*, 1999). The A2-S100A10 heterotetramer consists of two 36 kDa annexin II (p36) molecules linked to two 11 kDa S100A10 (p11) protein subunits

(Figure 1.12b) (Gerke and Moss, 2002). Annexin II is therefore essentially present in the cytosol as a monomer (p36) while the A2-S100A10 heterotetramer complex (AIIt) is able to associate with the plasma membrane (Thiel *et al.*, 1993; Waisman, 1995).

A low-resolution crystal structure confirmed that the two annexin II molecules (red) are structured on either side of the two S100A10 subunits (dark and light blue) in the centre (Figure 1.12b) (MacLeod *et al.*, 2003). According to Lewit-Bentley *et al*. (2000), the high resolution crystal structure of the core domain of annexin II illustrated that the one S100 (p11) subunit binds to the N-terminal region (yellow) of annexin II (Figure 1.12 a, b).



a)

b)

Figure 1.12 **(a)** The structure of the complex of p11 with the N terminal region of annexin II (Kraulis *et al.*, 1991). **(b)** The heterotetrameric complex (AIIt) of S100A10 (p11) with annexin II (p36) (Kraulis *et al.*, 1991).

1.10.2 Expression of the Annexin Family

Annexin II levels vary mainly when present in the form of the AS-S100A10 heterotetramer (Kaczan-Bourgois, 1996). Annexins are expressed in various tissues at different stages, with annexins I, II, V and VI being expressed in syncytiotrophoblast epithelium, and annexins II and V localised in intestinal epithelia (Figure 1.13) (Glenney *et al.*, 1987). In a study by Kaczan-Bourgois (1996), annexin II (p36) and S100A10 (p11) levels in the synciotrophoblast apical membrane were found to gradually increase during pregnancy. Since annexin II has been associated with endocytosis and exocytosis, it could possibly be implicated in transport systems in the brush-border membrane vesicles of the placental syncytiotrophoblast (Kaczan-Bourgois *et al*., 1996).

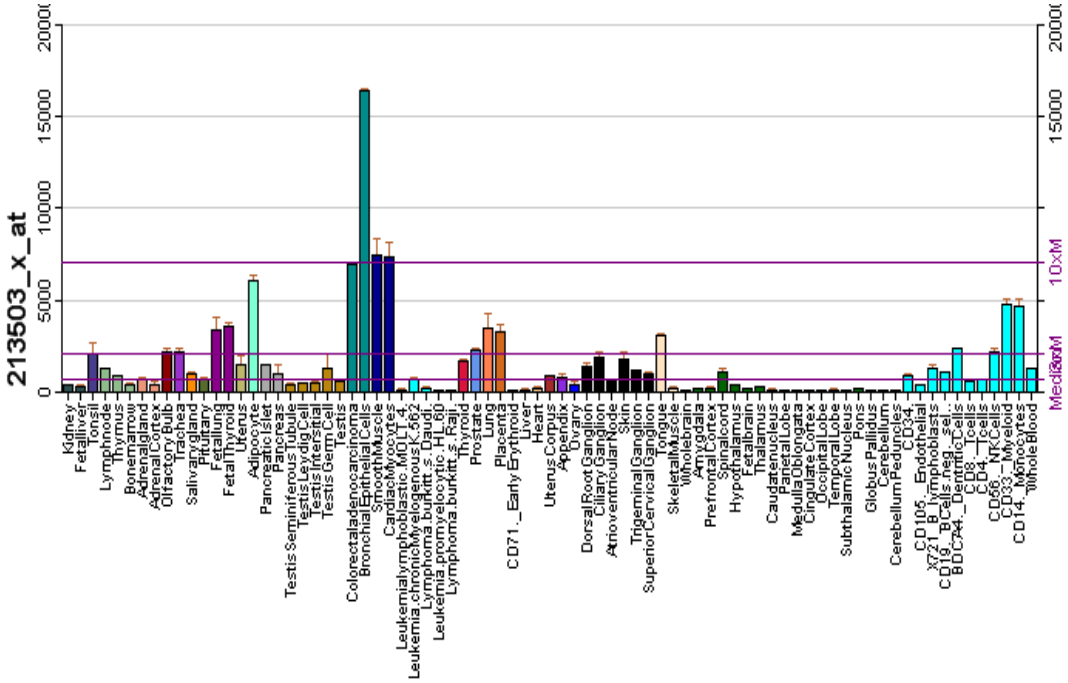


Figure 1.13 Expression profile of the *ANXA2* gene in various human tissues. (Retrieved from: GNF SymAtlas, http://biogps.gnf.org).

1.10.3 Function of Annexin II

Annexin II has been assigned many functions of which “membrane trafficking” is the result of its involvement in *in vitro* interactions and various effects on membranes (Brooks *et al*., 2002). Annexin II is known to play an essential role in a number of biological processes such as the regulation of cell growth and differentiation in the placenta throughout pregnancy, and signal transduction pathways (Creutz, 1992; Baldwin *et al.*, 2005; Aukrust *et al.*, 2006).

Another function of annexin II, receiving growing support in the literature, is its “anti-thrombic role at the endothelial cell surface” (Gerke and Moss, 2002). It is hypothesized that the reduced levels of functional annexin II on the surface of vascular endothelial cells may lead to a predisposition to cardiovascular disease. Additionally, numerous studies have suggested that annexin II plays an important role in fibrin homeostasis. A study by Fleury *et al.* (1993) indicated that human annexin II binds t-PA or plasminogen at the carboxyl-terminal lysines of S100A10, stimulating t-PA-dependent plasmin production. A2-S100A10 therefore functions as a plasminogen regulatory protein, increasing general fibrinolytic activity. A defective annexin II molecule would inhibit catalytic plasmin activity (Cesarman-Maus and Hajjar, 2005), preventing plasmin generation, which would subsequently prevent fibrin degradation, and ultimately would prevent a fibrin hemostatic clot from being dissolved.

1.11 Study Hypothesis

Based on the evidence provided, we hypothesized that DNA sequence variants in the *ANXA2* gene may affect PP13/annexin II binding and in this indirect manner, impact on PP13 transport and facilitate reduced PP13 protein levels observed in pre-eclamptic pregnancies. Currently, little is known about genetic variation in the *ANXA2* gene. This case-control study is therefore a platform to clarify our hypothesis, but also to serve as a foundation for future studies. Results acquired from this research will contribute towards understanding the association between specific genetic variants in the *ANXA2* gene and PE, in addition to contributing to the understanding of the pathogenesis of this complex disease.

1.12 Aim of this Study

The aim of this study was to characterise the sequence variation in the *ANXA2* gene in pre-eclamptic pregnancies.

This would be achieved by the following objectives:

1. searching for sequence variation in the *ANXA2* gene and comparing the frequencies of the identified variants in healthy and pre-eclamptic pregnancies,
2. performing detailed statistical analyses to determine a possible association between the *ANXA2* gene variants and PE,
3. performing *in silico* analysis to determine the predicted effects of the identified variants on annexin II function.

2 MATERIALS AND METHODS

MATERIALS

2.1 Study Cohort

Institutional approval was granted for this study by the Ethical Review Committee of the Faculty of Health Science of Stellenbosch University (N05/07/122), and written informed consent was obtained from all the participants (Appendix A). A complete clinical data set for each patient is available at Tygerberg Hospital.

Maternal and fetal patient and control cohorts were recruited from Tygerberg Hospital in the Western Cape, South Africa. The underlying etiology of PE is unkown and maternal and fetal samples were included in the patient study cohort since genetic variation in the *ANXA2* gene in either maternal or fetal samples could contribute to our understanding of the pathology of PE. The maternal patient cohort (n=120) consisted of 71 Mixed Ancestry and 49 South African Black women who have clinically been diagnosed with early onset PE. Early onset severe PE was diagnosed when the onset of disease was before 34 gestational weeks. Pre-eclampsia was defined by two diastolic blood pressure measurements of 90 mmHg or more, at least 4 hours apart, coupled with significant proteinuria of 0.3 g/24 h or more or persistent 2+ on dipsticks. The smaller sample size for the fetal cohort (n=96) could be explained by poor pregnancy outcome characteristic of PE. These two cohorts had 80 combined maternal-fetal pairs. A sample of approximately 10 ml maternal whole blood was drawn from antecubital fossa veins and up to 5 ml fetal cord blood samples were collected at delivery. Fetal cord blood samples collected from 54 healthy pregnancies (Mixed Ancestry) were used as the control cohort. Whole blood and cord blood samples were collected in blood collection tubes containing the anti-coagulant ehtylenediaminetetraacetic (EDTA) (Becton, Dickson and Company, New Jersey, USA) and stored at -20°C until DNA extraction.

DETAILED EXPERIMENTAL PROCEDURES

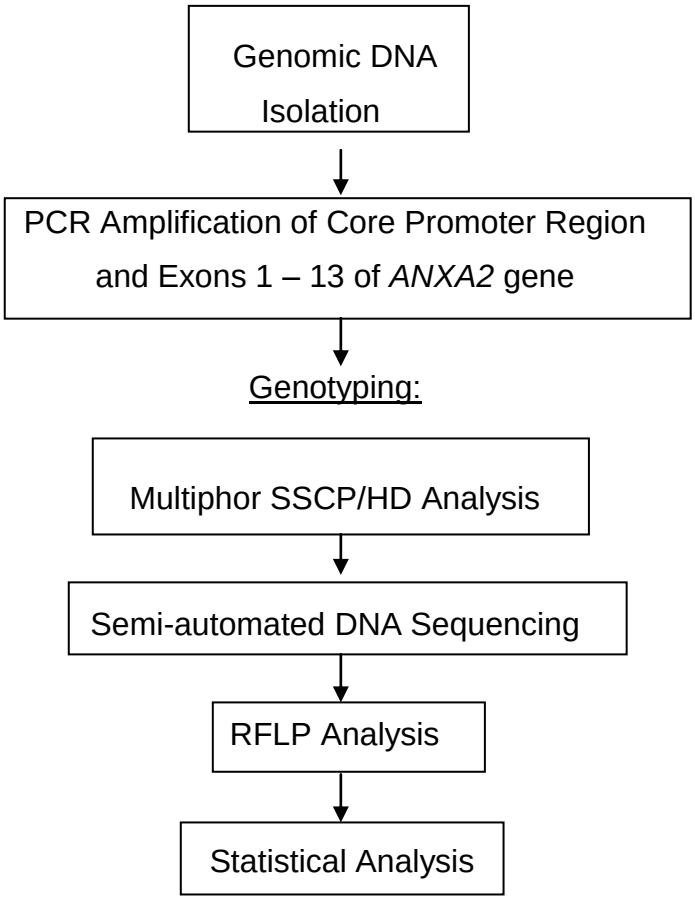


Figure 2.1 Overview of methods that were employed in this study.

2.2 Genomic DNA Isolation from Whole Blood Samples

Genomic DNA was extracted from whole blood from 120 women and 96 offspring diagnosed with PE, as well as 54 healthy control individuals using the Miller *et al*. (1988) protocol (Appendix B). This cohort included 80 pre-eclamptic maternal-fetal pairs. In those fetal samples of low volume, DNA were extracted from 300 µl cord blood using the Pure Gene®DNA isolation Kit (Gentra Systems™, Minneapolis, USA) (Appendix B). All DNA samples were finally diluted in 50 µl TE buffer.

2.2.1 Quantification of Isolated Genomic DNA

Genomic DNA samples were gently shaken for 20 minutes on a vortex prior to quantification, using the NanoDrop® ND-100 Spectrophotometer (Rockland, Delaware, USA). DNA samples were diluted with dH_2O according to their individual concentrations to a final concentration of 30 ng/µl. These dilutions were used for PCR amplification as described in section 2.3.2.

2.3 Polymerase Chain Reaction (PCR)

2.3.1 Primer Design

The *ANXA2* gene reference sequence (accession number: NM_001002858) was obtained from NCBI Entrez Nucleotides Database (http://www.ncbi.nlm.nih.gov/nucleotide/, January 2008). All primers were designed using Integrated DNA Technologies Software, Primer Quest (http://www.idtdna.com) (Figure 2.2). The NCBI Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/) was used to examine primer specificity sequence similarity. PCR amplification of the three intronless pseudogenes reported for the *ANXA2* gene, along with other genes with high sequence similarity to the *ANXA2* gene was prevented by designing primers flanking the coding regions of the gene.

Characteristics of the designed primers are presented in Table 2.2. Primers were also designed to amplify the 5' promoter area of the *ANXA2* gene, while a mutagenic primer was designed for exon 9.

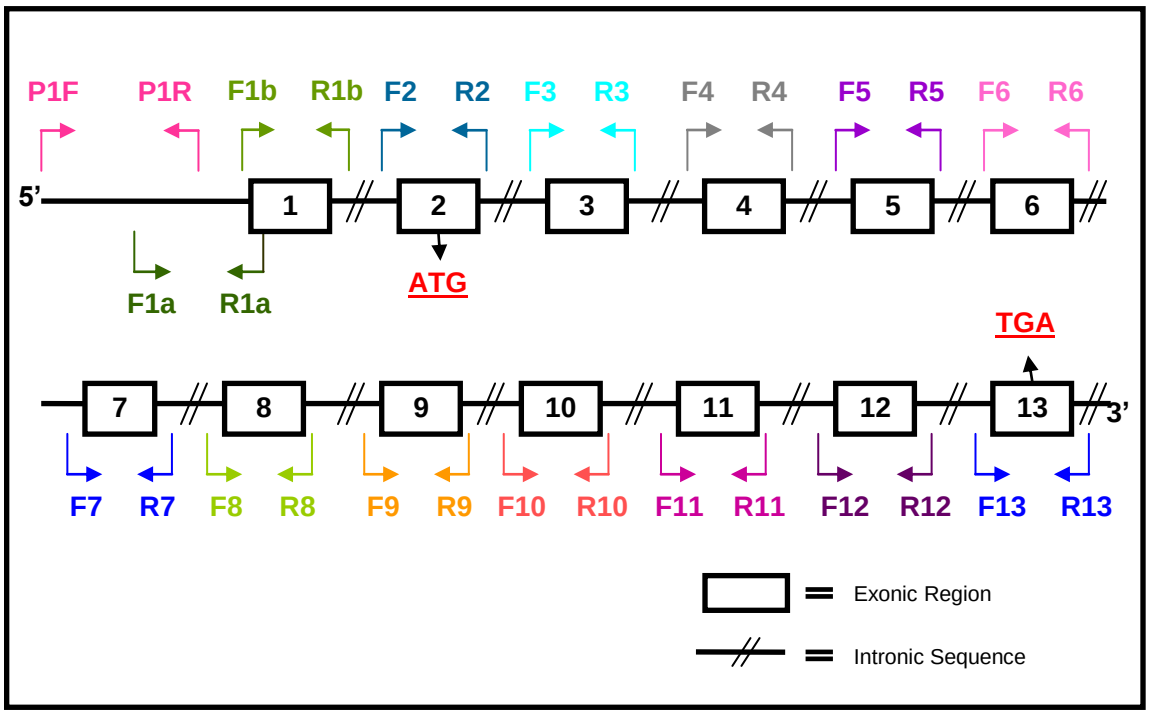


Figure 2.2 Schematic diagram of the *ANXA2* gene indicating primer pairs relative to the exons and introns. Abbreviations: ATG, start codon; F, forward; R, reverse; TGA, stop codon.

2.3.1.1 Primer Design for Core Promoter Region and Exon 1

Since the translation start site (ATG) for the *ANXA2* gene is located within exon 2, exon 1 was considered as part of the promoter. A core promoter region of approximately 600 bp for the *ANXA2* gene was selected and overlapping primers were designed (http://www.idtdna.com). Three primer pairs were designed to amplify fragments of approximately 300 bp, since fragments smaller than 400 bp are most ideal for resolution on the Multiphor SSCP/HD system. Primers F1b and R1b amplified the entire exon 1 and this amplicon thus overlapped with the F1a/R1a amplicon (3 bp) (Figure 2.3).

ggatcctggctcccggctttggcctcccagcccaggctcggggcttcctggcc
ctcggaggccggggcgggccagg**ctttgcagcacagggcgaat**ctccctccct
ccccgggcgccccccgggaactggcggccagaggttaagctgacatgtttgca
gacggtaatggggcgctggatgccaacagccccagctcaggggagtccagatt
tatgtttcttggaacaattctttcctccattgcattaacgccaaaaaaaaaaa
aaaaaaaaaaaaaacacttaatcaagcccagagtctactactaaagtcttgtt
aatttat**taacccaagccgaggctgag**agctcgacgtggcacttaagaagtaa
aatgaggaaagg**aggaagaaacggggtccctgt**gggcctcggggcagcctcgg
ccgggctttctcggagggcagggccaggggcgctggggccgctcccgcgtggc
gcggctcgggagcgttccagggcgcggtccctgcgggccggcgggccggggtg
agtcacccctgacttggggtgggagcgggccgtgtaaggggaggcggggcgg
ggcggggcgggcctcgcctcgcctagggaggatgtggcgggtataaaagcccc
acccagg**c**agccggct**ctGCTCAGCATTTGGGGACGCTCTCAGCTCTCGGCG**
CACGGCCCAGGTAAGCGGGGCGCGCCCTGCCCGCCCGCGATGGGCCGCCAGCT
AGCGGGGTGTGGAGACGCTGGGAAGAAGGgtacggcctggcagggaggccccg
gcaactggccccggaggccgtcgggggcaggggctgaagggacggccggcac
tggagagtctttggggaggggctcagggacaaggtgtttggggcaggggtga
ggggccaccggccctcgggccgggagggtgcagggcagaggtggtgggggtc
gggct**tctctaagaagccctagcgggag**cggggcgaggcggcgtggggagagg
ggcggagggaggccccgcattccagacgcgccaggcgccgccgaggcagggcg

Figure 2.3 PCR amplification of promoter and exon 1 with three primer sets: Fragment P1 primer pair (indicated in green), fragment 1a primer pair (indicated in bold and underlined) and fragment 1b primer pair (indicated in blue). Exon 1 is indicated in upper case.

2.3.1.2 Mutagenic Primer Design for Exon 9

A mutagenic forward primer for exon 9 was designed to include a restriction endonuclease recognition site for digesting with *Tsp509I* to genotype the intronic SNP c.589-5 C>T (Table 2.1) (Figure 2.4).

Table 2.1 Mutagenic primer sequences for the characterisation of SNP c.589-5 C>T.

Primer	Sequence (5'-3')	Annealing Temp (ºC)	Amplicon Size (bp)	Restriction Enzyme	Fragments Sizes (bp)
9MF	gtagttaacttttcctggttggttgtaat	54	209	*Tsp509I*	C allele: 209
9R	gtccatgttgggatggagt				T alelle: 182, 27

atctgtgcctaaggaagactgcttgtgattaagacttttgacctaaatagtag
gtctctgcccatg**tctgcagtataatcatctttcctc**gtttggcattttgaat
gtgtagctttttttgttttttttgcatgacatttttttgacgttacattgga
tgt**gtagttaacttttcctggttggttgtaatc/t**gtagGATCTCTATGACGC
TGGAGTGAAGAGGAAAGGAACTGATGTTCCCAAGTGGATCAGCATCATGACCG
AGCGGAGCGTGCCCCACCTCCAGAAAGgtgggcactgtgctgagggagtctca
gtgccttggggccctggctggggcggctgggtggat**gtccatgttgggatgga**
gtctgccagaatgaggttctgtctccgcttggcac

Figure 2.4 PCR amplification of exon 9 with mutagenic forward primer (indicated in red) and the original reverse primers (indicated in green). The SNP IVS8-5 (C/T) is indicated in yellow highlight. The a-nucleotide highlighted in blue is a c-nucleotide in the original *ANXA2* sequence.

2.3.2 PCR Parameters and Programmes

The standard polymerase chain reaction (PCR) contained 1X GoTaq Buffer (Promega, Madison WI, USA), 30 ng of genomic DNA, 1.5 to 2.5 mM $MgCl_2$ (Promega, Madison WI, USA) (Table 3), 200 µM dNTPs (dATP, dCTP, dGTP, dTTP) (Fermentas International Inc., Hanover, USA), 0.2 to 0.4 µM forward and reverse primers (Table 3), 0.5 to 1.0 U of GoTaq Polymerase enzyme (Promega, Madison WI, USA) (Table 2.3), and the appropriate amount of ddH_2O to reach a final reaction volume of 50 µl. PCR reactions were performed using GeneAmp® PCR Systems 2700 (Applied Biosystems, California, USA). The standard PCR cycle consisted of initial denaturation at 95˚C for 3 minutes, followed by 35 cycles of denaturation at 95˚C for

20 seconds, annealing at the T_a for the specific PCR primers sets (Table 2.2) for 30 seconds, elongation at 72˚C for 30 seconds and a final elongation at 72˚C for 5 minutes.

Dimethyl Sulfoxide (DMSO) (Sigma-Aldrich®, Steinheim, Germany) to a final concentration of 5% was added to the reaction mix of amplicon 1b, containing exon 1 and amplicon P1, containing the majority of the *ANXA2* 5'UTR, in order to successfully optimise the reaction and increase the yield of the PCR product (Table 2.3). Betaine (Sigma-Aldrich®, Steinheim, Germany) (final concentration of 0.5 M) was added to the reaction mix of the amplicon containing exon 6 to optimise and improve its amplification (Table 2.3).

Table 2.2 Oligonucleotide primers used for the amplification of relevant exons in the *ANXA2* gene.

Position	Primer	Sequence (5'-3')	T_m (ºC)	T_a (ºC)	Product Size (bp)
5'UTR	P1F	ctttgcagcacagggcgaat	59	61	328
	P1R	acagggaccccgtttcttcct	60		
5'UTR/ Exon 1	1aF	taacccaagccgaggctgag	60	58	366
	1aR	cgccgagagctgagagcgt	62		
Exon 1	1bF	ctgctcagcatttggggacg	59	62	329
	1bR	cccgctagggcttcttagaga	58		
Exon 2	2F	tcgctgtgcttcagagtaaag	55	54	216
	2R	cttccttagatgtacaaggatag	50		
Exon 3	3F	tacctacctagttacctagtatg	50	56	325
	3R	ccaacgtatggcgatttgag	54		
Exon 4	4F	cttgacaccttctatgtaagcc	53	54	261
	4R	cgcatgagagccacaattca	56		
Exon 5	5F	ggtcaaagactcacaacctgc	56	54	312
	5R	actgaaaatcaccaaattacactcaga	55		
Exon 6	6F	ggatgctaacagacccaatc	56	54	302
	6R	actaatgcaggcatcttagcc	55		
Exon 7	7F	ctgttctgcttctcagagtc	53	55	273
	7R	aggggatttagttaattcactcc	52		
Exon 8	8F	taaataatggcgtacatggactg	53	55	246
	8R	ggagtcatttgtcacacagag	53		
Exon 9	9F	tctgcagtataatcatctttcctc	52	54	305
	9R	actccatcccaacatggac	54		
Exon 10	10F	ctaagctcctaggatgctgtg	55	55	300
	10R	attaacaggatggccatctgtc	55		
Exon 11	11F	ccattgcattaggtcacttctg	54	54	260
	11R	catccatgaatcaaggag	47		
Exon 12	12F	gatgctcaacctgtagacgc	56	56	279
	12R	tcatcattctgccaggccac	58		
Exon 13	13F	tgtgcgttgaggctacttctg	57	54	233
	13R	gggactgttattcgcaagctg	56		

T_m = Melting Temperature ; T_a = Annealing Temperature

Table 2.3 Final primer, *Taq Polymerase*, $MgCl_2$, DMSO and Betaine concentrations in the PCR reaction for the amplification of the *ANXA2* gene.

Amplicon	Primers (µM)	Taq (U)	$MgCl_2$ (mM)	DMSO (%)	Betaine (M)
P1	0.2	0.5	1.5	5	-
1a	0.2	0.5	1.5	5	-
1b	0.4	1.0	1.5	-	-
2	0.4	0.5	1.5	-	-
3	0.4	0.5	2.0	-	-
4	0.4	0.5	2.0	-	-
5	0.4	0.1	2.0	-	-
6	0.4	0.5	2.5	-	0.5
7	0.4	0.5	1.5	-	-
8	0.4	0.5	1.5	-	-
9	0.4	0.5	2.0	-	-
10	0.4	0.5	1.5	-	-
11	0.4	1.0	2.0	-	-
12	0.3	0.5	1.5	-	-
13	0.3	0.5	1.5	-	-

2.3.2.1 Visualisation of PCR Products

Amplification was confirmed by agarose gel electrophoresis on a 1.5% (w/v) agarose gel with 1X TBE buffer (Appendix C) containing 0.003% (v/v) Ethidium Bromide (Sigma-Aldrich Inc., St Louis, USA). Five µl of Cressol (Appendix A) was added to 5 µl of the amplification product and loaded onto the agarose gels. The products were resolved at 120 Volts for 45 minutes in 1X TBE buffer (Appendix C). A 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas International Inc., Hanover, USA) was used concurrently to estimate the size of the PCR products. The gels were examined under UV light and photographed with the Multigenius Bio Imaging System (Syngene, Cambridge, UK).

2.4 Mutation Detection

2.4.1 Multiphor Single Strand Conformational Polymorphism / Heteroduplex (SSCP/HD) Analysis

Amplicons were screened for conformational variation using Multiphor Single Strand Conformational Polymorphism/Heteroduplex (SSCP/HD) analysis (Liechti-Gallati *et al*., 1999) (Appendix B). Three µl of each PCR product was mixed with 3 µl of fomamide-based loading dye (Appendix C) and denatured for 5 minutes at 95°C. The mix was then quenched on ice, allowing the formation of stable homo- and heterodimers. Three µl of this reaction mix was then loaded onto a Multiphor SSCP/HD gel, consisting of 12% (w/v) acrylamide-PDA gel mix (Appendix C), 10% (w/v) ammonium persulphate (APS) and N, N, N'N'-tetramethylethylene-diamine (TEMED). Electrophoresis was performed at optimised temperatures, for optimised duration, at 355 Volts on the Pharmacia LKP 2117 Multiphor^{TM} II Eletrophoresis System (Amersham Pharmacia Biotech, Amsterdam, UK) (Table 2.4).

Following electrophoresis conformational bands on the Multiphor SSCP/HD gels (Appendix B) were visualised via a two-step 0.1% silverstaining method. Firstly the Multiphor SSCP/HD gel plates were stained in a 0.1% (w/v) Silver Nitrate (Appendix C) solution for 10 minutes, after which plates were washed with dH_2O to remove excess solution from the gels. The second step involved staining Multiphor SSCP/HD gels for 10 minutes with Developing Solution (Appendix C). The gels were lifted onto clean Watmann filter paper and DNA fragments could be visualised immediately.

2.4.2 Semi - automated DNA Sequencing

Conformational variants detected on Multiphor SSCP/HD gels were subsequently characterised by bi-directional semi-automated DNA sequencing. Successfully amplified PCR products chosen for semi-automated DNA sequencing were purified using Sure Clean (Bioline. London,

UK) (Appendix B). To ensure successful purification of the PCR products, concentrations were determined using the NanoDrop® ND-100 Spectrophotometer (Rockland, Delaware, USA). The Central Analytical Facility (Stellenbosch University, Stellenbosch, SA) performed all sequencing and electrophoresis reactions by using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, California, USA) followed by electrophoresis on the ABI 3130XL Genetic Analyzer (Applied Biosystems, California, USA).

Sequences were analysed using BioEdit Sequence Alignment Editor Software v7.0.9.0 (Hall, 1999). The sequences were aligned to the subsequent reference sequences (accession number: NM_001002858) obtained from NCBI Entrez Nucleotides Database (http://www.ncbi.nlm.nih.gov/nucleotide/, 2008) using the ClustalW v1.4 program in BioEdit Sequence Alignment Editor Software Program and analysed for differences from the reference sequences.

Table 2.4 Detection methods used to screen the *ANXA2* gene.

Amplicon	Detection Method	Multiphor Temperature (ºC)	Duration
P1	Multiphor SSCP/HD, Semi-automated DNA Sequencing	17	3h
1a	Multiphor SSCP/HD, Semi-automated DNA Sequencing	13	4h
1b	Multiphor SSCP/HD, Semi-automated DNA Sequencing	9	3h45min
2	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	4	3h20min
3	Multiphor SSCP/HD, Semi-automated DNA Sequencing	17	3h30min
4	Multiphor SSCP/HD, Semi-automated DNA Sequencing	13	2h
5	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	4	3h20min
6	Multiphor SSCP/HD	21	3h
7	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	17	3h30min
8	Multiphor SSCP/HD	21	3h15min
9	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	13	3h30min
10	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	5.5	3h
11	Multiphor SSCP/HD, Semi-automated DNA Sequencing	17	3h15min
12	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	17	3h30min
13	Multiphor SSCP/HD, RFLP, Semi-automated DNA Sequencing	13	3h15min

2.4.3 Restriction Fragment Length Polymorphism (RFLP) Analysis

Genotypes were confirmed by Restriction Fragment Length Polymorphism (RFLP) analysis of the PCR products. The nature of the restriction fragments for each amplicon was determined (when the sequence variant abolished or created a restriction endonuclease recognition site) using restrictionmapper.org, http://insilico.ehu.es/restriction and BioEdit Sequence Alignment Editor Software v7.0.9.0 (Hall, 1999). Restriction reactions consisted of 10 µl of the PCR product, 1X relevant restriction enzyme buffer (Table 2.5) and 5 U of the relevant restriction enzymes with added dH_20 to bring the final reaction volume to 20 µl. The incubation of the PCR products with relevant restriction enzymes was performed according to the manufacturers' recommendations for 16 hours (Table 2.5).

Table 2.5 Restriction enzymes, buffers and incubation temperatures used for RFLP analysis.

Exon	Variant	Buffer	Restriction Enzyme	Incubation Temperature	Recognition Sequence (5'-3')
2	c.-11-43 G>A	B	***Mbo* II	37ºC	GAAGA$(N)_8$^
	c.-11-13 A>T	Tango	***Rsa* I	37ºC	GT^AC
9	c.589-5 C>T	1	**Tsp509I*	65ºC	^AATT
	c.682+49 C>T	3	**Aci* I	37ºC	C^CGC
10	c.683-56 G>A	3	**Bcl* I	50ºC	T^GATCA
12	c.934 G>A; p.Gly312Ser	3	**BceA* I	37ºC	ACGGC$(N)_{12}$^
13	c.*1057	4	**Hph* I	37ºC	GGTGA$(N)_8$^
	c.975 C>T; p.Gly325Gly	4	**Hph* I	37ºC	GGTGA$(N)_8$^

Restriction Enzyme suppliers:

* New England Biolabs Inc., Beverly, USA

** Fermentas International Inc., Hanover, USA

2.4.3.1 Visualisation of Restriction Enzyme Digested Products

The products of the restriction enzyme digestions were resolved on 3% (w/v) agarose gels with 1X TBE buffer (Appendix C) containing 0.003% (v/v) Ethidium Bromide (Sigma-Aldrich Inc., St Louis, USA). A 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas International Inc., Hanover, USA) was used to estimate the size of the restriction fragments. The gel electrophoresis was performed at room temperature and at 90 Volts for optimised duration. An ultraviolet light (UV) system (Multi Genius Bio Imaging System, Syngene, Cambridge, UK) was used for visualisation of the DNA fragments (Fragment sizes are shown on the gels in the following section).

2.4.4 Heteroduplex Single-Stranded Conformational Polymorphism (HEX-SSCP) Analysis

Heteroduplex Single-Stranded Conformational Polymorphism (HEX-SSCP) analysis (Kotze *et al.*, 1995) was used to screen and confirm genotyping in selected exons by preparing 12% (w/v) PAA gels (Appendix C) containing 7.5% (w/v) urea, 1.5X TBE (Appendix C), 0.1% (w/v) ammonium persulphate (APS) and 0.1% (v/v) N, N, N'N'-tetramethylethylene-diamine (TEMED).

PAA gels were cast, left to polymerise and the wells were cleaned with 1.5X TBE before the gel systems were assembled and filled with 1.5X TBE. The HEX-SSCP gel system was kept at 4°C for approximately 1 hour before loading the PCR products. PCR products (20 µl) mixed with 15 µl bromophenol blue loading dye (Appendix C) were denaturated at 95°C for 10 minutes and cooled immediately on ice. Fifteen µl of PCR product was loaded onto the PAA gels and eletrophoresis was performed at a constant temperature of 4°C at 250 Volts for 18 hours on the Hoefer SE 600 Basic Slab (30 cm) vertical apparatus.

Following electrophoresis, gels were removed from the gel apparatus and stained in a 0.01% (v/v) Ethidium Bromide (Sigma-Aldrich Inc., St Louis, USA)

solution for 10 minutes and destained in dH_2O for 3 minutes. DNA fragments were visualised and photographed using an ultraviolet light (UV) system (Multi Genius Bio Imaging System, Syngene, Cambridge, UK).

2.5 Statistical Analysis

Allele and genotype frequencies were determined by counting, and tested for deviation from the Hardy-Weinberg Equilibrium (HWE) with Chi-square analysis in Tools For Population Genetic Analysis (TFPGA) Software v1.3 (Miller, 1997). Fischer's exact test was performed when less than five individuals of a certain genotype were observed. Differences between the allele and genotype frequencies were verified in Microsoft® Office Excel (2003) by way of 2X2 and 2X3 contingency tables. In order to determine a possible association between genetic variants in the *ANXA2* gene and PE, an analog to Fischer's exact test and haplotype analysis were performed within Haploview, v3.31 (Barett *et al*., 2005). *P*-values of less than 0.05 were considered significant. A LOD score ≥ 3 and a correlation coefficient (r^2-value) of ≥ 0.4 were considered a significant indication of linkage disequilibrium (LD).

3 RESULTS

3.1 Patient Demographics

A total of 120 pre-eclamptic samples was available for genetic screening of the *ANXA2* gene. Clinical patient demographics for participants are shown in Table 3.1. Clinical patient information was obtained in the form of questionnaires completed by authorised hospital personnel (Appendix D).

Table 3.1 Demographic features for the maternal study cohort.

Patient demographic features	Mean	Range	Normal
Age (years)	22.5	14-43	n/a
Gravidity	1.25	1-5	n/a
Parity	0.2	0-3	n/a
Systolic blood pressure (mmHg)	159.6	140-210	120
Diastolic blood pressure (mmHg)	106.4	80-140	80
Gestational age at complication (weeks)	29	25-34	n/a
Gestational age at delivery (weeks)	30.25	22-38	38-40
Fetal birth weight (grams)	1330	324-2496	Appendix E

mmHg = Millimeter of Mercury

Maternal patient demographics revealed that most mothers were primi-gravidae patients, which was consistent with the mean maternal age observed (22.5 years). As expected, systolic and diastolic blood pressure levels of pre-eclamptic mothers were elevated compared to healthy pregnant mothers. The median gestational age at complication was very early (29 weeks) and correlated with the mean gestational age at delivery (30.25 weeks), since clinical intervention was needed immediately to attempt saving mother and neonate. Early delivery as a consequence of complication will affect the outcome of fetal birth weight directly, which explains the low fetal birth weight (1330 g) observed. From the patient demographics it is evident that most cases were “early-onset” ie, PE was diagnosed after 20 weeks of pregnancy, but before 34 weeks, suggesting that this is a relatively ‘homogenous’ clinical group.

3.2 Screening of the *ANXA2* gene

Conformational variants on the Multiphor SSCP/HD system were identified in ten of the 15 *ANXA2* gene amplicons screened (Table 3.2). No variants were identified in exons 2, 3, 6, 7, 10 and 11 of the *ANXA2* gene. The sequence variants detected in the various regions of the *ANXA2* gene, the PCR fragments they occurred in, their nature and reference number (if previously documented) are indicated in Table 3.2. The nomenclature used to describe the *ANXA2* gene and sequence variants in this study was according to the recommendations by the Human Genome Committee (HUGO), (http://www.genenames.org/guidelines.html) and the Human Genome Variation Society (HGVS), (http://www.hgvs.org/mutnomen/).

Table 3.2 Variants identified in the *ANXA2* gene.

Region	Amplicon	Variant	Nature	Ref#	Reference
5'UTR	**1a**	c.-442 (5'UTR)	C>G	-	**NOVEL**
		c.-191 (5'UTR)	G>C	-	**NOVEL**
		c.-189_-188ins(5'UTR)	GCCGG	-	**NOVEL**
		c.-135 (5'UTR)	C>G	-	**NOVEL**
	1b	c.-92 (5'UTR)	A>T	-	**NOVEL**
		c.-31 (5'UTR)	T>C	rs12904657	NCBI SNP database
Coding	**4**	c.222; p.Ala74Ala	C>T	-	**NOVEL**
	5	c.292; p.Val98Leu	G>T	rs1059688	NCBI SNP database
	9	c.600; p.Asp110Asp	C>T	-	**NOVEL**
	12	c.934; p.Gly312Ser	G>A	-	**NOVEL**
	13	c.975;p.Gly325Gly	C>T	rs1244554	NCBI SNP database
Non-coding	**1b**	c.-12+75	C>A	rs12904756	NCBI SNP database
	2	c.-11-43	G>A	rs11858864	NCBI SNP database
		c.-11-13	A>T	rs11855679	NCBI SNP database
		c.48+67	C>T	rs3743268	NCBI SNP database
	5	c.244-42	G>C	-	**NOVEL**
		c.244-76	C>G	-	**NOVEL**
	7	c.449-17	G>A	rs12898604	NCBI SNP database
		c.528+38	C>T	-	**NOVEL**
	9	c.589-5	C>T	-	**NOVEL**
		c.682+49	C>T	-	**NOVEL**
	10	c.683-56	G>A	rs11633619	NCBI SNP database
	13	c.961-30	A>G	-	**NOVEL**
		c.961-24	C>G	-	**NOVEL**
3'UTR	**13**	c.*1057	A>G	-	**NOVEL**

Ref# = Reference number as retrieved from the Entrez SNP database, (http://www.ncbi.nlm.nih.gov/SNP/).

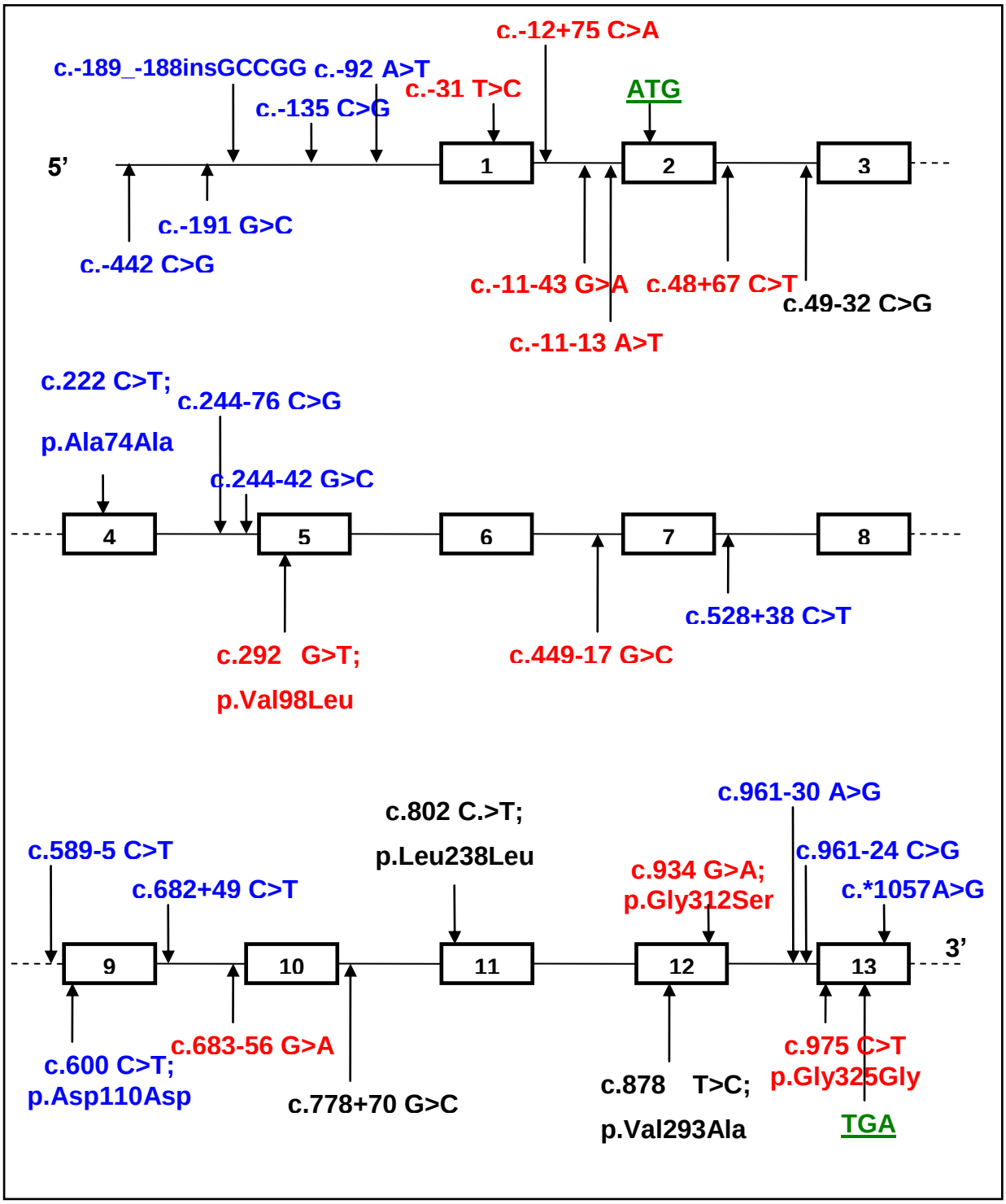


Figure 3.1 Schematic diagram (not drawn to scale) of the *ANXA2* gene indicating positions of previously documented variants identified in this study (indicated in red), novel variants identified in this study (indicated in blue) and previously documented variants, not identified in this study (indicated in black).

Fifteen PCR fragments were used to screen the *ANXA2* gene, including 600 bp upstream into the 5'UTR. Twenty-five sequence variants were identified in this study of which 15 were novel and ten previously documented (Figure 3.1). The variants were grouped according to their location and will be discussed accordingly.

3.2.1 Variants in the *ANXA2* 5'UTR

Five novel (c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG, c.-135 C>G and c.-92 A>T) and one previously reported (c.-31 T>C) variant were identified in the *ANXA2* 5'UTR following Multiphor SSCP/HD analysis.

Multiphor SSCP/HD analysis of amplicon 1a, containing part of the 5'UTR and part of exon 1 (Appendix F, Figure 1a) revealed four novel variants, namely c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG and c.-135 C>G in the *ANXA2* 5'UTR. Sequencing electropherograms for SNPs c.-442 C>G, c.-191 G>C and c.-135 C>G are presented in Appendix F, Figure 1b. All variants were genotyped by Multiphor SSCP/HD analysis, as the conformations were very distinctive.

The DNA sequence representative of the novel 5 bp insertion 188 bp upstream of the translational start site (ATG) located in exon 2 was compared to the wild type sequence in Appendix F, Figure 2 to indicate the position and nature of this insertion. The c.-189_-188insGCCGG insertion revealed a unique conformation in the heteroduplex section of the gel. The heterozygous status of the c.-189_-188insGCCGG insertion was more prevalent in the total maternal cohort (4.6%) compared to the fetal (1.2%) and control (1.9%) cohorts. No individuals homozygous for the insertion were observed.

Multiphor SSCP/HD analysis of amplicon 1b, encompassing exon 1 (Appendix F, Figure 3), revealed one novel SNP c.-92 A>T and one previously documented SNP c.-31 T>C (dbSNP: rs12904657). These two SNPs are

localized upstream of the ATG within exon 2 and were therefore classified as variants located in the *ANXA2* 5'UTR.

No restriction endonuclease recognition sites could be found to confirm SNP c.-31 T>C. Several patient and control individuals were subsequently subjected to HEX-SSCP analysis (Figure 3.2) to confirm genotypes observed on the Multiphor SSCP/HD system, since the conformations were not always clear and easy to distinguish from each other. In Figure 3.2 conformations 2, 5 and 7 as allocated on the Multiphor SSCP/HD gels appeared similar on the HEX-SSCP gels, although the DNA sequences revealed that these are different from each other. Multiphor conformations 1, 4 and 6 appeared identical on the HEX-SSCP gels, but were in actual fact different according to the DNA sequencing results. Genotypes for SNP c.-31 T>C were therefore allocated according to Multiphor typings. In contrast, Multiphor conformations for variant c.-92 A>T were clear and were therefore used to genotype the patient and control cohorts.

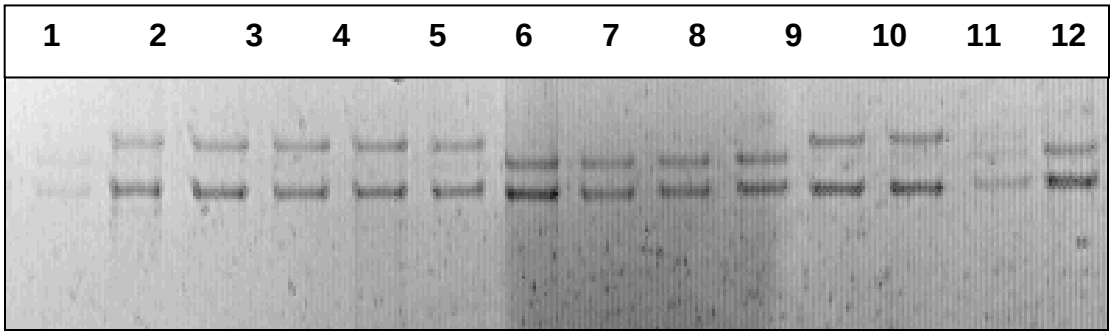


Figure 3.2 HEX-SSCP analysis of specific conformations observed with Multiphor SSCP/HD analysis. Lane 1 = conformation 1; lanes 2-6 = conformations 2 and 5; lanes 7-9 = conformations 3 and 4; lane 10 = conformation 6; lanes 11, 12 = conformation 7 according to Multiphor SSCP/HD analysis.

3.2.2 Coding Variants

Three novel (c.222 C>T; p.Ala74Ala, c.600 C>T; p.Asp110Asp and c.934 G>A; p.Gly312Ser) and two previously reported (c.292 G>T; p.Val98Leu and c.975 C>T; p.Gly325Gly) variants were identified within the coding regions of the *ANXA2* gene following Multiphor SSCP/HD analysis.

Multiphor SSCP/HD analysis of the amplicon containing exon 4, revealed a novel synonymous exonic SNP at amino acid position 74 (c.222 C>T; p.Ala74Ala) (Appendix F, Figure 5a). No restriction enzyme digest was performed to confirm genotypes observed on Multiphor SSCP/HD gels, since all three variant conformations detected, were subjected to semi-automated DNA sequencing (Appendix F, Figure 5b). Three heterozygous C/T control individuals were observed (5.7%), but the variant allele was absent from maternal and fetal patient cohorts.

One previously documented exonic SNP c.292 G>T; p.Val98Leu (dbSNP: rs1059688) was detected following bi-directional DNA sequencing of the conformational variants identified by Multiphor SSCP/HD analysis of the amplicon containing exon 5 (Appendix F, Figures 6a and 6b). The non-synonymous exonic SNP Val98Leu is responsible for a change from Valine to Leucine at amino acid position 98, which are both hydrophobic amino acids with non-polar side chains. None of the patient or control individuals were found to be homozygous for this SNP.

Multiphor SSCP/HD analysis of the amplicon containing exon 9, revealed one novel synonymous SNP c.600 C>T; p.Asp110Asp (Appendix F, Figure 8a). SNP c.600 C>T; p.Asp110Asp occurred at a very low frequency, as only one heterozygous individual was detected in the control cohort (1.9%). Semi-automated DNA sequence analysis was employed to confirm the genotype (Appendix F, Figure 8b).

Multiphor SSCP/HD analysis of the amplicon containing exon 12 revealed the presence of a novel exonic SNP c.934 G>A; p.Gly312Ser (Appendix F, Figure 10). This non-synonymous exonic SNP is responsible for a change from the amino acid glycine, a hydrophobic amino acid with non-polar side chains to serine, which has polar side chains. Genotypes determined on the Multiphor SSCP/HD system were confirmed with a *BceAI* (New England Biolabs, USA) restriction enzyme digest (Figure 3.3). Digestion fragments of 216 bp and 62 bp resolved on a 3% agarose gel, represented homozygous status for the A allele, while three fragments of 187 bp, 62 bp and 26 bp represented homozygous status for the G allele. This SNP was detected at a very low frequency in the total maternal patient group (2.5%), the total fetal patient group (3.4%) and the control group (7.7%).

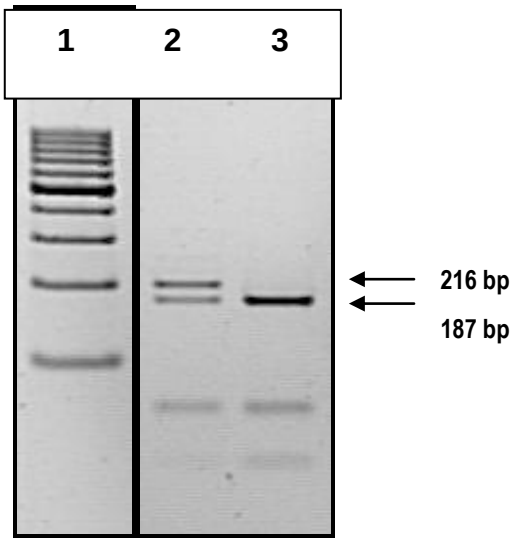


Figure 3.3 *BceAI* restriction enzyme analysis of c.934 G>A; p.Gly312Ser in the amplicon containing exon 12 resolved on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lane 2 = GA, lane 3 = GG.

Multiphor SSCP/HD analysis of the amplicon containing exon 13 revealed one previously documented SNP c.975 C>T; p.Gly325Gly (dbSNP: rs12442554) (Appendix G, Figure 11). This synonymous SNP (c.975 C>T; p.Gly325Gly) located at amino acid position 325 is responsible for a silent mutation.

Conformation 4 could not be distinguished from conformation 1 on Multiphor SSCP/HD gels and a *HphI* (New England Biolabs, USA) restriction enzyme digest was used to confirm SNP c.975 C>T; p.Gly325Gly and thereby differentiate between these conformations and minimise mistyping of this locus (Figure 3.4). Digestion fragments of 60 bp and 173 bp resolved on a 3% agarose gel indicated individuals homozygous for the C allele. Three fragments 60 bp, 62 bp and 111 bp indicated individuals homozygous for the T allele.

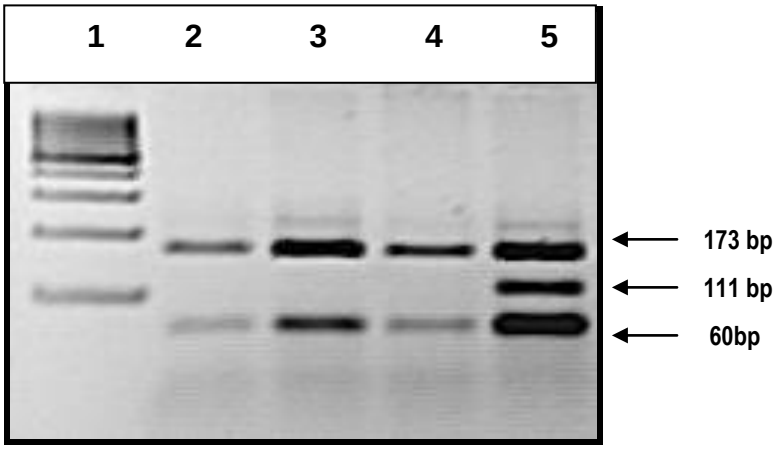


Figure 3.4 *HphI* restriction enzyme analysis of c.975 C>T; p.Gly325Gly in the amplicon containing exon 13 resolved on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lanes 2, 3, 4 = CC, lane 5 = CT.

3.2.3 Intronic Variants

Following Multiphor SSCP/HD analysis, seven novel (c.244-42 G>C, c.244-76 C>G, c.528+38 C>T, c.589-5 C>T, c.682+49 C>T, c.961-30 A>G, c.961-24 C>G) and six previously identified (c.-12+75 C>A, c.-11-43 G>A, c.-11-13 A>T, c.48+67 C>T, c.449-17 G>A, c.683-56 G>A) variants were identified in the non-coding regions flanking the exons in the *ANXA2* gene.

One previously documented intronic SNP c.-12+75 C>A (rs12904756) within intron 1 was identified upon Multiphor SSCP/HD analysis of amplicon 1b,

containing exon 1 (Appendix F, Figure 3). Multiphor conformations of this SNP were clear and therefore used to perform genotyping in the patient and control cohorts.

Multiphor SSCP/HD analysis of the amplicon containing exon 2 revealed three previously documented DNA sequence variants viz, c.-11-43 G>A (dbSNP: rs11858864), c.-11-13 A>T (dbSNP: rs11855679), located within intron 1, and c.48+67 C>T (dbSNP: rs3743268) located within intron 2 (Appendix F, Figure 4).

RFLP analysis using the restriction enzymes *MboII* (Fermentas, Burlington, USA) and *RsaI* (Fermentas, Burlington, USA) confirmed the genotypes observed on the Multiphor SSCP/HD gels for the following two variants; c.-11-43 G>A and c.-11-13 A>T (Figure 3.5). In the presence of the c.-11-43 G>A SNP, *MboII* cleaved the A allele into two fragments (34 bp and 182 bp) while individuals homozygous for the G allele could be identified by fragments of 22 bp and 194 bp on 3% agarose gels. *RsaI* digestion for variant c.-11-13 A>T, yielded fragment sizes of 12 bp, 52 bp and 152 bp, and these products were visualised on 3% agarose gels for the identification of the A homozygosity status. Individuals homozygous for the T allele could be identified by fragments of 12 bp and 204 bp. SNP c.48+67 C>T was genotyped according to Multiphor typings, since no restriction endonuclease recognition sites could be identified. The homozygous status of the variant allele of SNP c.-11-43 G>A occurred at a higher frequency in the control cohort (46.3%) compared to the maternal (36.4%) and fetal cohorts (36.2%). The frequency of SNP c.-11-13 A>T in the maternal and fetal patient and control cohorts were identical to that observed for SNP c.-11-43 G>A. This could potentially be explained by LD between these two closely-located SNPs in this population. The variant allele T of SNP c.48+67 C>T was only observed in one heterozygous patient in the maternal group, at a frequency of 0.9% but was absent in the control and fetal groups.

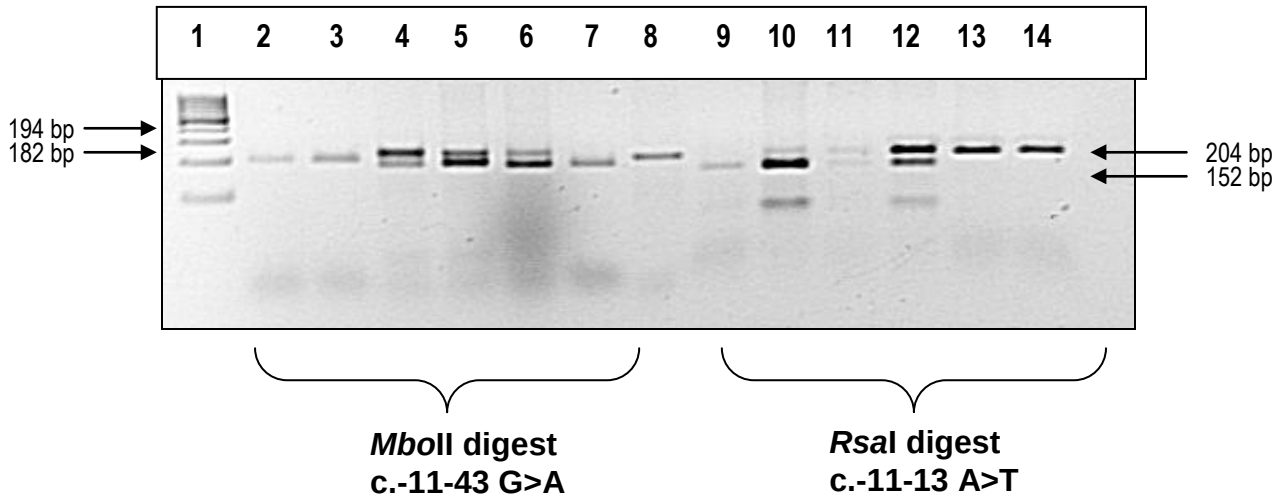


Figure 3.5 *MboII* and *RsaI* restriction enzyme analysis of c.-11-43 G>A and c.-11-13 A>T in the amplicon containing exon 2, resolved on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lanes 2, 3 = GG; lanes 4, 5 = GA; lanes 6, 7 = AA; lanes 9, 10 = AT; lanes 11, 12 = AT, lanes 13, 14 = TT.

Two novel intronic SNPs (c.244-76 C>G and c.244-42 G>C) were detected within intron 4 following bi-directional DNA sequencing of the conformational variants identified by Multiphor SSCP/HD analysis of the amplicon containing exon 5 (Appendix F, Figure 6). None of the patient or control individuals was found to be homozygous for any of these SNPs.

Multiphor SSCP/HD analysis of the amplicon containing exon 7 revealed a previously documented SNP c.449-17 G>A (dbSNP: rs12898604) and novel SNP c.528+38 C>T within introns 6 and 7, respectively (Appendix F, Figure 7). Genotypes on the Multiphor SSCP/HD gels could be distinguished with certainty and were thus assigned according to Multiphor typings.

Multiphor SSCP/HD analysis of the amplicon containing exon 9 revealed two novel SNPs c.589-5 C>T and c.682+49 C>T within introns 8 and 9, respectively (Appendix F, Figure 8). The Multiphor conformational variant used to distinguish individuals heterozygous for the SNP c.589-5 C>T could

not always be genotyped with certainty. All samples were therefore subjected to a *Tsp509I* (New England Biolabs, USA) restriction enzyme digest and visualised on 3% agarose gels to identify and confirm genotypes (Figure 3.6). In the presence of the T allele *Tsp509I* cleaved the amplicon into two smaller restriction digestion fragments (27 bp and 182 bp) and individuals homozygous for the C allele could be identified by an uncut fragment of 209 bp.

AciI (New England Biolabs, USA) restriction digest, resolved on 3% agarose gels confirmed the SNP c.682+49 C>T by revealing fragments 105 bp and 200 bp in size (representing homozygosity for the T allele). Homozygosity for the C allele was reflected by three fragments (31 bp, 74 bp, 200 bp) (Figure 3.7).

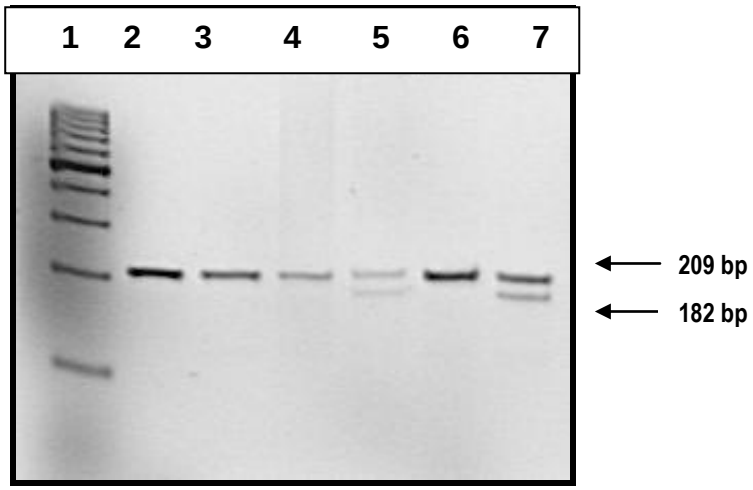


Figure 3.6 Mutagenic primers amplified a 209 bp fragment for a *Tsp509I* restriction enzyme digest that enabled identification of SNP c.589-5 C>T in the amplicon containing exon 9 on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lane 2 = undigested fragment (209 bp), lanes 3, 4, 6 = CC, lanes 5, 7 = CT.

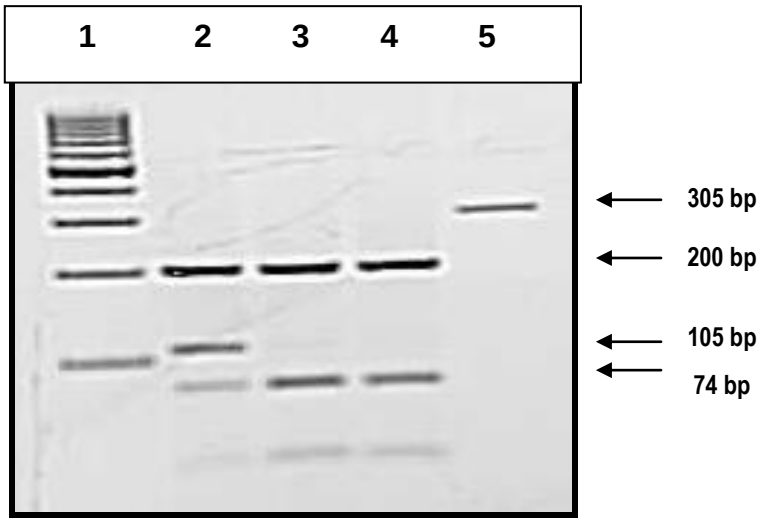


Figure 3.7 *AciI* restriction enzyme analysis of c.682+49 C>T in the amplicon containing exon 9 resolved on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lanes 2 = CT, lanes 3, 4 = CC, lane 5 = undigested fragment (305 bp).

Multiphor SSCP/HD analysis of the amplicon containing exon 10 revealed several unique conformations at several temperatures during optimization, but once genotyped at large scale no variation could be detected. Running temperatures and duration, the amount of PCR product and loading dye were adjusted, but nonetheless no variation could be detected. The different conformations identified during optimization were therefore subjected to bi-directional DNA sequencing, which revealed the presence of the previously documented SNP c.683-56 G>A (dbSNP: rs11633619) within intron 9 (Appendix F, Figure 9).

This locus was genotyped using a restriction enzyme for digestion with *BclI* (New England Biolabs, USA) of the amplicon containing exon 10 (Figure 3.8). Digestion fragments 21 bp, 49 bp and 230 bp, resolved on 3% agarose gels, indicated homozygosity for the A allele. Homozygous status for the G allele could be recognized by fragments of 21 bp and 279 bp.

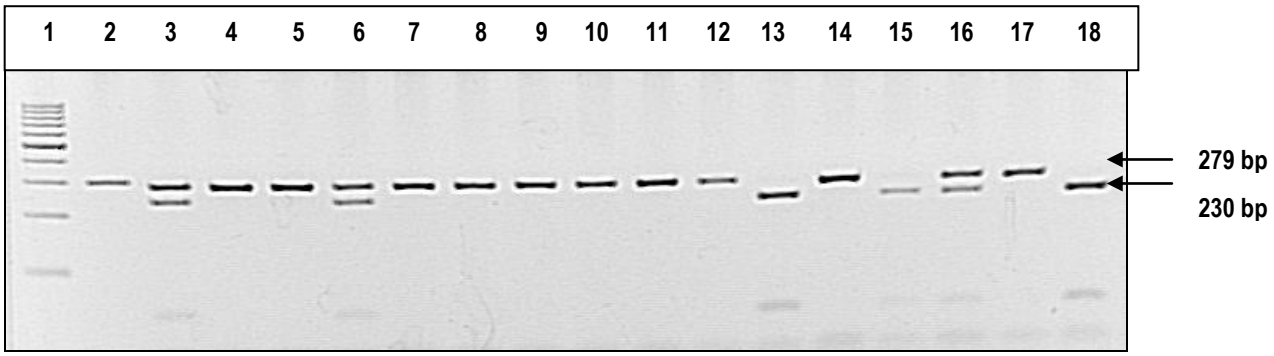


Figure 3.8 *BclI* restriction enzyme analysis of c.683-56 G>A in the amplicon containing exon 10 resolved on a 3% (w/v) agarose gel. Lane 1 = 100 bp molecular weight marker (O'Generuler^{TM}, Fermentas); lane 2 = undigested fragment (300 bp), lanes 3, 6, 16 = GA, lanes 4, 5, 7-12, 14, 17 = GG; lanes 13, 15, 18 = AA.

Multiphor SSCP/HD analysis of the amplicon containing exon 13 revealed two novel intronic SNPs c.961-30 A>G and c.961-24 C>G within intron 12 (Appendix F, Figure 11). SNP c.961-30 A>G was genotyped according to Multiphor typings, which could easily be distinguished from other conformations. Genotypes for SNPs c.961-24 C>G was confirmed by semi-automated DNA sequencing (Appendix F, Figure 11b), since only one control sample showed heterozygous status. No homozygous status for these two SNPs was identified.

3.2.4 Variants in the *ANXA2* 3'UTR

Multiphor SSCP/HD analysis of the amplicon containing exon 13 revealed the novel SNP c.*1057 A>G in the *ANXA2* 3'UTR (Appendix F, Figure 11). SNP c.*1057 A>G is located 37 bp downstream the stop codon in the 3'UTR in exon 13. Only one heterozygous genotype was observed in the maternal patient cohort (0.9%), indicating that this SNP occurred at a very low frequency in this population.

3.3 Statistical Analysis

Allele and genotype frequencies were determined for all loci in the *ANXA2* gene, and analysed for deviation from the Hardy-Weinberg Equilibrium (HWE) with Chi-square analysis in Tools For Population Genetic Analysis (TFPGA) Software v1.3 (Miller, 1997). The frequencies for 23 of the 25 variants identified in the *ANXA2* gene were in HWE (Appendix G, Tables 1-12, together with the *P*-values). *P*-values of less than 0.05 were considered as significant and are highlighted in red (Appendix G, Tables 1-12).

Statistical significant associations between patient and control cohorts were determined by an analog to Fischer's exact test and verified by way of 2X3 contingency tables in Microsoft® Office Excell (2003) and are presented in Appendix G, Tables 1-12. Patient cohorts were subdivided into maternal and fetal patient cohorts and compared to a control group. Patients were further stratified into subgroups according to their ethnicity, namely Mixed Ancestry and South African Black maternal and fetal groups. *P*-values of less than 0.05 were considered as significant and are highlighted in blue in Appendix G, Tables 1-12.

3.3.1 Variants in the *ANXA2* 5'UTR

Genotype and allele frequencies for four (c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG and c.-135 C>G) of the six DNA sequence variants identified in the 5'UTR, showed no deviation from HWE (*P*>0.05) in any of the maternal and fetal patient or control study cohorts.

Genotype and allele frequencies for SNP c.-92 A>T in the total maternal cohort deviated from HWE (0.0431), however no deviation was observed in the total fetal and control cohorts, or any of the maternal and fetal subgroups (*P*>0.05). The genotype distribution for SNP c.-31 T>C in all maternal, fetal and control study groups deviated from HWE (*P*<0.05).

No significant differences in genotype and allele frequencies were observed between maternal patient and control cohorts or the fetal patient and control cohorts for sequence variants c.-442 C>G, c.-191 G>C, c.-189_-188insGCCGG and c.-135 C>G.

Statistical significance was observed at locus c.-92 A>T when comparing the total maternal patient group (p=0.039) and the Black fetal patient subgroup (p=0.019) to the control group. The minor allele T for this SNP occurred at a higher frequency in the total (16.4%), Mixed Ancestry (15.9%) and Black (17.1%) maternal patient groups, compared to the control (7.8%) group.

Statistical significant differences in genotype frequencies between the total maternal patient and control cohorts (p=0.039), total fetal patient and control cohorts (p=0.003) and Black fetal patient and control cohorts (p=0.007) were observed for SNP c.-31 T>C. The variant allele C was present at a higher frequency in the total maternal patient cohort (62.1%), the Mixed Ancestry maternal patient cohort (61.4%) and the Black maternal patient cohort (63.4%) compared to the frequency observed in the control cohort (44.1%).

3.3.2 Coding Variants

Genotype and allele frequencies observed in the maternal, fetal and control study cohorts for all variants identified in the coding regions of the *ANXA2* gene were in HWE (*P*>0.05).

Only one of the five exonic SNPs revealed a significant difference in genotype and allele frequencies between the patient and control study cohorts. A significant difference in allele frequencies was observed between the total maternal and control cohorts (p=0.008), the Mixed Ancestry maternal and control cohorts (p=0.035) and the total fetal and control cohorts (p=0.019) for SNP c.222 C>T; p.Ala74Ala located within exon 4.

3.3.3 Intronic Variants

Two variants (c.-12+75 C>A and c.244-76 C>G) identified in the non-coding regions of the *ANXA2* gene showed deviation from HWE in certain subgroups. Genotype and allele frequencies for all the other DNA sequence variants identified in the non-coding regions of the *ANXA2* gene showed no deviation from HWE ($P>0.05$) in the maternal and fetal patient or the control study cohorts.

The genotype distribution for SNP c.-12+75 C>A deviated from HWE in the Black maternal patient group (0.0426), although none of the other study groups deviated from HWE at this locus. Genotype and allele frequencies observed in the Mixed Ancestry fetal patient group for SNP c.244-76 C>G located within intron 4 deviated from HWE (0.0000). One heterozygous C/G individual (2.3%) and one individual homozygous for the G allele (2.3%) were observed in the Mixed Ancestry fetal patient group at locus c.244-76 C>G.

The only intronic loci to reveal a significant difference in genotype and allele frequencies between the maternal patient and control and fetal patient and control cohorts were the variants c.244-76 C>G, c.449-17 G>A and c.589-5 C>T.

A significant difference was identified between the Black maternal patient group and the control group (p=0.018) at locus c.244-76 C>G. Statistical significant differences were observed in genotype and allele frequencies between the Black maternal patient group and the control group (p=0.032), along with a significant difference in allele frequency between the Black fetal patient and control group (p=0.030) at locus c.449-17 G>A.

Marginal association with PE was identified at locus c.589-5 C>T in the total fetal patient cohort (p=0.041), Mixed Ancestry fetal patient cohort (p=0.041) and the Black fetal patient cohort (p=0.041), but the variant allele T was

absent in the control cohort. No significant differences was detected between maternal and control cohorts.

3.3.4 Variants in the *ANXA2* 3'UTR

Genotype and allele frequencies for SNP c.*1057 A>G identified in the 3'UTR of the *ANXA2* gene showed no deviation from HWE (*P*>0.05) in the maternal and fetal patient or the control study cohorts.

No significant differences in genotype and allele frequencies were observed between any of the patient and control cohort for SNP c.*1057 A>G.

3.4 Comparison of the Genotype and Allele Frequencies in Different Ethnic Populations

Genotype and allele frequencies in European, Asian and Sub-Saharan African populations of nine previously documented loci identified in this study, were compared to the frequencies observed in this study population (Appendix H, Table 1). These frequencies were retrieved on 27 July 2009 from the Entrez SNP database, (http://www.ncbi.nlm.nih.gov/SNP/).

3.5 Haplotype Analysis

Haplotype analyses were performed on maternal and fetal patient cohorts independently within each subgroup (Mixed Ancestry and Black) within Haploview, version 3.31 (Barett *et al*., 2005). SNP c.-31 T>C (marker 6) located within exon 1 was excluded from haplotype analysis, because of its deviation from the HWE in both patient and control cohorts (P<0.05). All loci included in haplotype analysis are presented in Appendix H, Table 2. Results of the analysis on the Mixed Ancestry and Black maternal patient cohorts using the four gamete rule (Wang *et al*., 2002) are presented in Table 3.3. The same analyses were replicated on the Mixed Ancestry and Black fetal patient cohorts and the results are presented in Table 3.4.

In a previous study by Bogdanova *et al*. (2007) a haplotype in the *ANXA5* promoter was identified in a German population and associated with RPL, another poor placental disease like PE. Variants identified in the 5'UTR of the *ANXA2* gene were subjected to haplotype analysis in attempt to identify a similar haplotype in the promoter region of *ANXA2* that could possibly be associated with PE.

3.5.1 Maternal and Control Cohorts

The four gamete rule within Haploview identified two intronic SNPs, c.-11-43 G>A and c.-11-13 A>T, indicating a significant level of LD between the loci in both the Mixed Ancestry and Black maternal patient cohorts. LD between these two SNPs was demonstrated by D' value = 1.0, LOD score = 51.05, r^2 value = 1.0 in the Mixed Ancestry maternal patient subgroup and by D' value = 1.0, LOD score = 43.56, r^2 = 1.0 in the Black maternal patient subgroup (indicated in red) (Table 3.3). LD was absent from other haplotype blocks identified in this analysis.

Table 3.3 Haplotype analysis of the maternal patient and control cohorts employing the four gamete rule.

Study	**Haplotype Blocks**	**1**	**2**	**3**	**4**
Cohort	**Markers**	**2-7**	**8-9**	**10-19**	**20-25**
Mixed Ancestry	**D'**	1.0	**1.0**	1.0	1.0
	LOD	0.13	**51.05**	0.01	0.11
	r^2	0.001	**1.0**	0.0	0.005
Study	**Haplotype Blocks**	**1**	**2**	**3**	**4**
Cohort	**Markers**	**1-7**	**8-9**	**11-16**	**17-25**
South African Black	**D'**	0.383	**1.0**	1.0	1.0
	LOD	0.21	**43.56**	0.03	0.01
	r^2	0.015	**1.0**	0.0	0.0

3.5.2 Fetal and Control Cohorts

The four gamete rule revealed a significant level of LD between SNPs c.-11-43 G>A and c.-11-13 A>T, in both the Mixed Ancestry fetal (D' value = 1.0, LOD score = 47.21 and r^2 value = 1.0) and Black fetal (D' value = 1.0, LOD score = 40.19 and r^2 value = 1.0) patient subgroups (indicated in red) (Table 3.4). These previously identified intronic SNPs are located 30 bp from each other within intron 1. LD was absent from the other haplotype blocks identified in the analysis of the fetal patient and control cohorts.

Table 3.4 Haplotype analysis of the fetal patient and control cohorts employing the four gamete rule.

Study	**Haplotype Blocks**	**1**	**2**	**3**	
Cohort	**Markers**	**1-7**	**8-9**	**11-25**	
Mixed Ancestry	**D'**	1.0	**1.0**	1.0	
	LOD	0.11	**47.21**	0.01	
	r^2	0.001	**1.0**	0.0	
Study	**Haplotype Blocks**	**1**	**2**	**3**	**4**
Cohort	**Markers**	**2-7**	**8-9**	**11-15**	**16-25**
South African Black	**D'**	1.0	**1.0**	1.0	1.0
	LOD	0.07	**40.19**	0.18	0.0
	r^2	0.001	**1.0**	0.003	0.0

3.6 In Silico Analysis

3.6.1 Promoter Variants

In silico analysis of the variants identified in the 5'UTR was performed using the internet based transcription factor binding site prediction databases TESS, (http://www.cbil.upenn.edu/cgi-bin/tess), MatInspector (http://www.genomatrix.de/cgi-bin/matinspector_prof/), AliBaba2.1 (http://www.gene-regulation.com/pub/programs/alibaba2/index.html) and JASPAR CORE (http://jaspar.genereg.net/). The results of these predictions are presented in Table 3.5.

The putative transcription factor binding sites predicted to be created or abolished in the presence of each variant were compared between the different databases. The concurrent results from more than one database would increase the strength of any predictions. In this study, the abolishment of a putative Sp1 site in the presence of the c.-191 G>C SNP as predicted by TESS and AliBaba2, was the only corresponding results between the four databases employed. Other transcription factor binding sites that were frequently predicted to be abolished in the presence of variants identified in the 5'UTR were the sites NF-1 and BRCA1, which will be included in the discussion.

Table 3.5 Putative transcription factor binding sites in the promoter region of the *ANXA2* gene, predicted to be created or abolished in the presence of each of the six variants identified in the 5'UTR.

	TESS		MatInspector		Alibaba2		JASPAR CORE	
	Creates	**Abolished**	**Creates**	**Abolished**	**Creates**	**Abolished**	**Creates**	**Abolished**
c.-442 C>G	CTCF,CAC1, CAC2, CACCCC binding factor	T-Ag Gal4	-	-	-	-	M2F1_1-4	-
c.-191 G>C	-	ETF (2X) CP1,LSF,**Sp1**, NF-S GCF,GAL4	XCPE1	PLAG1,ZF9 EGR1,KKLF	WT1	NF-1 (2X) **Sp1**	-	-
c.-189_-188insGCCGG	-	-	XCPE1 KLF6,Sp2 MAZ	-	WT1,Sp1 (2X) ETF, appaB, AP,1ph	NF-1,SRY Sp1,YY1,AP	TFAP2A (2X)	-
c.-135 C>G	T-Ag,Sp1 LBP-1	-	-	-	AP-2alpha c-Myc	-	-	BRCA1
c.-92 A>T	RAF	Sp1 ABF1	DMTE	-	-	-	-	-
c.-31 T>C	-	TEF2	-	TIEG BKLF	Sp1	-	-	BRCA1 CREB1

3.6.2 Coding Variants

Five loci were identified with Multiphor SSCP/HD analysis within the coding regions of the *ANXA2* gene. SNP c.*1057 A>G is located in exon 13, 37 bp downstream of the stop codon and was therefore excluded from *in silico* analysis.

Exonic SNPs were analysed using the software program SIFT (Sorting Intolerant From Tolerant) (http://blocks.fhcrc.org/sift/SIFT/html) to determine the putative effect of amino acid substitutions on the function of the annexin II protein. Of the six SNPs analysed c.934 G>A; p.Gly312Ser was predicted to affect protein function with a score of 0.01, a normalized probability score < 0.05 was considered to be deleterious. PMUT (http://mmb2.pcb.ub.es:8080/PMut/), a software program based within MMB (Molecular Modeling and Bioinformatics Group) (http://mmb.pcb.ub.es/) confirmed the prediction obtained with SIFT, in that SNP c.934 G>A; p.Gly312Ser might have a possible pathological effect on the function of annexin II. PMut predicted this with a NN output value = 0.8816, where an output > 0.05 was considered as pathological.

Above mentioned exonic SNPs were analyzed using internet based program ESE Finder version 3.0 (http://rulai.cshl.edu/tools/ESE/) (Cartegni *et al*., 2002). The creation of an exonic splicing enhancer (ESE) motif was predicted in the presence of the variant T allele at the c.292 G>T; p.Val98Leu locus. This software program is based on the fact that SR (Serine/Arginine) proteins such as SF2/ASF, SRp40, SRp55 AND SC35 are known to bind to these motifs (Smith *et al*., 2006; Wang *et al*., 2005b).

3.6.3 Intronic Variants

Thirteen variants identified in the non-coding regions of the *ANXA2* gene were subjected to *in silico* analysis to predict their possible roles in alternative splicing. The software program ASSP (Alternative Splice Site Predictor)

(http://www.es.embnet.org/~wang/asp.html) predicted that only one intronic variant, c.244-76 C>G could affect alternative splicing by the abolishment of an acceptor site in the presence of the variant allele G, this result however was not confirmed by the software program NetGene2 (http://www.cbs.dtu.dk/services/NetGene2). NetGene2 additionally predicted the abolishment of a donor site in the presence of the c.449-17 G>A variant allele A and the creation of an acceptor site in the presence of the c.961-30 A>G variant allele G.

4 DISCUSSION

4.1 Genetic Data Analysis

Sixteen of the 25 variants identified in the *ANXA2* gene and the proximal promoter region are novel, of which five are located within the 5'UTR, three are exonic, seven are intronic, and one is located within the 3'UTR. Furthermore, significant associations were identified between four novel (c.-92 A>T, c.222 C>T; p.Ala74Ala, c.244-76 C>G and c.589-5 C>T) and one previously documented (c.449-17 G>A) variants and PE.

Most variants in the human genome occur within intronic regions or the promoter area of genes (Haga *et al.*, 2002; Hughes *et al.*, 2003). Results obtained in this study substantiate these findings as 13 of the 25 variants identified were located within non-coding regions, and only six variants identified in the 5'UTR, five identified in the coding regions of the gene and one identified in the 3'UTR. Despite this knowledge, SNPs in the coding regions of genes have previously been considered as the most significant SNPs to investigate, as these SNPs were believed to be the ones to most likely have a functional effect. Three of the five exonic SNPs identified in this study were synonymous SNPs which correlate with findings by Haga *et al*. (2002) that allele frequencies of silent mutations within coding regions of genes have been documented to be much higher than non-synonymous SNPs.

Seven of the 25 variants identified in this study were responsible for C to T transitions, of which four formed CG dinucleotides. This, however, was expected because of the mutability of CG dinucleotides, which are also known as mutation hotspots (Bird, 1987; Cooper and Youssoufian, 1987). DNA methylation of CG dinucleotide sequences account for the conversion of cytosines to 5'mC, which in turn become uracils upon deamination. When the deaminated base has not been repaired, the uracil becomes a thymidine (Bird, 1987; Cooper and Youssoufian, 1987), resulting in a C to T transition. Improved understanding of mutational processes such as these hotspots,

could promote our knowledge of genetic processes such as replication, transcription and recombination, consequently contributing to our knowledge of disease mechanisms (Duret, 2009).

Knowledge about variation in the *ANXA2* gene is limited and all variants identified in this case-control association study would thus contribute to our understanding of the association between specific genetic variants in the *ANXA2* gene and PE.

4.1.1 *ANXA2* 5'UTR and 3'UTR

The 5'UTR and 3'UTR of human genes vary from gene to gene and are consequently exceptionally diverse, contributing to the difficulty in analysing and characterising these regions. These untranslated areas are known to contain various regulatory elements to which transcription factors would bind for functional transcriptional activation. Mutations in the 5'UTR and 3'UTR could influence the structure of chromatin, directly affecting access of transcription factors and RNA polymerase II to chromatin and therefore the binding of these factors to DNA. Mutations in transcription factors and transcription factor binding sites have been associated with disease (Remmers *et al*., 2007; Edberg *et al*., 2008). It is therefore clear that the analysis of the *ANXA2* 5'UTR and 3'UTR would play a critical role in determining whether variants in these regions could be associated with severe early onset PE in the South African population.

i) HWE Analysis and Significant Associations

The deviation from HWE at SNP c.-92 A>T in the total maternal pre-eclamptic patient cohort could be an indication of a genuine genetic association between this locus and the pathophysiology of PE (Schaid and Jacobsen, 1999). In this study, the deviation from HWE could therefore be a result of the genetic mechanism of the disease causing an association of the minor allele T with PE, because of a rare dominant disease susceptibility allele. The significant

associations identified in the total maternal (p=0.039) and Black fetal (p=0.019) patient cohorts between the locus c.-92 A>T and PE corroborates the deviation from HWE in the total maternal patient cohort, since an association between this SNP and the underlying etiology of PE could mean that the minor allele T play a role as a susceptibility factor to severe early onset PE.

A significant departure from HWE at locus c.-31 T>C in the maternal and fetal patient and control study cohorts was identified in this study. Possible reasons for deviation from HWE could be because of assertive mating, inbreeding, selection, random genetic drift, gene flow, population stratification, mutations, DNA quantity or quality and genotyping errors (Hartl *et al*., 1997). In this study, low DNA quantity or quality due to degradation were considered a possible reason. This could thus explain the lack of heterozygous individuals observed in this study population as it could be the result of allelic dropout or false alleles (Taberlet *et al*., 1996). A more recent study by Pompanon *et al*. (2005) emphasized the reality of human handling and experimental errors and the consequences of these genotyping errors on results. In this study however, experimental procedures, genotyping and statistical analysis were performed with immense care and precision to ensure accurate and reliable results.

Unfortunately no allele or genotype frequencies for the SNP c.-31 T>C are electronically available for the African-American and Sub-Saharan African populations. A previous study by Hassan *et al*. (2003) demonstrated a significant difference in genotype and allele distributions in selected cytokines between African-American and Caucasian populations. According to genotype frequencies retrieved from the Entrez SNP database, (http://www.ncbi.nlm.nih.gov/SNP) the minor allele C of SNP c.-31 T>C was present in the Caucasian population at a frequency of 92.0%, compared to the frequency of 62.3% observed in the South African population. In accordance with the study by Hassan *et al*. (2003) this study substantiates that differences

in allele and genotype frequencies are found between various ethnic population groups.

The significant associations observed in both the maternal and fetal patient study cohorts between the SNP c.-31 T>C and PE were not taken into account, due to the significant departure from the HWE in both patient and control study cohorts. Other than SNPs c.-92 A>T and c.-31 T>C, none of the other variants identified in the proximal promoter region or the 3'UTR of the *ANXA2* gene in the patient and the control study cohorts, deviated from the HWE or demonstrated a significant deviation in the allele and genotype frequencies between the control and maternal and fetal PE patient study cohorts.

ii) *In silico* Analysis

Transcription factors play a critical role in the regulation of transcription from DNA to RNA through RNA polymerase II. In a study by Iwakuma *et al*. (2005) dysfunctional transcriptional regulation by the disruption of transcription factors as a result of mutations in the DNA binding domain of p53 was implicated in severe disease phenotypes such as cancer. Recently, the focus on promoter analysis and characterisation has expanded tremendously, as knowledge regarding the transcriptional regulation of genes has been limited. Information on how this process could be affected by the disruption or activation of transcription factors may explain the association of variants c.-92 A>T and c.-31 T>C with the study cohorts and PE.

Promoter analysis in this study revealed conflicting results, with only TESS and AliBaba2 corresponding in the prediction that the SNP c.-191 G>C would create a Sp1 site. Sp1 (specificity protein 1) is known to function either as an activator or as a basal promoter element (Bouwman and Philipsen, 2002; Nichols *et al*., 2003), depending on the nature of the promoter, which binds and acts through CG boxes in order to regulate transcription in multiple genes (Suske, 1999). The *ANXA2* promoter region is particularly CG-rich and the

creation of a Sp1 site in the promoter region could therefore affect the transcriptional activity of the *ANXA2* gene. However, no association was identified between this SNP and PE and the predicted effect would therefore not be strong.

Two other transcription factors that were frequently predicted to be abolished in the presence of the variants located within the *ANXA2* 5'UTR were NF-1 and BRCA1. Alibaba2 predicted that NF-1 would be abolished in the presence of the variants c.191 G>C and c.-189_-188insGCCGG and JASPAR CORE predicted that BRCA1 would be abolished in the presence of the variants c.-135 C>G and c.-31 T>C.

NF-1 (nuclear factor 1) is a constitutively expressed transcription factor that can either be activated or repressed during transcription. A recent study by Zhao and Ennion (2006) confirmed NF-1 functioning as a repressor, by demonstrating that mutations in NF-1 resulted in a reduced $P2X_1$ promoter activity. BRCA1 (breast cancer susceptibility gene 1) was an interesting transcription factor to investigate, since it has been implicated in chromatin folding and is known for its role in DNA-repair, rather than the transcriptional process itself (Ye *et al*., 2001). The mechanism by which BRCA1 functions is different to that of Sp1, since it does not bind to DNA sequences specifically, but associates with DNA independently and has a preference for abnormal DNA structures (Paull *et al*., 2001).

Variant alleles of the SNPs identified in the *ANXA2* proximal promoter of the *ANXA2* gene could affect transcriptional regulation by creating or disrupting putative transcription factor binding sites. In the case of the creation of a putative site (eg, in this study the creation of a putative Sp1 site caused by the SNP c.-191 G>C), the binding of a corresponding transcription factor could effect transcriptional regulation by acting either as an activator or a repressor. The abolishment of a transcription factor binding site (eg, in this study, the abolishment of the putative BRCA1 caused by variants c.-191 G>C and c.-189_-188insGCCGG and NF-1 sites caused by the SNPs c.-135 C>G and

c.-31 T>C) would in theory cause that the corresponding transcription factor cannot bind and 'normal' transcriptional regulation may be disrupted.

These predictions suggest that the above-mentioned SNPs could affect the transcriptional activity of the *ANXA2* gene and in doing so, affect the up- and down regulation of annexin II expression. A variation in the levels of the annexin II protein could contribute to the binding of this protein to PP13, and consequently affect the transport of PP13 across the maternal-fetal interface. Although these predictions give an indication of the putative effect of the variants in the *ANXA2* 5'UTR and the possible role it could play in the progression of this disease, it is recommended that the effect of gene regulation should be validated experimentally.

The SNP c.*1057 A<G, located within the 3'UTR, was identified in one patient, and only in a heterozygous state. Although this SNP was present at an extremely low frequency, variants in the 3'UTR have previously shown to influence mRNA stability (Wang *et al*., 2005a; Orsi, 2007). Unlike the case for the 5'UTR, no *in silico* tools are available to predict the putative effect of the c.*1057 A>G SNP on the regulation of translation initiation in the *ANXA2* gene. Since the 3'UTR is proposed to enhance the translational efficiency of mRNA and play a role in the transport of mRNA from the nucleus to the cytoplasm (Sachs *et al*., 1997), variants in the *ANXA2* 3'UTR could affect annexin II protein formation. The predicted functional effect of SNP c.*1057 A>G on annexin II however needs to be validated experimentally.

4.1.2 Coding Regions

A study by Cargill *et al*. (1999) indicated that only a few SNPs in the coding sequences of genes would be present at a high frequency. However, SNPs in regulatory and coding regions of the genes, especially those responsible for amino acid substitutions are thought to have the largest impact on protein function (Collins *et al.*, 1997). The identification of variants in the coding region of the *ANXA2* gene is therefore crucial, because it could have a

functional impact on the annexin II protein and thereby contribute to the phenotype of the disease.

i) HWE Analysis and Significant Associations

A statistically significant association between the coding SNP c.222 C>T; p.Ala74Ala located within exon 4, and severe early onset PE was identified in the total maternal (p=0.008), Mixed Ancestry maternal (p=0.035) and total fetal (p=0.019) patient study cohorts. The minor allele T was present in three of the 54 control individuals, but was completely absent from all maternal and fetal cohorts, implying a possible protective effect. It could be hypothesised that the T allele contributes to a reduced risk for the development of severe early onset PE in this specific study population. PE, however, is a complex disease and numerous variants in multiple genes would play a role in the pathology of PE. Although single variants have previously been associated with disease, the frequency of these alleles remain low and the association demonstrated between this SNP and PE needs to be confirmed in other studies.

ii) *In silico* Analysis

Although a significant association was identified between variant c.222 C>T; p.Ala74Ala and PE, *in silico* analysis revealed no predicted functional or pathological effect of this variant on the annexin II protein. *In silico* analysis, however, revealed interesting results for variants that showed no significant association with PE.

As reported in section 3.4.3, the exonic SNP c.934 G>A; p.Gly312Ser was predicted to affect protein function by two independent software programs SIFT and PMUT. These programs predicted that the Glycine to Serine amino acid substitutions at residue 312 in the annexin II protein could have a pathological affect on the function of annexin II. The SNP could consequently affect the structure and function of the *ANXA2* gene in cellular, molecular and

physiological processes. Although results obtained by *in silico* analysis are only an indication of the effect of the variants involved, several studies have confirmed results obtained by these predictive programs (Volinia *et al*., 2005; Skalsky *et al*., 2007).

Exonic splicing elements (ESE) are motifs that play an imperative role in recognising alternative splice sites via serine arginine (SR) proteins. These elements occur frequently in the human genome and are present in most exons (Wu *et al*., 2005). The creation of an ESE motif, as predicted by ESE Finder in the presence of the variant allele T at SNP c.292 G>T; p.Val98Leu, could affect mRNA splicing by inaccurate and ineffective recognition of the exon boundaries.

Pre-mRNA splicing has been shown to be influenced by variants such as missense, nonsense and silent DNA sequence variants (Pagani and Baralle, 2004). These variants could alter the functioning of splicing machinery, resulting in exon skipping or the activation of cryptic splice sites. The accuracy and efficiency of splicing are directly affected by these splicing abnormalities, resulting in the formation of defective or aberrant mRNA isoforms which could be associated with the pathophysiology of PE.

The translation of aberrant mRNA isoforms could affect the final protein product, which would subsequently impact the annexin II/PP13 binding. Impaired binding could influence the transport of PP13 across the maternal-fetal interface which would explain reduced PP13 protein levels observed in pre-eclamptic pregnancies and contribute to our understanding of this complex disease. Functional analyses are thus recommended to validate the predictions and effects of the variants identified in the *ANXA2* coding region.

4.1.3 Intronic Regions

Introns were once considered “junk” DNA, as they are removed from the genome and were not considered to play a role in gene expression. In a study by Nott *et al.* (2003) the significant role of introns in molecular evolution and regulation of gene expression was emphasized. This study employed a luciferase-based reporter system in which contructs containing the TPI intron 6 in two different locations within the open reading frame of the *Renilla* luciferase indicated that intron 6 significantly improved mRNA accumulation. Another study identified a significant association between multiple intronic SNPs in intron 7 of the TSH Receptor (*TSHR*) gene and Graves’ disease in a Japanese cohort (Hiratani *et al*., 2005). Previously SNPs located within intronic regions of the *ANXA2* gene were associated with sickle cell osteonecrosis (Baldwin *et al*., 2005). These studies have proven the relevance of investigating intronic SNPs in various candidate genes associated with disease phenotypes and pathology.

i) HWE Analysis and Significant Associations

The deviation from HWE at locus c.-12+75 C>A in the Black maternal patient cohort (0.0426) and c.244-76 C>G in the Mixed Ancestry fetal patient cohort (0.0000) could be explained by the population being stratified into smaller subgroups. Analysis of the subgroups allowed studying of genotype and allelic distribution between the different ethnic populations, which could contribute to our understanding of associations of variants present in specific populations with PE. The significantly reduced sample size however, affected the statistical power and could be the reason for deviation from HWE in some of the smaller subgroups. It is therefore recommended that genotyping of these SNPs should be replicated in larger Mixed Ancestry and Black populations.

A significant difference in the allele frequency between the Black maternal patient subgroup and control group (p=0.018) was identified at intronic SNP

c.244-76 C>G, since the minor allele G was present at a higher frequency (11.5%) in the patient cohort, compared to the control cohort (2.9%). Although this SNP was present in the Black population at a low frequency this SNP could represent a possible susceptibility to PE in the South African Black population.

The minor allele A of SNP c.449-17 G>A occurred more frequently in the control cohort (17.9%), compared to the Black fetal patient cohort (3.0%) and the Black maternal patient cohort (3.2%), suggesting a possible protective effect. This association suggests that the minor allele A could confer reduction in the risk of developing PE or play a protective role in the development of PE, which if investigated in different larger populations, could contribute to our knowledge of the underlying mechanism of PE.

The intronic SNP, c.589-5 C>T located within intron 8 revealed a significant difference in allele frequencies between all fetal patient and control cohorts (p=0.041). The complete absence of this SNP in the control cohort, compared to its frequency in the fetal patient cohort (3.9%), suggests a possible susceptibility to PE. Again the minor allele T was present at a very low frequency and along with the marginal association with PE it is recommended that the allele frequency of this SNP be determined in a larger population before a conclusion can be drawn about the role this SNP would play as a susceptibility allele to PE.

ii) *In silico* Analysis

In silico analysis of the SNP c.589-5 C>T, using the internet based programs Alternative Splice Site Predictor (ASSP) and NetGene2 as described in section 3.5.2, predicted no effect on alternative splicing. Predictions made by both these programs were contradictory, which suggests that programs for *in silico* analysis should be employed only to give an indication of the possible effects of variants on alternative splicing. Functional analysis using

approaches such as minigene assays are thus recommended to validate the predictions and effects of these SNPs (Baralle *et al.*, 2003).

4.1.4 Technique Evaluation

In an attempt to eliminate genotyping errors as the reason for departure from HWE, HEX-SSCP analysis was employed to confirm genotypes for the SNP c.-31 T>C observed by Multiphor SSCP/HD analysis. The results indicated in section 3.2.1 show that Multiphor SSCP/HD analysis was more sensitive than HEX-SSCP analysis, since conformational variants which were identified on the Multiphor gels and confirmed with bi-directional DNA sequencing, were not detected with the HEX-SSCP technique. Although Multiphor SSCP/HD analysis proved to be more sensitive, RFLP analysis using mutagenic primers and bi-directional DNA sequencing are recommended in order to ensure accurate and reliable genotype and allele frequencies for SNP c.-31 T>C in this study cohort.

Multiphor SSCP/HD analysis was used to screen most of the *ANXA2* gene, the only exception being the amplicon representing exon 10. Although Multiphor SSCP/HD analysis proved to be more sensitive than HEX-SSCP analysis, genotypes could not be discerned with confidence and the SNP identified by DNA sequencing was genotyped by RFLP analysis. It is therefore a possibility that other variants present in samples that were not subjected to DNA sequencing of this amplicon, were overlooked. In future it is thus recommended that this amplicon is screened for DNA variations using an alternative method of mutation detection.

In this study, conformational variants could be identified on the SSCP and HD components of Multiphor SSCP/HD gels, but the presence of more than one variant in an amplicon complicated the observation and clear visualisation of the different conformations on the gels. Consequently, conformations often needed to be confirmed with bi-directional DNA sequencing and RFLP analysis. Apart from direct bi-directional DNA sequencing, other screening

methods such as denaturing high-performance liquid chromatography (dHPLC) and high resolution melt (HRM) analysis are therefore recommended for future use (Fassano *et al*., 2005; Takano *et al*., 2008). These high-throughput techniques are ideal for screening large genes for novel DNA mutations (eg, SNPs, small deletions and insertions) with high sensitivity and efficiency.

4.1.5 Haplotype Analysis

Haplotype analysis, using the four gamete rule, revealed a significant level of LD between SNPs c.-11-43 G>A and c.-11-13 A>T in the Mixed Ancestry and Black maternal study cohorts (D' = 1.0, LOD > 43.56, r^2 = 1.0) and the Mixed Ancestry and Black fetal study cohorts (D'= 1.0, LOD > 40.19, r^2=1.0). These SNPs are located a mere 30 bp from each other within intron 1, which might explain the reason for the strong LD between them. The SNPs represent tagSNPs that could be used to genotype one SNP with the purpose of predicting the genotype of another SNP (or SNPs) in LD, allowing for effective high throughput genotyping (Sobrino *et al*., 2005). In future instead of two restriction enzyme digests, as was performed in this study, genotypes for one SNP could thus be deduced from genotyping the other SNP by a restriction enzyme digest.

Genotype frequencies at these loci (determined in these two South African population groups) were compared to genotype frequencies reported in Nigerian (YRI) family groups (HapMap). No LD could be demonstrated between SNPs c.-11-43 G>A and c.-11-13 A>T in the Nigerian population (D'=0.773, LOD=2.78, r^2=0.36). Although genotype and allele frequencies in different African populations are expected to be very similar, differences in populations of different geographical and ethnical heritage occur quite frequently as a result of the genetic diversity between African populations (Chakravarti, 2001).

Recurrent pregnancy loss (RPL) has previously been associated with haplotype M2, which comptrises of four SNPs (-19G/A, 1A/C, 27T/C, 76G/A) in the 5'UTR of the *ANXA5* gene (Bogdanova *et al.,* 2007). Recently, a study by Tiscia *et al*. (2009) confirmed the association of haplotype M2 with RPL, but furthermore, demonstrated that haplotype M2 is associated with gestational hypertension and PE. None of the other variants identified in this study were found to be in significant LD with each other. It is however possible that several of these SNPs may be in LD with variants in other genomic regions that were not screened in this study.

4.1.6 Low Frequency SNPs

The variants identified in this study are present at a very low frequency in both patient and control study cohorts. Possible reasons for the absence of homozygous genotypes (for the variant allele) could be because the minor alleles of these variants are present at such a low frequency, that it could not be detected by genotyping 120 pre-eclamptic maternal and 94 fetal individuals (Glatt *et al*., 2001). Sixteen of the 25 variants identified are novel variants that have not previously been characterized in the South African Black and Mixed Ancestry populations. Consequently, it could be hypothesized that these variants do not occur in other specific ethnic groups or that different susceptibility loci could be implicated in the risk of developing PE in different population groups.

4.1.7 Comparison of the Genotype and Allele Frequencies in Different Ethnic Populations

Population frequencies can be influenced by various factors. There are significant differences in allelic frequencies of variants in numerous genes in different cultural and ethnic groups (Kreek *et al.*, 2005). It is therefore important to consider ethnicity. The allelic frequencies obtained in this study population can therefore be compared to that of similar studies previously performed in different populations and ethnic groups.

The minor allele for the previously documented SNP c.-31 T>C (dbSNP: rs12904657) located within the 5'UTR of the *ANXA2* gene was present in the total maternal patient population at a frequency of 62.1% which correlates with the frequency observed in the Caucasian population (73.0%). Unfortunately no frequencies in the Asian and Sub-Saharan African populations were available and it could thus not be used to determine how similar these frequencies are.

As presented in Appendix H, Table 1, the genotype and allele frequencies for the three SNPs c.975 C>T; p.Gly325Gly, c.-11-43 G>A and c.48+67 C>T were present at a lower frequency in this study than in the European, Asian and Sub-Saharan African populations and SNP c.-11-13 A>T occurred at a higher frequency in this population compared to the other three populations. SNPs c.449-17 G>A and c.683-56 G>A were more prevalent in this study compared to the Sub-Saharan African population, but was present less often when compared to the European and Asian populations. Although a difference in the frequencies could be observed between the different populations, the frequencies of four of these SNPs (c.-11-13 G>A, c.48+67 C>T, c.449-17 G>A, c.683-56 G>A) in the two South African populations used in this study, was very similar to the frequencies observed in the Sub-Saharan populations. The correlation between these frequencies was expected, since the study participants represented Mixed Ancestry and Black women.

Genotype and allele frequencies for only one previously documented exonic SNP c.975 C>T; p.Gly325Gly, also identified in this study, were available on the Entrez SNP database, (http://www.ncbi.nlm.nih.gov/SNP/). The frequency of the variant allele T was found to be 56.7% in the European, 50.0% in the Asian and 44.2% in the Sub-Saharan African populations (Appendix H, Table 1). Interestingly, in the current study, the T allele was present in the total maternal patient and total control cohorts at a frequency of 0.8% and 0.6%, respectively. The frequency of the SNP c.975 C>T; p.Gly325Gly observed in the South African population groups used in this study was thus very low

compared to other populations. The variant allele was present in the Mixed Ancestry patient population at a frequency of 1.4% and 1.9% in the control cohort, but was completely absent from the Black population. A recent study by Coop *et al*. (2009) suggested that this would most likely not be as a result of selection based on geographical area, but that migration, population history and drift could be the cause of high allelic differences between populations.

4.2 Limitations of this Study

i) Sample Size

The relatively small population size of 120 maternal and 94 fetal patient individuals along with 54 control individuals, could explain the reason for the lack of homozygosity and the low frequencies of heterozygotes for a number of SNPs identified in this study. This limits detection of possible association between these polymorphisms and PE. The cohort size could therefore be increased in an attempt to increase the genotype and allele frequency.

Different population sizes were used in similar case-control association studies investigating candidate genes involved in PE. A cohort of 89 pre-eclamptic maternal patients and 349 control women was used in a study by Yu *et al*. (2006) that demonstrated an association between the Asp289Glu variant in the *eNOS* gene and PE, whereas a study by Goddard *et al*. (2007) employed the MassARRAY^{TM} system to genotype 775 SNPs in 190 candidate genes in a population comprising 394 pre-eclamptic maternal patients and 324 control women.

Although our cohort was not as large as those previously documented, the size of this population is sufficient to detect variants that are present at a high frequency. The identification of variants that occur at a lower frequency within the *ANXA2* gene, along with the identification of statistical significant associations between these variants and PE would require replication in an extended cohort.

The Mixed Ancestry maternal patient population (n=71) was considered as the focus group in this study and the Black patients (n=49) were included to increase the cohort. The control study cohort however consisted of only Mixed Ancestry individuals and it is therefore highly recommended that in future a control study cohort of ethnically matched, healthy pregnant women for the South African Black population is included to ensure more accurate comparisons in genotype and allele frequencies between patient and control groups.

ii) Multiple Testing

A further limitation of this study is the need for correction for multiple testing, since 25 markers were identified within the *ANXA2* gene in these two South African populations (Mixed Ancestry and Black). Analysing the data obtained in this study, (involving multiple comparisons) could affect the rejection of the H_0 hypothesis, along with confidence intervals. Repeated use of statistical tests is often responsible for an increased rate of type I errors. The false positive rate associated with multiple tests could be controlled by several tests and techniques. It is therefore recommended that multiple testing should be performed by employing Bonferroni corrections such as in a study by Unoki's *et al*. (2000), where statistical analysis of 33 SNPs were subjected to Bonferroni's corrections to ensure the correct conclusions on rejecting the H_0. In certain cases, where this technique is too conservative, the Post-hoc testing of ANOVA could be employed.

iii) Complexity of the Disease

PE is a multifaceted disease and currently no cure is available. Screening a candidate gene such as *ANXA2* contributes to insight and knowledge about the pathology of the disease. PE however, remains a complex disease and cannot be explained by variation(s) in a single gene. Multiple candidate genes should therefore be studied simultaneously to identify a possible association between genetic variants and the pathogenesis of PE. Applying a

technique that would enable high throughput genotyping could accelerate analysis of variants in numerous genes involved in healthy and complicated pregnancy. Multi-disciplinary approaches to analysing and investigating several factors involved in the disease process would broaden the spectrum of possible causes. The results of such a study could be useful in a clinical setting, as they may provide valuable insight into the mechanisms underlying PE.

4.3 Future Research and Conclusions

The research hypothesis for this study was to determine whether variants in the *ANXA2* gene confer susceptibility or protection in PE, and in doing so expand our knowledge and understanding of the pathogenesis of this complex disease.

The original aims for this study as stated in section 1.12 were successfully achieved;

- the *ANXA2* gene and 600 bp of the proximal promoter area were successfully characterised,
- sixteen of the 25 variants identified in the *ANXA2* gene in the South African population, were novel variants,
- the genotype and allele frequencies for these variants were identified in the Mixed Ancestry and South African Black study cohorts.
- significant association was identified between one exonic SNP (c.222 C>T; p.Ala74Ala), three intronic SNPs (c.244-76 C>G, c.449-17 G>A, c.589-5 C>T) and one SNP in the *ANXA2* 5'UTR (c.-92 A>T) and PE,
- the SNP c.-92 A>T located within the *ANXA2* 5'UTR demonstrated significant association with PE in the total maternal patient cohort (p=0.039) and Black fetal patient cohort (p=0.019),
- the exonic SNP (c.222 C>T; p.Ala74Ala) demonstrated a significant association with PE in the total maternal patient cohort (p=0.008), the

Mixed Ancestry maternal patient cohort (p=0.035)and total fetal patient cohort (p=0.019),

- significant association was identified between the following intronic SNPs and PE: SNP c.244-76 C>G demonstrated a significant association in the Black maternal patient cohort (p=0.018), SNP c.449-17 G>A demonstrated a significant association in the Black maternal patient cohort (p=0.032) and Black fetal patient cohort (p=0.030) and SNP c.589-5 C>T demonstrated a significant association in all the fetal cohorts (p=0.041),
- significant LD (D' = 1.0, LOD > 40.19, r^2 = 1.0) was demonstrated between two intronic SNPs (c.1-54 G>A and c.1-24 A>T) in the total maternal and fetal patient groups and subgroups,
- *in silico* analysis identified three SNPs, c.-191 G>C in the 5'UTR, c.934 G>A; p.Gly312Ser in exon 12 and c.292 G>T; p.Val98Leu in exon 5 that warrants experimental validation of the role these SNPs could play in the structure and function of annexin II in cellular and molecular processes.
- the three SNPs c.975 C>T; p.Gly325Gly, c.-11-43 G>A and c.48+67 C>T were present at a lower frequency in this study than in the European, Asian and Sub-Saharan African populations and SNP c.-11-13 A>T occurred at a higher frequency in this population compared to the other three populations.
- SNPs c.449-17 G>A and c.683-56 G>A were more prevalent in this study compared to the Sub-Saharan African population, but were less common compared to the European and Asian populations.

Although 25 variants were identified in the *ANXA2* gene, association could only be demonstrated at only five loci. The frequencies of these SNPs in the South African pre-eclamptic population were extremely low and although they would most likely not have a functional role in the pathology of PE, these SNPs could serve as susceptibility loci. Identifying SNPs associated with PE, remains important since it enhances the ability for early identification of individuals at risk of severe ealy onset PE.

This is the first study, to our knowledge, to screen the *ANXA2* gene in a pre-eclamptic population and by doing this a general view of the *ANXA2* gene on the molecular level was obtained. The findings in this study therefore contributed tremendously to our knowledge and understanding of the role *ANXA2* plays in healthy and pre-eclamptic pregnancies and set a good foundation for future studies.

Suggestions for future research would include the following:

i) In addition to screening the *ANXA2* gene in a larger cohort, this study should be replicated in other population groups to allow for comparison of genotype and allele frequencies.

ii) Conventional microarray technology would facilitate the study of genes that are differentially expressed in PE patients compared to healthy pregnant women. In a study by Reimer *et al*. (2002) placental tissue of PE patients was analysed for differential expression in multiple genes using DNA microarray analysis, revealing that obesity related genes, such as leptin, were up-regulated during PE. Although this approach has been employed before, *LGALS13* and *ANXA2* genes were not included in those studies. A similar study, including these specific genes, could therefore contribute to our understanding of their possible effect and function in the pathology of PE.

iii) The SNP c.-191 G>C identified in the *ANXA2* 5'UTR, should be further analysed in an attempt to determine its functional effect on the regulation of the *ANXA2* gene. Chromatin immunoprecipitation assays (ChiP) (Wells and Farnham, 2002) and electromobility shift assays (EMSA) (Prusty and Das, 2005) could be used to analyse and characterise the putative transcription factor binding sites predicted to be affected in the presence of the c.-191 G>C SNP.

iv) Functional analysis, employing minigene assays of SNPs c.934 G>A; p.Gly312Ser and c.292 G>T; p.Val98Leu should be performed to determine their possible functional effect on the annexin II protein (Baralle *et al.*, 2003). Structural modelling of annexin II and PP13 with and without these variants by X-ray crystallography could contribute to the understanding of how protein binding between annexin II and PP13 is affected by functional variants.

v) Annexin II levels should be measured in the sera of PE maternal patients and compared to levels in healthy pregnant mothers in the South African population. In Europe, PP13 levels are currently being measured using ELISA and is performed during the first trimester of pregnancy (Diagnostic Technologies Limited, Israel). This allows early prediction and identification of women who will subsequently develop PE, allowing early intervention and prophylaxis. Although this is currently not financially viable in South Africa, comparing measured annexin II levels with genotypes of SNPs identified in the *ANXA2* gene in pre-eclamptic patients and healthy pregnancies could ellucidate the association of variation in the *ANXA2* gene and PE. Significant differences in annexin II levels in PE patient and control individuals would contribute to our knowledge about the function of the *ANXA2* gene. Annexin II levels should furthermore be measured concurrently with PP13 levels and analysed in combination with genetic screening results of the *ANXA2* and *LGALS13* genes to expand the accuracy and efficiency of genotyping tests.

Healthy pregnancy involves a complex interplay of various bio-molecules to ensure normal development. The etiology underlying PE is still unknown, emphasizing the importance of identifying genes predisposing to this life threatening disease. Screening candidate genes, such as *ANXA2,* in combination with functional and biochemical analyses, could enhance our understanding of the underlying pathology of PE and contribute to the

development of an early biomarker. Early detection of PE would facilitate early intervention, hospital care and monitoring of pregnancies which would ultimately result in better pregnancy outcome and decreased mortality and morbidity rates.

5 REFERENCES

Aukrust I, Evensen L, Hollas H, Berven F, Atkinson RA, Travè G, Flatmark T, Vedeler A (2006) Engineering, biophysical characterisation and binding properties of a soluble mutant form of annexin A2 domain IV that adopts a partially folded conformation. *J Mol Biol* 363: 469-481

Bohn H, Kraus W, Winckler W (1983) Purification and characterisation of two new soluble placental tissue proteins (PP13 and PP17). *Oncodev Biol Med* 4:343 - 350

Baldwin C, Nolan VG, Wyszynski DF, Ma Q, Sebastiani P, Embury SH, Bisbee A, Farrell J, Farrer L, Steinberg MH (2005) Association of klotho, bone morphogenic protein 6, and annexin A2 polymorphisms with sickle cell osteonecrosis. *Blood* 106: 372-375

Baralle M, Baralle D, De Conti L, Mattocks C, Whittaker J, Knezevich A, ffrench-Constant C, Baralle FE (2003) Identification of a mutation that perturbs *NF1* gene splicing using genomic DNA samples and a minigene assay. *J Med Genet* 40(3): 220-222

Barett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21: 263-265

Barondes SH, Cooper DNW, Gitt MA, Leffer H (1994a) Galectins. Structure and function of a large family of animal lectins. *J Biol Chem* 269: 20807-20810

Barondes SH, Castronovo V, Cooper DNW, Cummings RD, Drickamer K, Feizi T, Gitt MA, Hirabayashi J, Hughes C, Kasai K, Leffler H, Li F-T, Lotan R, Mercurio AM, Monsigny M, Pillai S, Poirer F, Raz A, Rigby PW, Rini JM, Wang JL (1994b) Galectins: a family of animal β-galactoside-binding lectins. *Cell* 76: 597-598

Benirschke K, Kaufmann P (1991) Early development of the human placenta. In: Benirschkle K, Kaufmann P (eds) Pathology of the Human Placenta. Springer-Verlag. New York.

Bidon N, Brichory F, Bourguet P, Le Pennec JP, Dazord L (2001) Galectin-8: a complex sub-family of galectins (Review). *Int J Mol Med* 8: 245-250

Bird AP (1987) CpG islands as gene markers in the vertebrate nucleus. *Trends Genet* 3: 342-347

Bogdanova N, Horst J, Chlystun M, Croucher PJP, Nebel A, Bohring A, Todorova A, Schreiber S, Gerke V, Krawczak M, Markoff A (2007) A common haplotype of the *Annexin A5* (*ANXA5*) gene promoter is associated with recurrent pregnancy loss. *Hum Mol Gen* 16(5): 573-578

Bouwman P, Philipsen S (2002) Regulation of the acivity of Sp1-related transcription factors. *Mol Cel Endocrinol* 195: 27-38

Brooks ND, Grundy JE, Lavigne N, Derry MC, Restall CM, MacKenzie R, Waisman DM, Pryzdial LG (2002) Ca^{2+}-dependent and phospholipid-independent binding of *annexin 2* and *annexin 5*. *Biochem J* 367: 895-900

Buhimschi CS, Norwits ER, Funai E, Richman S, Guller S, Lockwood CJ, Buhimschi IA (2005) Urinary angiogenic factors cluster hypertensive disorders and identify women with severe pre-eclampsia. *Am J Obstet Gynecol* 192: 734-741

Burger O, Pick E, Zwickel J, Klayman M, Meiri H, Slotky R, Mandel S, Rabinovitch L, Paltieli Y, Admon A, Gonen R (2004) Placental protein 13 (PP-13): effects on cultured trophoblasts, and its detection in human body fluids in normal and pathological pregnancies. *Placenta* 25: 608-622

Burri PH, Hluschuk R, Djonov V (2004) Intussusceptive angiogenesis: its emergence its characteristics, and its significance. *Dev Dyn* 231: 474 -488

Caniggia I, Winter J, Lye SJ, Post M (2000) Oxygen and placental development during the first trimester: Implications for the pathophysiology of pre-eclampsia. *Placenta 21, Suppl A, Trophoblast Research*, 14: S25-S30

Cargill M, Altshuler D, Ireland J, Sklar P, Ardlie K, Patil N, Lane CR, Lim EP, Kalyanaraman N, Nemesh J, Ziaugra L, Friedland L, Rolfe A, Warrington J, Lipshutz

R, Daley GQ, Lander ES (1999) Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nature Genetics* 22: 231-238

Cartegni L, Chew SL, & Krainer AR (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet* 3(4): 285-298

Carty DM, Delles C, Dominiczak AF (2008) Novel biomarkers for predicting Preeclampsia. *Trends Cardiovasc Med* 18: 186-194

Cesarman-Maus G, Hajjar KA (2005) Molecular mechanisms of fibrinolysis. *Br J Haematol* 129: 307-321

Chafetz I, Kuhnreich I, Sammar M, Tal Y, Gibor Y, Meiri H, Cuckle H, Wolf M (2007) First-trimester placental protein 13 screening for preeclampsia and intrauterine growth restriction. *Am J Obstet Gynecol* 197: 35.e1-35.e7

Chakravarti A (2001) To a future of genetic medicine. *Nature* 409: 822-823

Chappell LC, Seed PT, Briley AL, Kelly FJ, Lee R, Hunt BJ, Parmar K, Bewley SJ, Shennan AH, Steer PJ, Poston L (1999) Effect of antioxidants on the occurrence of PE in women at increased risk: a randomised trial. *Lancet* 354: 810-816

Chappell S, Morgan L (2006) Searching for genetic clues to the causes of pre-eclampsia. *Clinical Science* 110: 443-458

Chikosi AB, Moodley J, Pegoraro RJ, Lanning PA, Rom L (1999) 5.10 methylenetatrahydrofolate reductase polymorphism in black South African women with pre-eclampsia. *Br J Obstet Gynecol* 106: 1219-1220

Clark DE, Smith SK, He Y, Day KA, Licence DR, Corps AN, Lammoglia R, Charnock-Jones DS (1998) A vascular endothelial growth factor antagonist is produced by the human placenta and released into the maternal circulation. *Biol Reprod* 59: 1540-1548

Collins FS, Guyer MS, Chakravarti A (1997) Variations on a Theme: Cataloging Human DNA Sequence Variation. Science 278(5343): 1580-1581

Conde-Agudelo A, Belizan JM (2000) Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. *Br J Obstet Gynecol* 107(1): 75-83

Coop G, Pickrell JK, Novembre J, Kudaravalli S, Li J, Absher D, Myers RM, Cavalli-Sforza LL, Feldman MW, Pritchard JK (2009) The role of geography in human adaptation. P*LoS Genet* 5(6): 1-16

Cooper DN, Youssoufian H (1987) The CpG dinucleotide and human genetic disease. *Hum Genet* 78(2): 151-155

Creutz CE (1992) The annexins and exocytosis. *Science* 258: 924-931

Cross JC (2003) The genetics of pre-eclampsia: a feto-placental or maternal problem? *Clin Genet* 64 : 96-103

Davey DA, MacGillivray I (1988) The classification and definition of the hypertensive disorders of pregnancy. *Am J Obstet Gynecol* 158: 892-898

Dekker G, Sibai B (2001) Primary, secondary and tertiary prevention of pre-eclampsia. *Lancet* 357 : 209-215

Department of Health National Committee on Confidential Enquiries into Maternal Deaths (NCCEMD) Saving Mothers: Fourth report on confidential enquiries into maternal deaths in South Africa 2005-2007

Dizon-Townson DS, Nelson LM, Easton K, Ward K (1996) The factor V Leiden mutation may predispose women to severe preeclampsia. *Am J Obstet Gynecol* 175: 902-905

Duret L (2009) Mutation patterns in the human genome: more variable than expected. *PLOS Biol* 7(2): 217-219

Edberg JC, Wu J, Langeveld CD, Brown EE, Marion MC, McGwin G, Petri M, Ramsay-Glodman R, Reveille JD, Frank SG, Kaufman KM, Harley JB, Alarcon GS, Kimberley RP (2008) Genetic variation in the CRP promoter: association with systemic lupus erythematosus. *Hum Mol Genet* 17(8): 1147-1155

Empson M, Lassere M, Craig JC, Scott JR (2002) Recurrent pregnancy loss with antiphospholipid antibody:a systemic review of therapeutic trials. *Obstet Gynecol* 99: 135-144

Fassano T, Bocchi L, Pisciotta L, Bertolini S, Calandra S (2005) Denaturing high-performance liquid chromatography (DHPLC) in the detection of mutations of ABCA1 gene in familial HDL deficiency. *J of Lipid Research* 46: 817-822

Fleury V, Loyau S, Lijnen HR, Nieuwenhuizen W Angles-Cano E (1993) Molecular assembly of plasminogen and tissue-type plasminogen activator on an evolving fibrin surface. *Eur J Biochem* 216: 549-556

Gerke V, Moss SE (1997) Annexins and membrane dynamics. *Biochim Biophys Acta* 1357: 129-154

Gerke V, Moss SE (2002) Annexins: From structure to function. *Physiol Rev* 82: 331-371

Glatt CH, DeYoung JA, Delgado S, Service SK, Giacomini KM, Edwards RH, Risch N, Freimer NB (2001) Sreening a large reference sample to identify very low frequency sequence variants: comparisons between two genes. *Nat Genet* 27: 435-438

Glenney JR, Tack B, Powell MA (1987) Calectins: Two distinct Ca^{2+} regulated phospholipid and actin binding proteins isolated from lung and placenta. *J Cell Biol* 104: 503-511

Goddard KAB, Tromp G, Romero R, Olson JM, Lu Q, Xu Z, Parimi N, Nien JK, Gomez R, Behnke E, Solari M, Espinoza J, Santolaya J, Chaiworapongsa T, Lenk GM, Volkenant K, Anant MK, Salisbury BA, Carr J, Lee MS, Vovis GF, Kuivaniemi H (2007) Candidate-gene association study of mothers with pre-eclampsia, and their infants, analysing 775 SNPs in 190 genes. *Hum Hered* 63: 1-16

Grandone E, Margaglione M, Colaizzo D, Cappucci G, Paladini D, Martinelli P, Montanaro S, Pavone G, Di Minno G (1997) Factor V Leiden, C>T MTHFR

polymorphism and genetic susceptibility to pre-eclampsia. *Thromb Haemost* 77: 1052-1054

Haga H, Yamada R, Ohnishi Y, Nakamura Y, Tanaka T (2002) Gene-based SNP discovery as part of the Japanese Millennium Genome Project: identification of 190 562 genetic cariations in the human genome. *J Hum Genet* 47: 605-610

Hall DR, Grové D, Carstens E (2006) Early pre-eclampsia: what proportion of women qualify for expectant management and if not, why not? *Eur J Obstet Gynecol Reprod Biol* 128: 169-174

Hall DR, Odendaal HJ, Smith M (2000) Is prophylactic administration of magnesium sulphate in women with pre-eclampsia indicated prior to labour? *Br J Obstet Gynecol* 107: 903-908

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41: 95-98

Hartl DL, Clark AG (1997) Principles of Population Genetics (eds3) Sunderland, Massachusetts: Sinaur Associates, Inc Publishers.

Hassan MI, Aschner Y, Manning CH, Xu J, Aschner JL (2003) Racial differences in selected cytokine allelic and genotypic frequencies among healthy, pregnant women in North Carolina. *Cytokine* 21: 10-16

Hillermann R, Carelse K, Gebhardt GS (2005) The Glu298Asp variant of the endothelial nitric oxide synthase gene increases the risk of abruptio placentae in pre-eclampsia. *J Hum Genet* 81: 415-419

Hiratani H, Bowden DW, Ikegami S, Shirasawa S, Shimizu A, Iwatani Y, Akamizu T (2005) Multiple SNPs in Intron 7 of Thyrotropin Receptor Are Associated with Graves' Disease. *J Clin Endocrinol Metab* 90(5): 2898-2903

Huang KS, Wallenr BP, Mattaliano RJ, Tizard R, Burne C, Frey A, Hession C, McGray P, Sinclair LK, Chow EP, Browning JL, Ramachandran KL, Tang J, Smart JE, Pepinsky RP (1986) Two human 35 kd inhibitors of phospholipase A2 are related

to substrates of pp60(v-src) and of the epidermal growth factor receptor/kinase. *Cell* 46: 191-199

Huebner K, Cannizzaro LA, Frey AZ, Hecht BK, Hecht F, Croce CM, Wallner B (1988) Chromosomal localization of the human genes for lipocortin I and lipocortin II. *Oncogene Res* 2: 99-310

Hughes AL, Packer B, Welch R, Bergen AW, Chanock SJ, Yeager M (2003) Widespread purifying selection at polymorphic sites in human protein-coding loci. *Proc Natl Acad Sci USA* 100: 15754-15757

Hughes RC (1999) Secretion of the galectin family of mammalian carbohydrate-binding proteins. (Review) *Biochim Biophys Acta* 1473: 172-185

Hughes RC (2001) Galectins as modulators of cell adhesion. *Biochimie* 83: 667-676

Iwakuma T, Lozano G, Flores ER (2005) Li-Fraumeni syndrome: a p53 family affair. *Cell Cylce* 4(7): 865-867

Jonsdottir LS, Arngrimsson R, Geirsson RT, Sigvaldason H, Sigfusson N (1995) Death rates from ischeamic heart disease in women with a history of hypertension in pregnancy. *Acta Obstet Gynecol Scand* 74(10): 772-776

Kaczan-Bourgois D, Salles JP, Hullin F, Fauvel J, Moisand A, Duga-Neulat I, Berrebi A, Campistron G and Chap H (1996) Increased content of annexin II (p36) and p11 in human placenta brush-border membrane vesicles during syncytiotrophoblast maturation and differentiation. *Placenta* 17: 669-676

Kaiser T, Brennecke SP, Moses EK (2000) Methylenetetrahydrofolate reductase polymorphisms are not a risk factor for pre-eclampsia/eclampsia in Australian women. *Gynecol Obstet Invest* 50: 100-102

Kenneth L, Hall DR, Gebhardt S, Grové (2008) Late onset pre-eclampsia in not an innocuous condition? (Accepted for publication in *Hypertension in Pregnancy* December 2008)

Kim YM, Bujold E, Chaiworapongsa T, Gomez R, Yoon B H, Thaler HT, Rotmensch S, Romero R (2003) Failure of physiologic transformation of the spiral arteries in patients with preterm labor and intact membranes. *Am J Obstet Gynecol* 189(4) 1063-1069

Kotze MJ, Theart L, Callis M, Peeters A, Thiart R, Langenhoven E (1995) Nonradioactive multiplex PCR screening strategy for the simultaneous detection of multiple low-density lipoprotein receptor gene mutations. *PCR Methods Applic* 4: 352-356

Kraulis PJ (1991) MOLSCRIPT; A program to produce both detailed and schematic plots of protein structures. *J Appl Cryst* 24: 946-950

Kreek MJ, Nielsen DA, Butelman ER, LaForge KS (2005) Genetic influences on impulsivity, risk taking, stress responsivity and vulnerability to drug abuse and addiction. *Nat Neurosci* 8 (11): 1450-1457

Lewit-Bentley A, Rety S, Sopkova-de Oliveira Santos J, Gerke V (2000) S100–annexin complexes: some insights from structural studies. *Cell Biol Int* 24 : 799 – 802

Liechti-Gallati S, Schneider V, Neeser D, Kraemer R (1999) Two buffer PAGE system-based SSCP/HD analysis: a general protocol for rapid and sensitive mutation screening in cystic fibrosis and any other human genetic disease. *Eur J Hum Genet* 7: 590-598

Lobsanov YD, Rini JM (1997) Galectin structure. *Trends Glycosci Glycotechnol* 9: 145-154

Lorkowski S, Cullen P (2003) Analysing Gene Expression: A Handbook of Methods and Pitfalls (Vol1) WILEY-VCH Verlag GmbH &Co. KGaA, Weinheim

MacLeod TJ, Known M, Filipenko NR, Waisman DM (2003) Phospholipid-associated annexin A2-S100A10 heterotetramer and its subunits: Characterization of the interaction with tissue plasminogen activator, plasminogen, and plasmin. *J Biol Chem* 278: 25577-25584

Mann JI, Doll R, Thorogood M, Vessey MP, Waters WE (1976) Risk factors for myocardial infarction in young women. *Br J Prev Soc Med* 30(2): 94-100
Martini FH, Bartholomew EF (2003) Development and Inheritance In: Essentials of anatomy and physiology (eds3) Prentice Hall, Pearson Education, Inc. New Jersey.

Maynard SE, Min JY, Merchan J, Lim K, Li J, Mondal S, Libermann TA, Moragn JP, Sellke FW, Stillman IE, Epstein FH, Sukhatme VP, Karumanchi SA (2003) Excess placental soluble fms-like tyrosine kinase 1 (sFlt1) may contribute to endothelial dysfunction, hypertension, and proteinuria in preeclampsia. *J Clin Invest* 111: 649-658

Miller M (1997) Tools for population genetic analyses (TFGA) 1.3: a Windows program for the analysis of allozyme and molecular population genetic data. Computer software distributed by the author.

Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 16 : 1215-1215

Morgan T, Craven C, Lalouel JM, Ward K (1999) Angiotensinogen Thr235 variant is associated with abnormal physiologic change of the uterine spiral arteries in first-trimester decidua. *Am J Obstet Gynecol* 180: 95-102

Moss SE, Morgan RO (2004) The annexins. *Genome Biol* 5: 219.1-219.8

Naylor S (2003) Biomarkers: current perspectives and future prospects. *Expert Rev Mol Diagn* 3: 525-529

Nichols AF, Itoh T, Zolezzi F, Hutsell S, Linn S (2003) Basal transcriptional regulation of human damage-specific DNA-binding protein genes DDB1 and DDB2 by Sp1, E2F, N-myc and NF1 elements. *Nucleic Acids Res* 31: 562-569

Nicolaides KH, Bindra R, Turan OM, Chefetz I, Sammar M, Meiri H, Tal J, Chuckle HS (2006) A novel approach to first-trimester screening for early pre-eclampsia combining serum PP-13 and Doppler ultrasound. *Ultrasound Obstet Gynecol* 27: 13-17

Norris M, Perico N, Remuzzi G (2005) Mechanisms of disease: pre-eclampsia. *Nat Clin Prac Nephrol* 1: 98-114

Nott A**,** Meislin SH, Moore MJ (2003) A quantitative analysis of intron effects on mammalian gene expression. *RNA* 9: 607-617

Orsi NM (2007) Genetics of preterm labour. *Best Pract Res Clin Obstet Gynaecol* 21(5): 757-772

Ozturk O, Bhattacharya S, Saridogan E Jauniaux **E**, Templeton A (2004) Role of utero-ovarian vascular impedance: predictor of ongoing pregnancy in an IVF-embryo transfer programme. *Reprod Biomed Online* 9: 299-305

Pagani F, Baralle FE (2004) Genomic variants in exons and introns: identifying the splicing spoilers. *Nat Rev Genet* 5: 389-396

Paull TT, Cortez D, Bowers B, Elledge SJ, Gellert M (2001) From the Cover: Direct DNA binding by BRCA1. *Proc Natl Acad Sci USA* 98: 6086-6091

Perillo NL, Marcus ME, Baum LG (1998) Galectins: versatile modulators of cell adhesion, cell proliferation, and cell death. *J Mol Med* 76: 402-412

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. *Nat Rev Genet* 6: 847-846

Poston L, Briley AL, Seed PT, Kelly FJ, Shennan AH (2006) Vitamin C and vitamin E in pregnant women at risk for pre-eclampsia (VIP trial): randomized placebo-controlled trial. *Lancet* 367: 1145-1154

Prusty PK, Das BC (2005) Constitutive activation of transcription factor AP-1 in cervical cancer and suppression of human papillomavirus (HPV) transcription and AP-1 activity in HeLa cells by curcumin. *Int J Cancer* 113: 951-960

Rajkovic A, Mahomed K, Rozen R, Malinow MR, King IB, Williams MA (2000) Methylenetetrahydrofolate reductase. *Mol Genet Metab* 69: 33-39

Rand JH, Wu XX (1999) Antibody-mediated disruption of the annexin V antithrombotic shied: a new mechanism for thrombosis in the antiphospholipid syndrome. *Thromb Haemost* 82: 649-655

Rand JH, Wu XX, Guller S, Gil J, Guha A, Scher J, Lockwood CJ (1994) Reduction of annexin V (placental anticoagulant protein-1) on placental villi of women with antiphospholipid antibodies and recurrent spontaneous abortion. *Am J Obstet Gynecol* 171: 1566-1572

Redman CW, Sargent IL (2005) Latest advances in understanding preeclampsia. *Science* 308: 1592-1594

Reimer T, Koczan D, Gerber B, Richter D, Thiesen HJ, Friese K (2002) Microarray analysis of differentially expressed genes in placental tissue of pre-eclampsia**:** up-regulation of obesity-related genes. *Mol Hum Reprod* 8(7): 674-680

Remmers EF, Plenge RM, Lee AT, Graham RR, Hom G, Behrens TW, de Bakker PIW, Le JM, Lee H, Batliwalla F, Li W, Masters SL, Booty MG, Carulli JP, Padyukov L, Alfredsson L, Klareskog L, Chen WV, Amos CI, Criswell LA, Seldin MF, Kastner DL, Gregerse PK (2007) *STAT4* and the Risk of Rheumatoid Arthritis and Systemic Lupus Erythematosus. *N Engl J Med* 357: 977-986

Rety S, Sopkova J, Renourd M, Osterloh D, Gerke V, Tabaries S, Russo-Marie F, Lewit-Bentley A (1999) The crystal structure of a complex of p11 with the annexin II N-terminal peptide. *Nat Struct Biol* 6: 89-95

Roberts JM, Cooper DW (2001) Pathogenesis and genetics of pre-eclampsia. *Lancet* 357: 53-56

Rodie VA, Freeman DJ, Sattat N, Greer IA (2004) Pre-eclampsia and cardiovascular disease: metabolic syndrome of pregnancy? *Artherosclerosis* 175: 189-202

Sachs AB, Sarnow P, Hentze MW (1997) Starting at the beginning, middle and end: Translation initiation in eukaryotes*. Cell* 89: 831-838

Schaid DJ, Jacobsen SJ (1999) Biased tests of association: Comparison of allele freqeuncies when departing from Hardy-Weinberg proportions. *Am J Epidemiol* 149(8): 706-711

Serrano NC (2006) Immunology and genetics of preeclampsia. *Clin Dev Immunol* 13(2-4): 197-201

Serrano NC, Casas JP, Dı´az LA, Pa´ez C, Mesa CM, Cifuentes R, Monterrosa A, Bautista A, Hawe E, Hingorani AD, Vallance P, Lo´pez-Jaramillo P (2004) Endothelial NO Synthase Genotype and Risk of Preeclampsia: A Multicenter Case-Control Study. *Hypertension* 44: 702-707

Shankar R, Gude N, Cullinanel F, Brenneckel S, Purcell AW, Moses EK (2005) An emerging role for comprehensive proteome analysis in human pregnancy research. *Reproduction* 129: 685-696

Shu F, Sugimura M, Kanayama N, Kobayashi H, Terao T (2000) Immunohistochemical study of annexin V expression in placentae of preeclampsia. *Gynecol Obstet Invest* 49: 17-23

Sibai B, Dekker G, Kupferminc M (2005) Pre-eclampsia Seminar. *Lancet* 365: 785-799

Sibai BM, Caritis SN, Thom E, Klebanoff M, McNellis D, Rocco L, Paul RH, Romero R, Witter F, Rosen M, Depp R (1993) National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units. Prevention of pre-eclampsia with low-dose aspirin in healthy, nulliparous pregnant women *N Engl J Med* 329: 1213-1218

Sibai BM, Gordon T, Thom E, Caritis SN, Klebanoff M, McNellis D, Paul RH, and the National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units (1995) Risk factors for pre-eclampsia in healthy nulliparous women: A prospective multicenter study. *Am J Obstet Gynecol* 172 (2): 642-648

Skalsky RL, Samols MA, Plaisance KB, Boss IW, Riva A, Lopez MC, Baker HV, Renne R (2007) Kaposi's Sarcoma-Associated Herpesvirus encodes and ortholog of miR-155. *J Virol* 81(23): 12836-12845

Skjaerven R, Vatten LJ, Wilcox AJ, Ronning T, Irgens LM, Lie RT (2005) Recurrence of pre-eclampsia across generations: Exploring fetal and maternal genetic components in a population based cohort. *Br Med J* 331(7521): 877-882

Smith PJ, Zhang C, Wang J, Chew S L, Zhang MQ, Krainer AR (2006) An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet* 15: 2490-2508

Sobrino B, Brión M, Carracedo A (2005) SNPs in forensic genetics: a review on SNP typing methodologies. *Forensic Sci Int* 154:181-194

Sohda S, Arinami T, Hamada H, Yamada N, Hamaguchi H, Kubo T (1997) Methylenetetrahydrofolate reductase polymorphism and pre-eclampsia. *J Med Genet* 34: 525-526

Spano F, Raugei G, Palla E, Colella C, Melli M (1990) Characterization of the human lipocortin-2-encoding multigene family: its structure suggests the existence of a short amino acid unit undergoing duplication. *Gene* 95: 243-251

Spencer K, Cowans NJ, Chefetz I, Tal J, Meiri H (2007) First-trimester maternal serum PP-13, PAPP-A and second-trimester uterine artery Doppler pulsatility index as markers of pre-eclampsia. *Ultrasound Obstet Gynecol* 29: 128-134

Suske G (1999) The Sp-family of transcription factors. *Gene* 238: 291-300

Taberlet, P, Griffin S, Goossens B, Qestiau S, Manceau V, Ecaravage N, Waits LP, Bouvet J (1996) Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Res* 24: 3189–3194

Takano EA, Mitchell G, Fix SB, Dobrovich A (2008) Rapid detection of carriers with BRCA1 and BRCA2 mutations using high resolution melting analysis. *BMC Cancer* 8: 59

Teasedale F (1985) Histomorphometry of the human placenta in maternal preeclampsia. *Am J Obstet Gynecol* 152: 25-31

Than NG, Pick E, Bellyei S, Szigeti A, Burger O, Berente Z, Janaky T, Boronkai A, Kliman H, Meiri H, Bohn H, Than GN, Sumegi B (2004) Functional analyses of placental protein 13/galectin-13. *Eur J Biochem* 271: 1065-1078

Than NG, Sumegi B, Than GN, Berente Z, Bohn H (1999) Isolation of sequence analysis of a cDNA encoding human placental tissue protein 13 (PP13), a New Lysophospholipase, Homologue of Human Eosinophil Charcot-Leyden Crystal Protein. *Placenta* 20: 703-710

Thiel C, Osborn M, Gerke V (1993) The tight association of the tyrosine kinase substrate annexin II with the submembranous cytoskeleton depends on intact p11- and Ca (2+)-binding sites. *J Cell Sci* 103 (3): 733-742

Tiscia G, Colaizzo D, Chinni E, Pisanelli D, Scianname N, Favuzzi G, Marqaqione M, Grandone E (2009) Haplotype M2 in the Annecin A5 (ANXA5) gene and the occurrence of obstetric complication. *Thromb Haemost* 102(2): 309-313

Tubbergen P, Lachmeijer AMA, Althuisius SM, Vlak MEJ, van Geijn HP, Dekker GA (1999) Change in paternity: a risk factor for PE in multiparous women? *J Reprod Immunol* 45: 81-88

Unoki M, Furuat S, Onouchi Y, Watanabe O, Doi S, Fujiwara H,Miyatake A, Fujita K, Tamari M, Nakamura Y (2000) Association studies of 33 single nucleotide polymorphisms (SNPs) in 29 candidate genes for bronchial asthma: positive association a T924C polymorphism in thetromboxane A2 receptor gene. *Hum Genet* 106: 440-446

Verwoerd GR, Hall DR, Grove D, Maritz JS, Odendaal HJ (2002) Primipaternity and duration of exposure to sperm antigens as risk factors for pre-eclampsia. *Int J Gynaecol Obstet* 78: 121-126

Villar J, Say L, Shennan A, Lindheimer M, Duley L, Conde-Agudelo A, Merialdi M (2004) Methodological and technical issues related to the diagnosis, screening,

prevention and treatment of pre-eclampsia and eclampsia. *Int J Genaecolo Obstet* 85: S28-S41

Visegrady N, Than NG, Kilar F. Sumegi B, Than GN, Bohn H (2001) Homology modeling and molecular dynamics studies of human placental tissue protein 13 (galectin-13). *Prot Eng* 14: 875-880

Volinia S, Calin GA, Liu C, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris C, Croce CM (2005) A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 103(7): 2257-2261

Waisman DM (1995) Annexin II tetramer: structure and function. *Mol Cell Biochem* 149: 301-322

Wallenburg HC, Dekker GA, Makovitz JW, Rotmans P (1986) Low-dose aspirin prevents pregnancy-induced hypertension and pre-eclampsia in angiotensin-sensitive primigavidae. *Lancet* 1: 1-3

Wang J, Pitarque M, Ingelman-Sundberg M (2005) 3'UTR ploymorphism in the human CYP2A6 gene affects mRNA stability and enzyme expression. *Biochem Biophys Res Commun* 340(2): 491-497

Wang J, Smith PJ, Krainer AR, Zhang MQ (2005) Distribution of SR protein exonic splicing enhancer motifs in human protein-coding genes. *Nucleic Acids Res* 33(16): 5053-5062

Wang N, Akey JM, 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. *Am J Hum Gen* 71: 1227-1234

Ward K, Hata A, Jeunemaitre X, Helin C, Nelson L, Namikawa C, Farrington PF, Ogasawara M, Suzumori K, Tomoda S, Berrebi S, Sasaki M, Corvol P, Lifton RP, Lalouel JM (1993) A molecular variant in angiotensinogen associated with pre-eclampsia. *Nat Genet* 4: 59-61

Wells J, Farnham PJ (2002) Characterising transcription factor binding sites using formaldehyde crosslinking and immunoprecipitation. *Methods* 26: 48-56

Wu Y, Zhang Y, Zhang J (2005) Distribution of exonic splicing enhancer elements in human genes. *Genomics* 86: 329-336

Ye Q, Hu YF, Zhong H, Nye AC, Belmont AS, Li R (2001) BRCA1-induced large scale chromatin unfolding and allele-specific effects of cancer-predisposing mutations. *J Cell Biol* 155: 911-921

Yu CKH, Casas JP, Savvidou MD, Sahamey MK, Nicolaides KH, Hingorani AD (2006) Endothelial nitric oxide synthase gene polymorphism (Glu298Asp) and development of PE: a case-control study and a meta-analyses. *BMC Pregnancy Childbirth* 6(7): 1-9

Yu CKH, Papageorghiou AT, Parra M, Dias RP, Nicolaides KH (2003) Randomized controlled trial using low-dose aspirin in the prevention of pre-eclampsia in women with abnormal uterine artery Doppler at 23 weeks' gestation. *Ultrasound Obstet Gynecol* 22: 233-39

Zhao J, Ennion SJ (2006) Sp1/3 and NF-1 mediate basal transcription of the human $P2X_1$ gene in megakaryoblastic MEG-01 cells. *BMC Mol Biol* 7(10): 1-11

Electronic Resources

AliBaba2.1,http://www.gene-regulation.com/pub/programs/alibaba2/index.html

ASSP, http://www.es.embnet.org/~wang/asp.html

Department of Health, www.doh.gov.za

Entrez Nucleotide Database, http://www.ncbi.nlm.nih.gov/nucleotide/

Entrez SNP Database, http://www.ncbi.nlm.nih.gov/SNP/

ESE Finder, http://rulai.cshl.edu/tools/ESE/

GNF SymAtlas, http://biogps.gnf.org

HapMap: http://www.hapmap.org

Human Genome Committee (HUGO),http://www.genenames.org/guidelines.html

Human Genome Variation Society (HGVS), http://www.hgvs.org/mutnomen/

In silico Molecular Biology Experiments, http://insilico.ehu.es/restriction/

Integrated DNA Technologies, www.idtdna.com

JASPAR CORE, http://jaspar.genereg.net/

MatInspector,http://www.genomatrix.de/cgi-bin/matinspector_prof/

MMB, http://mmb.pcb.ub.es/

NCBI Basic Local Alignment Search Tool, http://ww.ncbi.nlm.nih.gov/BLAST/

NetGene2, http://www.cbs.dtu.dk/services/NetGene2

PMUT, http://mmb2.pcb.ub.es:8080/PMut/

RestrictionMapper, www.restrictionmapper.org

SIFT, http://blocks.fhcrc.org/sift/SIFT/html

TESS, http://www.cbil.upenn.edu/cgi-bin/tess

6 APPENDICES

Appendix A: Informed Consent Forms

FACULTY OF HEALTH SCIENCES
STELLENBOSCH UNIVERSITY

INFORMATION AND INFORMED CONSENT DOCUMENT FOR DNA ANALYSIS AND STORAGE

TITLE OF THE RESEARCH PROJECT: ***Profiling placental protein 13 / galectin-13 in pre-eclampsia.***

REFERENCE NUMBER: ……..……..…….

PRINCIPAL INVESTIGATOR: Drs R Hillermann & GS Gebhardt

Address: Departments of Obstetrics and Gynaecology and Genetics, University of Stellenbosch, Tygerberg and Paarl Hospital

DECLARATION BY OR ON BEHALF OF PARTICIPANT:

I, THE UNDERSIGNED,

……..…….. *(name)*

[ID No: ……..……] of ……..……

……..……

…… *(address).*

A. HEREBY CONFIRM AS FOLLOWS:

1. I was invited to participate in the above-mentioned research project of the Departments of Obstetrics and Gynaecology and Genetics, Faculty of Health Sciences, University of Stellenbosch.
2. The following aspects have been explained to me:

   Aim: Pre-eclampsia (high blood pressure in pregnancy) is a severe complication of pregnancy and it affects about 5% of all

pregnancies. Currently there is no way to predict who will be affected. We are investigating a protein normally secreted by the placenta (the afterbirth) as it seems that this protein may predict the development of later problems.

2.1 **Procedures:** I will be requested to provide information about my medical history. Blood (10*ml*, about 2 teaspoonfuls) will be collected from me at delivery of my baby together with the routine blood samples taken at birth. In addition 6 small biopsies (pea-size) will be taken from the afterbirth (the placenta) after delivery and before the placenta is routinely destroyed.

2.3 **Genetic considerations**

- The DNA may be stored for several years until the technology for meaningful analysis becomes available;
- The clinically relevant results of the (possible) analyses carried out on this material in the current study can be made known to me at my request and in accordance with the relevant protocol, if and when it becomes available. In addition, I authorise(s) the investigator(s) to make the information available to ……*(doctor's name),* the doctor involved in my care, as well as to the following family members ……
  ……
  …… *(names);*
- The DNA will be maintained indefinitely, unless I request to have it and/*or the stored clinical data destroyed by contacting the investigator conducting the present study, dr GS Gebhardt at 938 9131 or the Chairperson of the Research Subcommittee C/Ethics Committee at 9389111 if the former cannot be located;

- The analyses in the current study are specific to the condition or disease mentioned above and cannot determine the entire genetic make-up of an individual;
- Genetic analyses may not be successful in revealing additional information regarding some families or some family members;
- Even under the best conditions, current technology of this type is not perfect and could lead to unreliable results.

2.4 **Confidentiality:** My identity will be kept confidential throughout. Information will not be associated with my name. The research staff will use only a coded number, access will be limited to authorized scientists and any scientific publications, lectures or reports resulting from the study will not identify me by name.

2.5 **Voluntary participation: Participation** is voluntary and I may decline participation, or withdraw from the study at any time without any loss of benefits to which I am otherwise entitled. Future management at this or any other institution will not be compromised by refusal or withdrawal.

2.6 **Risks:** There are no more than minimal medical or psychological risks associated with this study:

- I may feel some pain associated with having blood withdrawn from a vein and may experience discomfort, bruising and/or slight bleeding at the site;
- Biopsies of the placenta is not painful as it is taken after delivery before routine destruction of the placenta
- As some insurance companies may mistakenly assume that my' participation is an indication of a higher risk of a genetic disease which could hurt my access to health or other insurance, no information about me or my family will be shared with such companies as this investigation

cannot be regarded as formal genetic testing for the presence or absence of certain genes.

2.7 **Benefits:**

- Although there may not be any direct benefits to me by participating at this stage, family members and future generations may benefit if the researchers succeed in scientifically delineating certain disorders further. Thereby the rational approach to the clinical diagnosis and therapy of its manifestations may be facilitated. The identification and location of the genes involved in such disorders, could in the end lead to the development of methods for prevention and to forms of new treatment aimed at curing or alleviating these conditions;
- In the unlikely event that the research may lead to the development of commercial applications, I or my heirs will not receive any compensation, but profits will be reinvested into supporting the cause of further research which may bring benefits to my family and to the community, such as health screening, medical treatment, educational promotions, etc;

2.8 **Permission for further studies:** Before my material is used in further projects in the future, the written approval of the Research Subcommittee C/Ethics Committee, Faculty of Health Sciences, will be obtained.

3. The information conveyed above was explained to me by ……..……..…….……..…… ….. *(name)* in English and I am fluent in this language

4. I was afforded adequate time to pose any questions and all questions were answered to my full satisfaction.

5. I was not pressurized to participate.

6. I will not be paid for participation, but reimbursement of travel costs will be considered (if applicable).

7. I will not incur any additional costs through participation.

8. I have/*has received a copy of this document for my records.

9. The Research Subcommittee C/Ethics Committee, Faculty of Health Sciences, Stellenbosch University, has approved recruitment and participation of individuals in this study on the basis of:
   - Guidelines on Ethics for Medical Research of the SA Medical Research Council;
   - Declaration of Helsinki;
   - International Guidelines : Council for International Organisations of Medical Sciences (CIOMS);
   - Applicable RSA legislation.

I. I HEREBY CONSENT VOLUNTARILY TO PARTICIPATE/*ALLOW THE POTENTIAL PARTICIPANT TO PARTICIPATE IN THIS STUDY:

Signed/*Confirmed at ……..……..…… on……...20……
(place *(date)*

……
*Signature or right thumb print of participant/*representative of participant* *Signature of witness*

DECLARATION BY OR ON BEHALF OF INVESTIGATOR(S):

I, ……..…*(name)*

declare that

- I explained the information in this document to ……. *(name of the patient/*participant)* and/or his/her representative …… *(name of the representative);*
- she/*he was encouraged and afforded adequate time to ask me any questions;
- this conversation was conducted in Afrikaans/*English/*Xhosa/*Other …….. and no translator was used/*was translated into ……*(language)* by ……..……..… *(name)*.

Signed at …….……. on ……..20….
(*place)* *(date)*

…….…… ……
Signature of investigator/representative of investigator *Signature of witness*

…… ……
Signature *of translator* *Signature of witness*

IMPORTANT MESSAGE TO PARTICIPANT/*REPRESENTATIVE OF PARTICIPANT:

Dear participant/*representative of participant,

Thank you very much for your participation in this study. Should, at any time during the study,

- an emergency arise as a result of the research, or
- you require any further information with regard to the study, kindly contact Dr GS Gebhardt at 938 0131, 938 4707 (after hours) or 87 21 711 (all hours) or come to the labour ward on the second floor, Tygerberg Hospital.

***Delete where not applicable**

Appendix B: Protocols

1. **Multiphor SSCP/HA analysis**
   (Modified protocol retrieved from Liechti-Gallati *et al.,* 1999)

Multiphor SSCP/Heteroduplex gel

- Wash plates 4X with 70% ethanol
- Add 80 µl plateglue and 8 µl 10% acetic acid on plate without blocks. Add 80 µl repelcote onto the plate that contains wells
- Wipe until resistance is felt and wash a few times with 70% ethanol
- Wipe spacers with 70% ethanol and place them on the plates
- Clamp on both short sides and the long area close to blocks (wells). Place upside down on bench
- Prepare gel mix
    - 15 ml gel mix (Appendix C)
    - 150 ml 10% APS (Appendix C)
    - 15 ml TEMED (Promega, Madison WI, USA)
- Pour quickly using a syringe
- Allow to set at room temperature for 30 minutes

Setting up the apparatus (Amersham Pharmacia Biotech, Amsterdam, UK)

- Separate plates
- Add dH_2O onto surface of Multiphor apparatus
- Slide plate with gel facing upwards over water
- Use 2 Whatmann buffer strips per side per 2 hour run soaked in TRIS-borate buffer (Appendix C)
- Place strips on both sides of gel close to the wells
- Wash electrodes of Multiphor gently each time before use

PCR products

- Take 3-5 µl PCR product and 3 - 5 µl SSCP loading dye (Appendix C)
- Denature at 95ºC for 5 min in PCR thermocycler
- Place on ice immediately
- Load 3 µl onto gel and run at optimised temperature (4ºC - 21ºC) for optimised duration at 355 V

Visualisation of bands

- Remove gel from apparatus
- Rinse 2X in dH_2O
- Incubate for 10 min at room temperature in solution 1 (Silver Nitrate solution) (Appendix C)
- Rinse 2X in dH_2O
- Incubate for 10 min at room temperature in solution 2 (Developing Soloution) (Appendix C)
- Rinse 2X with dH_2O
- Blot dry with paper towel
- Cover gel with Whatmann filter paper

2. Protocol for genomic DNA extraction from whole blood (Protocol retrieved from Miller *et al.,* 1988)

Day 1:

- Place ~10 ml blood in a 50 ml Falcon tube
- Add 30 ml cold Lysis Buffer (Appendix C)
- Place on ice for 15 min, shaking each 5 min interval
- Centrifuge @ 1500 rpm for 10 min
- Pour off supernatant, keeping pellet
- Add 10 ml PBS (Appendix C) and mix
- Centrifuge @ 1500rpm for 10 min
- Pour off supernatant keeping pellet
- Dissolve Pellet in: 30 ml Cell Lysis Buffer
  30 µl Proteinase K (10 mg/ml)
  300 µl 10% SDS
- Mix well and incubate overnight in 55 ºC waterbath

Day 2:

- Add 1 ml 6M NaCl and shake for 1 min
- Centrifuge @ 3500 rpm for 30 min
- Transfer supernatant to new tube and shake for 15 s
- Centrifuge @ 2500 rpm for 15 min
- Transfer supernatant to new tube without foam or pellet
- Add 2 volumes ice cold 100% ethanol to precipitate DNA
- Scoop out DNA and place in Eppendorf tube with 500 µl 70% ethanol
- Centrifuge @ 14 000 rpm for 10 min @ 4ºC
- Dissolve pellet in 200-800 µl ddH_2O/TE buffer depending on pellet size

3. Rapid DNA isolation from 300µl whole blood using the Pure Gene® DNA isolation Kit (Gentra Systems™, Minneapolis, USA)

Cell Lysis

- Add 300 µl whole blood to 900 µl RBC Lysis Solution and incubate for 1 min at room temperature, invert gently 10 times during incubation
- Centrifuge for 20 s at 13 000 - 16 000 g.
- Remove supernatant (leave ~10 µl liquid)
- Vortex pellet to resuspend
- Add 300 µl Cel Lysis Solution and pipet up and down

Protein Precipitation

- Add 100 µl Protein Precipitation Solution and vortex at high speed for 20 s
- Centrifuge for 1 min at 13 000 - 16 000 g (dark brown pellet forms)

DNA precipitation

- Transfer supernatant to a clean Eppendorf tube containing 300 µl 100% isopropanol and mix by inverting samples 50 times.
- Centrifuge at 13 000 - 16 000 g for 60 s (obtain small white pellet)
- Pour off supernatant and drain tube on clean absorbent paper
- Add 300 µl 70% ethanol and invert tube several times to wash pellet
- Centrifuge at 13 000 - 16 000 g for 60 s
- Carefully pour of ethanol, invert and drain tube for 30 min on clean towel paper

DNA Hydration

- Add 20 µl DNA Hydration Solution and dissolve for 30 min
- Store at 4 ºC

Appendix C: Solutions

1. DNA Extraction Method

Extraction Buffers and chemicals used:

Nuclear Lysis Buffer

1,211 g	Tris-Cl	0.01 M
23,4 g	NaCl	0.4 M
0.6 g	EDTA	0.002 M

pH = 8.2 with 1M NaOH

BTV (1 l) with dH_2O

Storage: 4ºC

Cell Lysis Buffer

8.3 g	NH_4Cl	0.155 M
1.1 g	$KHCO_3$	0.01 M
0.03 g	EDTA	0.0001 M

pH = 7.4 with HCl

BTV (1 l) with dH_2O

Storage: 4ºC

Phosphate Buffered Saline (PBS)

2 g	KCl	0.027 M
8 g	NaCl	0.137 M
1.14 g	Na_2HPO_4	0.008 M
0.2 g	KH_2PO_4	0.0015 M

BTV (1 l) with dH_2O

Storage: 4ºC

10% SDS

10g SDS

BTV (100 ml) with dH_2O

Storage: Room Temp (Prevent Precipitation)

Proteinase K

10 mg/ml dissolved in dH_2O

Storage: -20ºC

6 M NaCl

175.32 g NaCl

BTV (500 ml) with dH_2O

Storage: Room Temp

2. Agarose Gel Electrophoresis

10X TBE

108 g Trizma Base

55 g Boric Acid

7.445 g EDTA

BTV (1 l) with dH_2O. Dilute 1 in 10 for 1X TBE buffer.

Cressol

Stock Solution:

10 mg Cressol

Dissolved in 1 ml

Dilute:

200 µl Cressol Stock

3.4 g Sucrose

9.5 ml dH_2O

3. Multiphor SSCP/HD Analysis

40% acrylamide-PDA Solution

396 g Acrylamide
4 g Piperazine diacrylamide
BTV (1 l) with dH_2O

0.75 M TRIS-Formate Buffer

90.8 g Trizma Base
600 ml dH_2O
pH = 9.0 with formic acid
BTV (1 l) with dH_2O

41% Glycerol

41 ml Glycerol
59 ml dH_2O

TRIS-Borate Buffer

125.9 g Trizma Base
17.3 g Boric acid (H_3BO_3)
700 ml dH_2O (dissolve)
pH = 9.0
50 µl Bromophenol Blue(4% solution)
BTV (1 l) with dH_2O

SSCP loading dye

95% (47.5ml) Formamide
100 mM (0.16 g) NaOH
0.25% (0.125 g) Bromophenol Blue
0.25% (0.125 g) Xylene Cyanole
BTV (50 ml) with dH_2O

10% APS

0.15 g APS
BTV (1.5 ml) with dH_2O

Plateglue

200 µl 3-(Trimethoxysily) propyl methacrylate
50 ml 100% Ethanol

Gel Mix (X10)

53 ml 40% acrylamide-PDA
85 ml Tris-Formate Buffer
30 ml 41% Glycerol
Storage: 4ºC in dark

Silver Staining Solutions

Solution I

0.75 g silver nitrate ($AgNO_3$)
300 ml dH_2O
Storage: in dark for +/- 3 days

Solution II

4.5 g Sodium Hydroxide pellets (NaOH)
300 ml dH_2O
3 ml Formaldehyde solution

4. HEX-SSCP Analysis

12% PAA Gel Mix

12% (w/v) [40% Stock Solution with 1%C (99 Acrylamide : 1 Bisacrylamide)]

0.1% (w/v) APS

0.1% (v/v) TEMED

7.5% (w/v) Urea

Bromophenol blue loading dye

95% (v/v) formamide

20 mM EDTA

0.05% (w/v) xylene cyanol

0.05% (w/v) bromophenol blue

Appendix D: Patient Questionnaires

GENETIC ASPECTS OF PRE-ECLAMPSIA:

NAME:

HOSPITAL:

BIRTH DATE:

1. AGE:
2. RACE:
3. GRAVIDITY:PARITY:MISCARRIAGES:ECTOPIC

PREVIOUS PREGNANCIES

4. YEAR: .....
5. PARTNER: (NO) .....
6. G.A. COMPLICATIONS: .....
7. COMPLICATIONS: (LIST) .....
8. METHOD DELIVERY: .....
9. BIRTHWEIGHT: .....
10. G.A. –DELIVERY: .....
11. OUTCOME: (LIST) .....
12. NICU:YES=1, NO=2 .....
13. SMOKE: YES=1, NO=2 .....
14. ALKOHOL: YES=1, NO=2 .....

CURRENT PREGNANCY

15. VDRL:NEG=0, POS TREATED=1, NOT TREATED=2
16. BLOOD GROUP:
17. CERVICAL CYTOLOGY: NORMAL=1, ABNORMAL=2
18. CERVICAL CULTURES:
    NOT DONE=1, NEG=2, GONO=3, CHLAMYDIA=4, GBS=5, OTHER=6
19. URINE MCS: NEG=2, ASYMP. BACTERIA TREATED=1, NOT TREATED=2,
    NOT DONE=3, UTI=4
20. SF GROWTH: $<10^{th}$=1, NORMAL=2, $>90^{th}$=3

21. PROTEINURIA: (GESTATION)
22. BP: (ADMISSION)
23. AMNIOCENTESIS FOR KARYOTYPING: YES=1, NO=2 (GESTATION)
24. FETAL MOVEMENT:NORMAL=1, DECREASED=2
25. ULTRASOUND: YES=1, NO=2
26. DOPPLER:N=1, >95th =2, AEDV=3,REDF=4
27. G.A. WHEN COMPLICATIONS DEVELOPED:
28. COMPLICATIONS:……....
    .....
29. SMOKE:YES=1, NO=2
30. ALCOHOL:YES=1, NO=2
31. MEDICATION AT ANY TIME DURING PREGNANCY:
    FOLATE=1, Fe=2, ASPIRIN=3,
    ANTIHYPERTENSIVES=4,PYRIDOXIN=5,OTHER=6
32. GESTATION:
33. BIRTH WEIGHT:
    a.) MALE=1,FEMALE=2
    b.) DELIVERY TYPE:
34. APGAR SCORE:
35. OUTCOME: (LIST)
36. MORBIDITY: (LIST)
37. NICO:YES=1, NO=2
38. DAYS:
39. REASON:.....
40. SPECIAL INVESTIGATIONS:
    UREAM: ..... LDH: .....
    KREAT: ..... WCC: .....
    AST: ..... HB: .....
    ALT: ..... HKT: .....

GENERAL INFORMATION

ADDRESS:

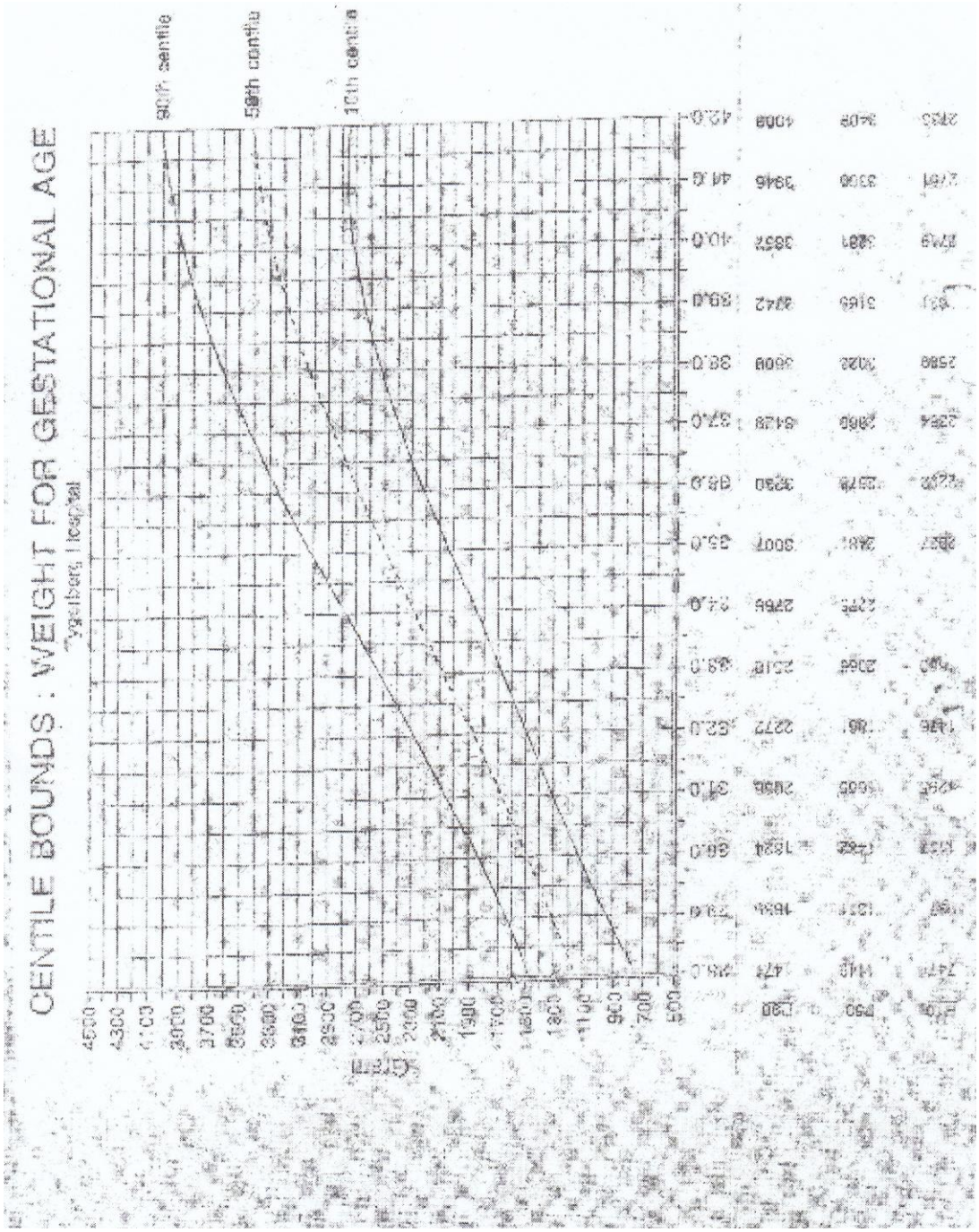
BORN AND RAISED:

41. YOURSELF: .....
42. YOUR MOTHER:….....
43. YOUR FATHER:…….....
44. CURRENT PARTNER: (FATHER OF YOUR CHILD).....
45. YOUR PARTNER MOTHER:.....
46. HAS SHE HAD TROMBOTIC EPISODES:.....
47. ANY COMPLICATIONS DURING YOUR PARTNER'S BIRTH:
    .....
48. PREVIOUS PARTNER(S): -CHILDHOOD AND PLACE OF BIRTH:
    .....
49. DO YOU HAVE SISTERS OR HALF SISTERS WHO HAVE THE FOLLOWING COMPLICATION PRETERM BIRTH=1, DEAD BABY=2, BLEEDING LATE IN PREGNANCY=3, HIGH BLOOD PRESSURE=4

    ADDRESS OF SISTER:

50. DID YOUR MOTHER HAVE ANY OF THE FOLLOWING:
    PRETERM BIRTH=1, DEAD BABY=2, BLEEDING LATE IN PREGNANCY=3, HIGH BLOOD PRESSURE=4, TROMBOSIS=5

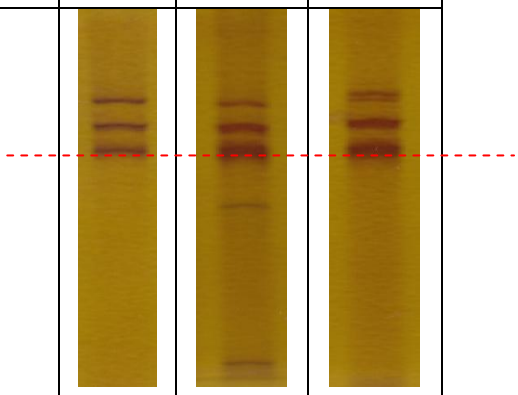
Appendix E: Fetal Growth Chart



Appendix F: Multiphor SSCP/HD gels and DNA Sequencing Electropherograms

Figures 1 – 11 represents the results of Multiphor SSCP/HD analysis of each PCR fragment in the format of a figure that indicates the various conformations detected on the Multiphor gels and the subsequent genotypes for the identified variants as detected by bi-directional DNA sequencing. Sequencing electropherograms representing the heterozygous status of each variant are included below the Multiphor gels.

a)

Conformations	1	2	3
			
c.-442 C>G	CC	CC	C**G**
c.-191 G>C	GG	G**C**	GG
c.-189_-188insGCCGG	WT	**HET**	WT
c.-135 C>G	CC	CC	C**G**

b)

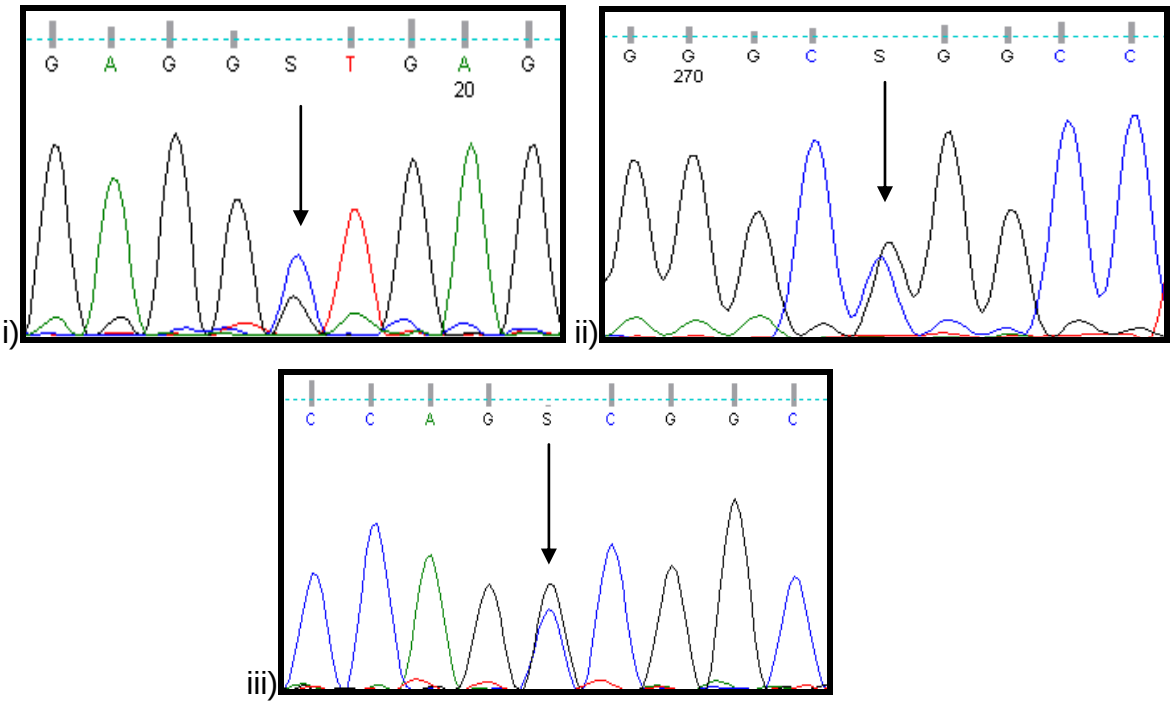
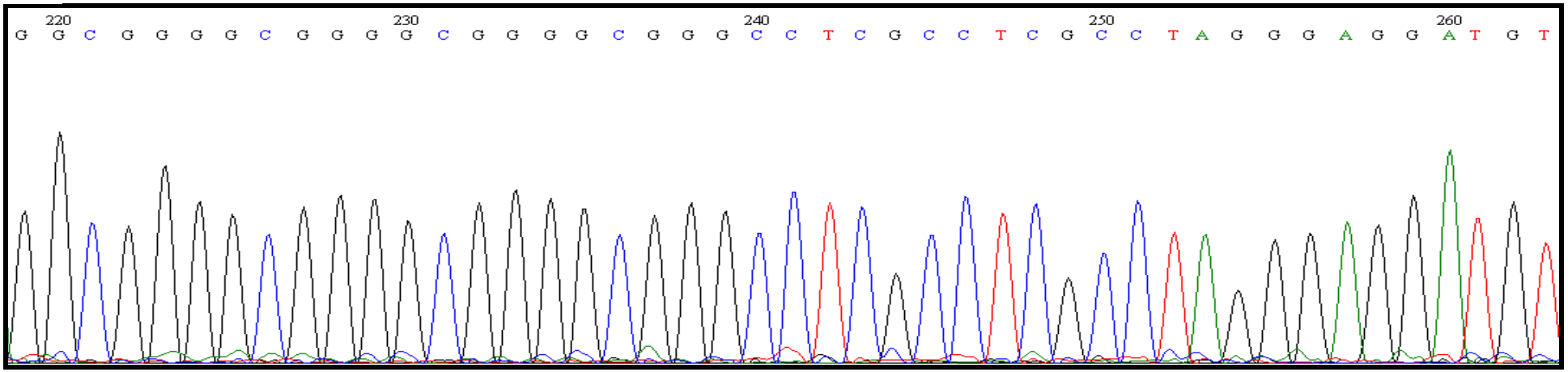


Figure 1. a) Conformational variants as detected with Multiphor SSCP/HD analysis in PCR fragment 1a (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating point of variation) for the SNPs (i) c.-442 C>G, (ii) c.-191 G>C and (iii) c.-135 C>G.

a)



b)

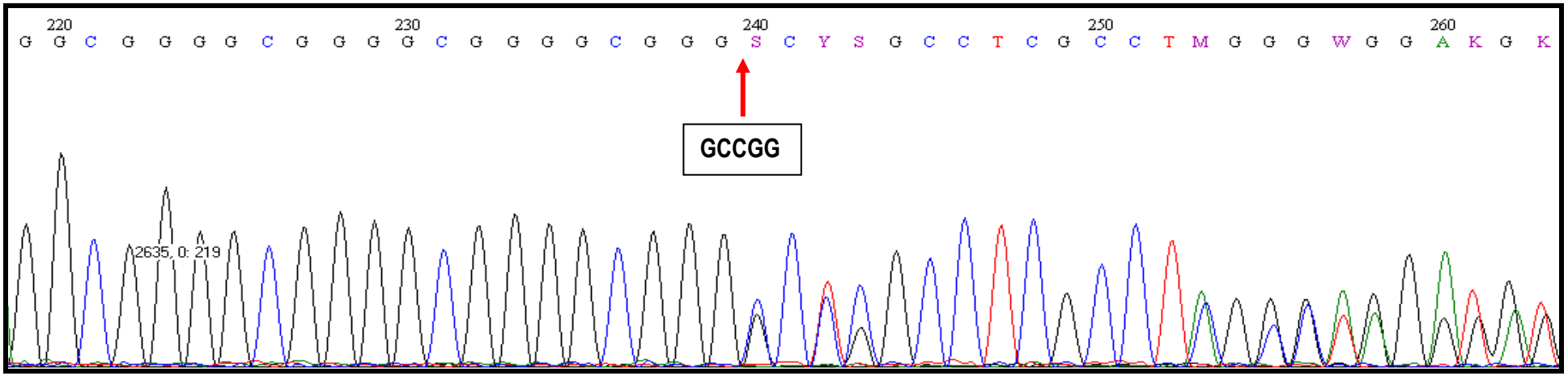
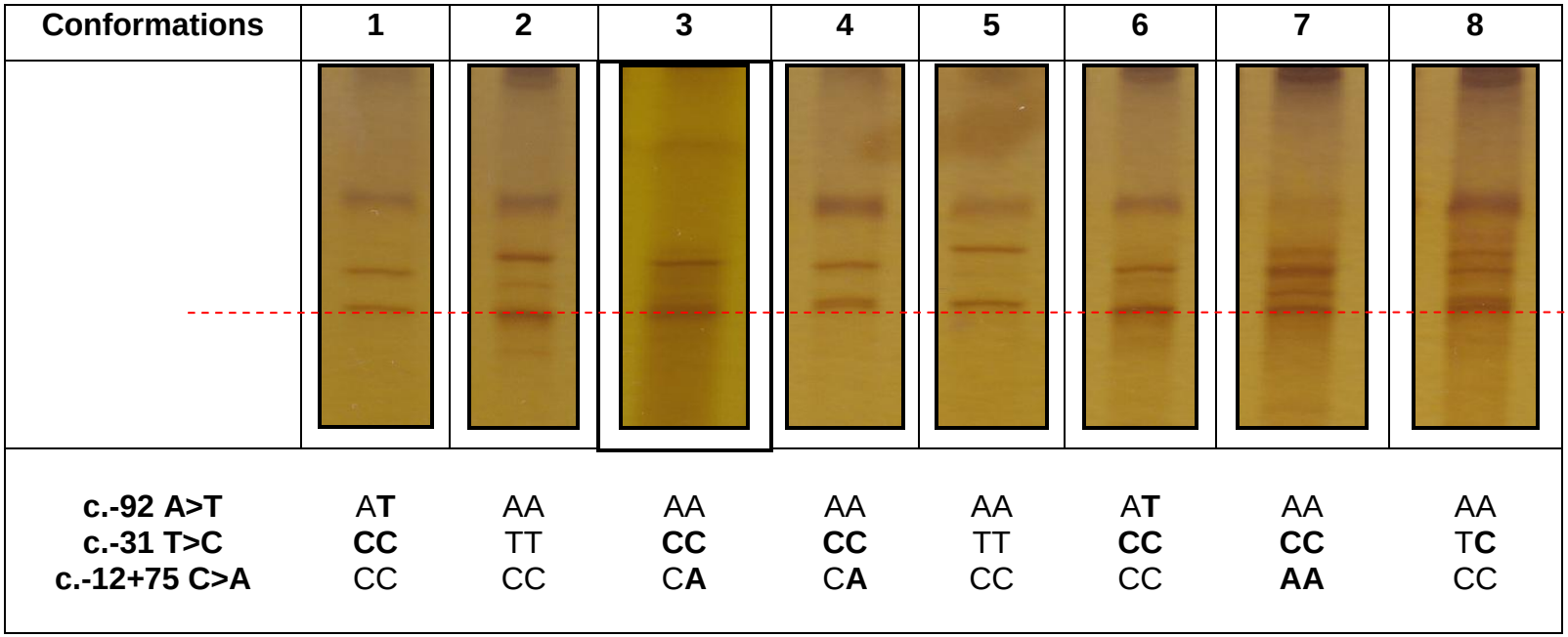


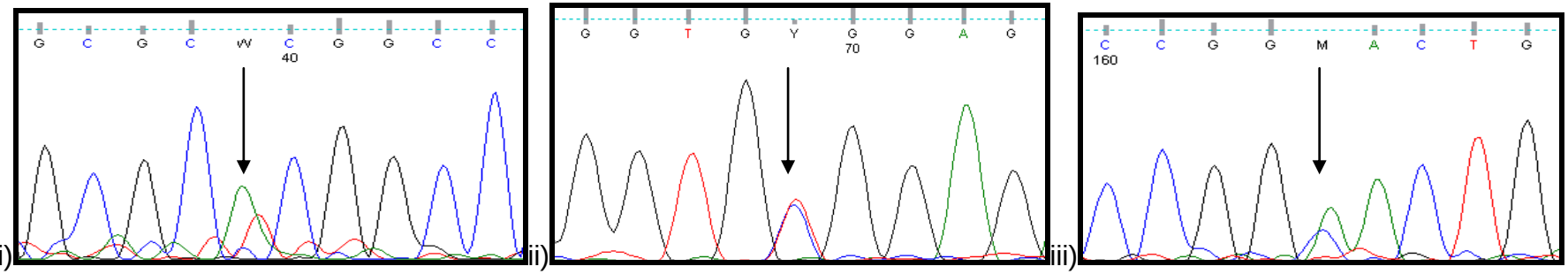
Figure 2. a) Electropherogram representing the homozygous wild type status for the PCR fragment 1a.
b) Electropherogram representing the heterozygous status (arrow indicating point of variation) for insertion c.-189_-188insGCCGG.

a)



Conformations	1	2	3	4	5	6	7	8
c.-92 A>T	A**T**	AA	AA	AA	AA	A**T**	AA	AA
c.-31 T>C	**CC**	TT	**CC**	**CC**	TT	**CC**	**CC**	T**C**
c.-12+75 C>A	CC	CC	C**A**	C**A**	CC	CC	**AA**	CC

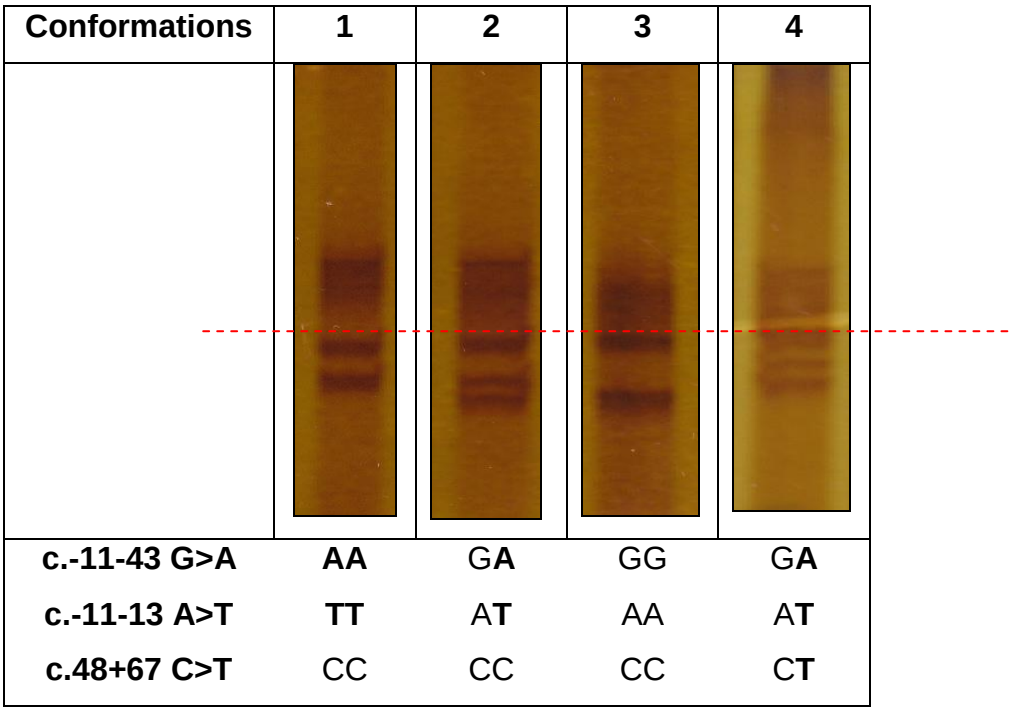
b)



i) ii) iii)

Figure 3. a) Conformational variants as detected with Multiphor SSCP/HD analysis in PCR fragment 1b (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating point of variation) for the SNPs (i) c.-92 A>T, (ii) c.-31 T>C and (iii) c.-12+75 C>A.

a)



Conformations	1	2	3	4
c.-11-43 G>A	**AA**	G**A**	GG	G**A**
c.-11-13 A>T	**TT**	A**T**	AA	A**T**
c.48+67 C>T	CC	CC	CC	C**T**

b)

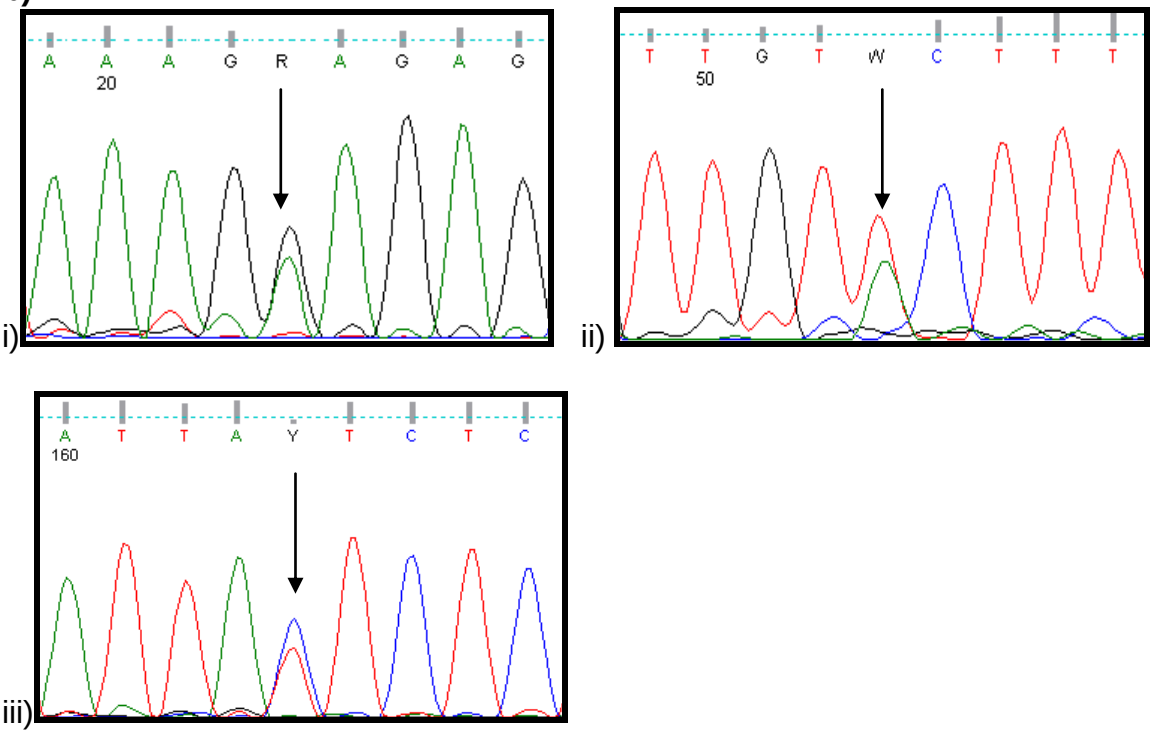


Figure 4. **a)** Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containing exon 2 (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating point of variation) for the SNPs (a) c.-11-43 G>A, (b) c.-11-13 A>T and (c) c.48+67 C>T.

a) **b)**

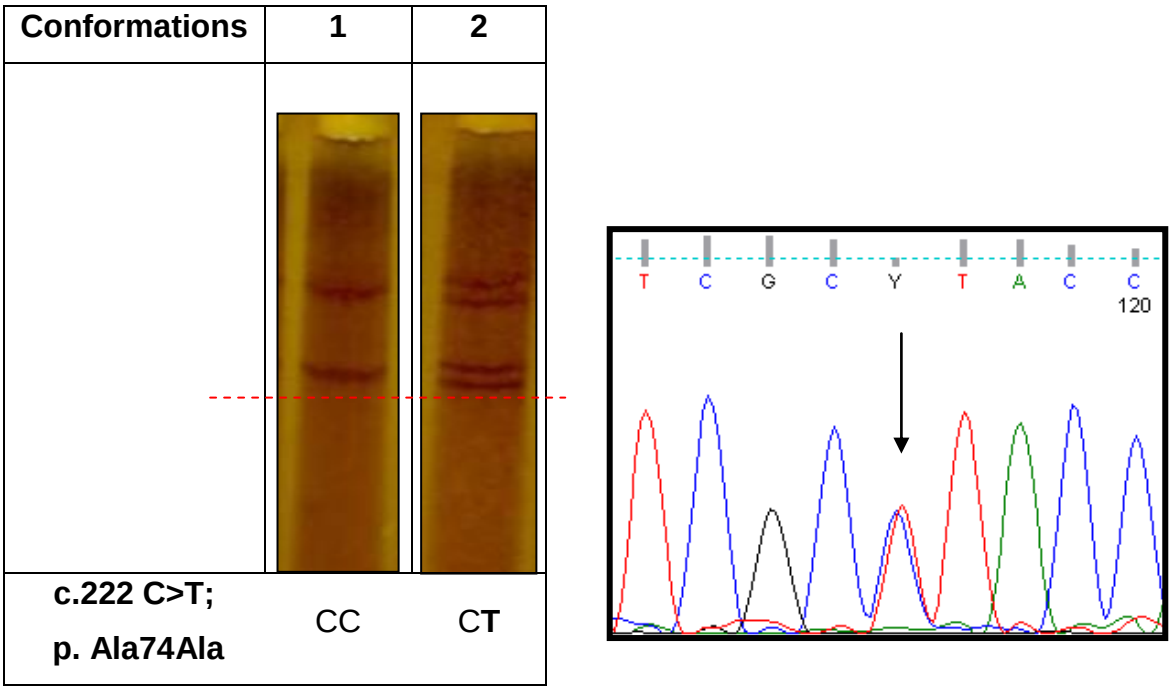
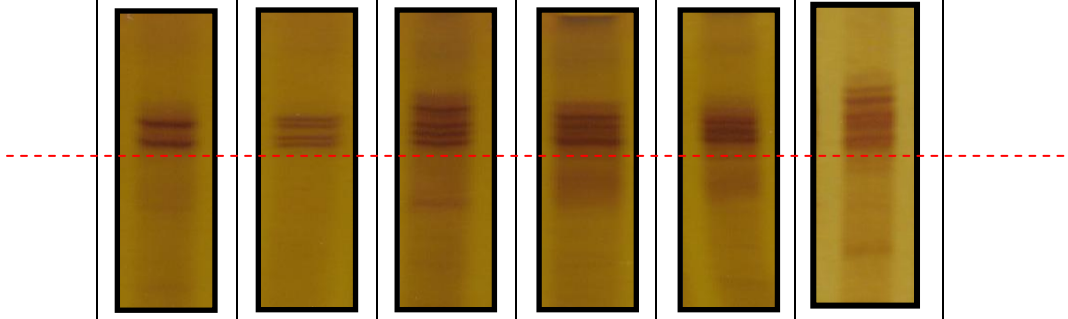


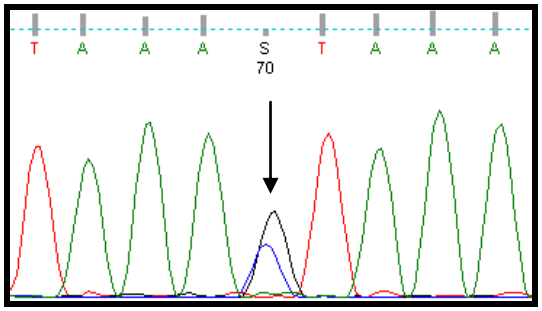
Figure 5. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the fragment containing exon 4 (variant allele indicated in bold).

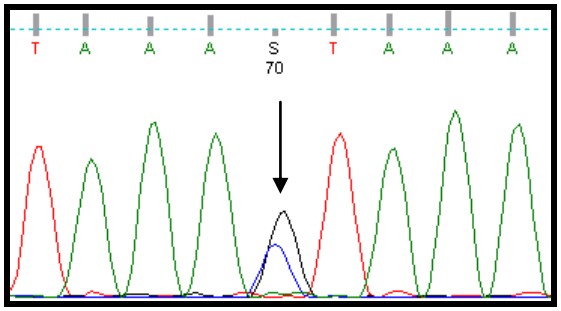
b) Electropherogram representing the heterozygous status (arrow indicating point of variation) for the SNP c.222 C>T; p.Ala74Ala.

a)

Conformations	1	2	3	4	5	6
						
c.244-76 C>G	CC	C**G**	C**G**	CC	CC	**GG**
c.244-42 G>C	GG	GG	G**C**	G**C**	GG	GG
c.292 G>T; p.Val98Leu	GG	GG	GG	GG	G**T**	GG

b)

i) 

ii) 

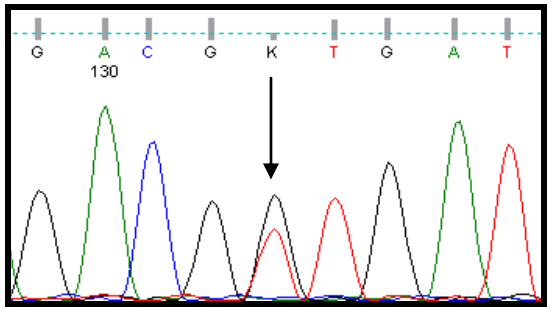
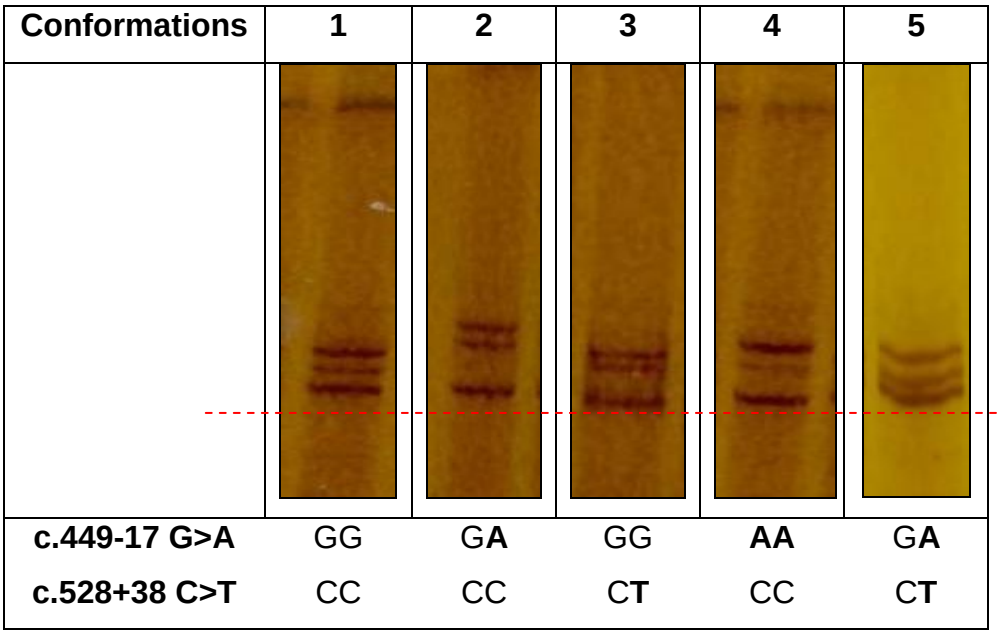
iii) 

Figure 6. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containning exon 5 (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating point of variation) for the SNPs (a) c.244-76 C>G, (b) c.244-42 G>C and (c) c.292 G>T; p.Val98Leu.

a)



b)

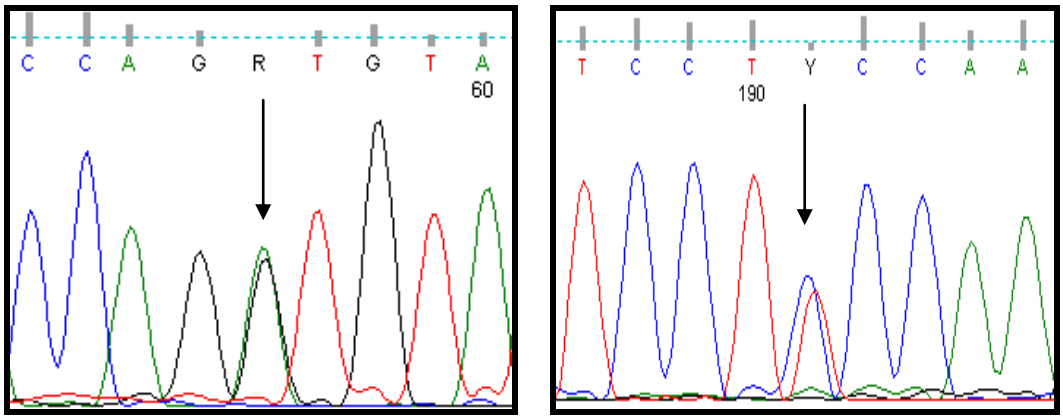
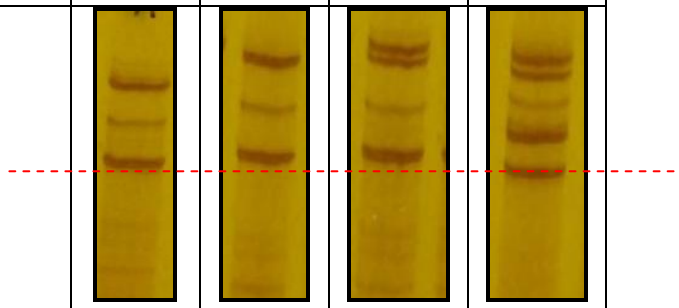


Figure 7. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containing exon 7(variant allele indicated in bold).

b) Electropherograms representing the heterozygous status (arrow indicating) for the SNPs (i) c.449-17 G>A and (ii) c.528+38 C>T.

a)

Conformations	1	2	3	4
				
c.589-5 C>T	CC	C**T**	CC	CC
c.682+49 C>T	CC	CC	C**T**	CC
c.600 C>T; p.Asp110Asp	CC	CC	CC	C**T**

b)

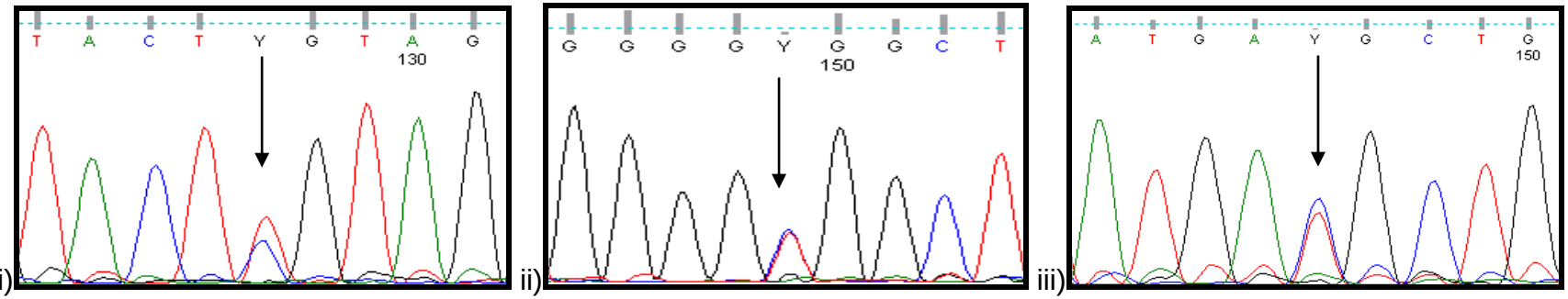


Figure 8. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containing exon 9 (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating) for the SNPs (i) c.589-5 C>T, (ii) c.682+49 C>T and (iii) c.600 C>T; p.Asp110Asp.

a)

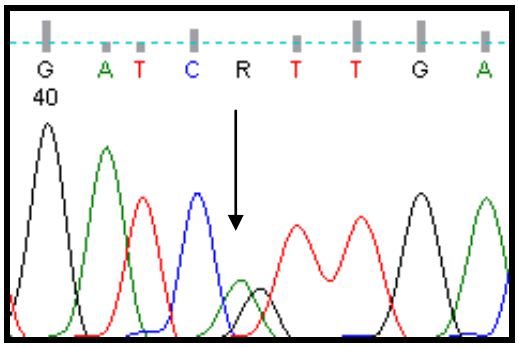
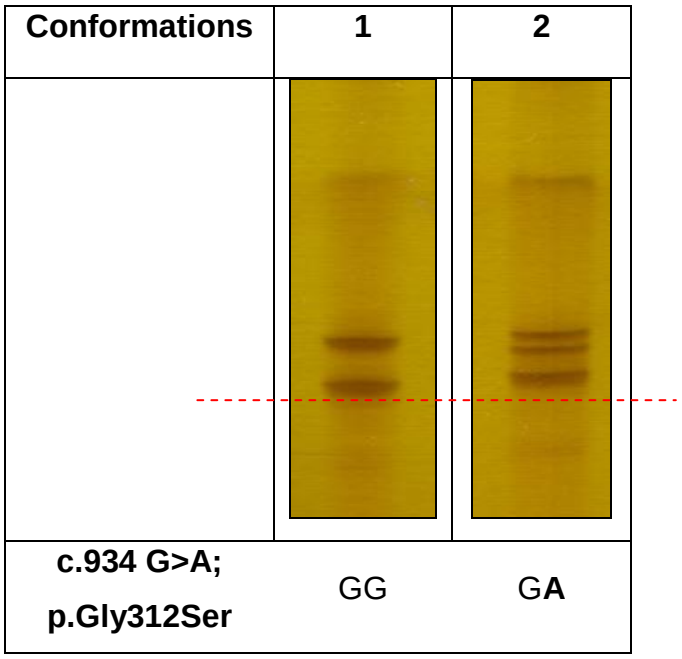


Figure 9. Electropherogram representing the heterozygous status (arrow indicating point of variation) for the SNP c.683-56 G>A.

a)



b)

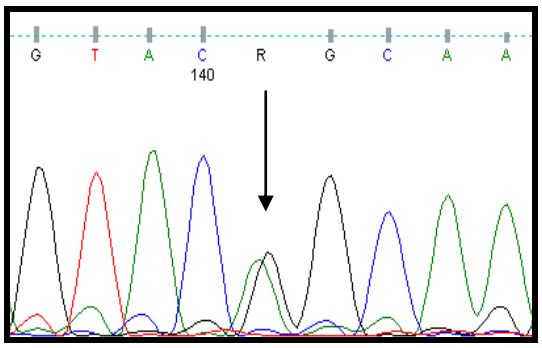
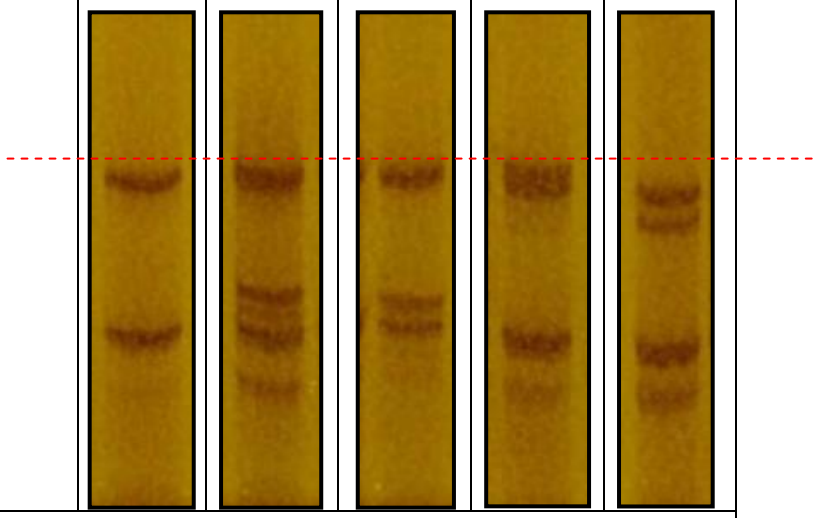
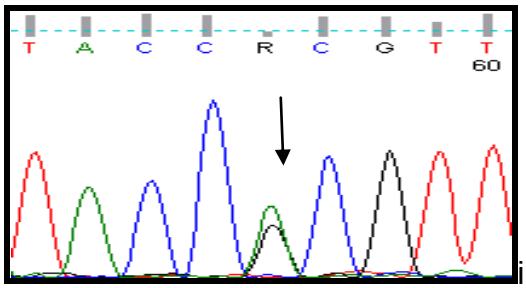


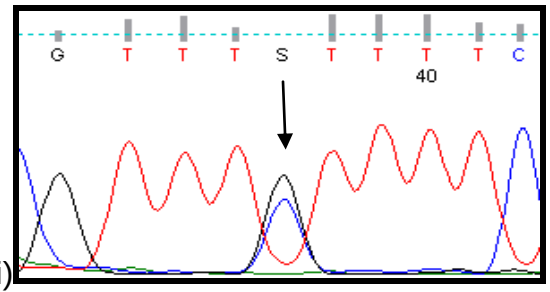
Figure 10. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containing exon 12 (variant allele indicated in bold). **b)** Electropherogram representing the heterozygous status (arrow indicating point of variation) for the SNP c.934 G>A; p.Gly312Ser.

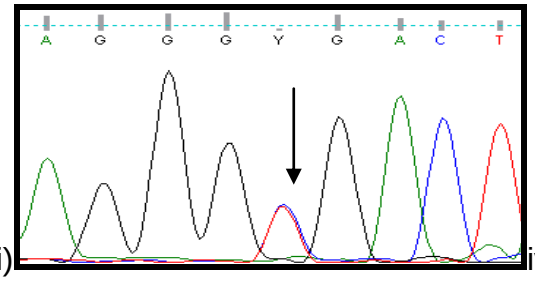
a)

Conformations	1	2	3	4	5
					
c.961-30 A>G	AA	A**G**	AA	AA	AA
c.961-24 C>G	CC	CC	CC	CC	C**G**
c.975 C>T; p.Gly325Gly	CC	CC	CC	C**T**	CC
c.*1057 A>G	AA	AA	A**G**	AA	AA

b)

i) 

ii) 

iii) 

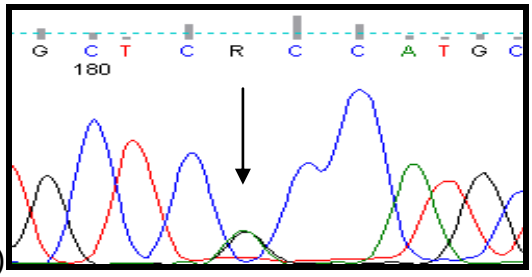
iv) 

Figure 11. a) Conformational variants as detected with Multiphor SSCP/HD analysis in the amplicon containing exon 13 (variant allele indicated in bold). **b)** Electropherograms representing the heterozygous status (arrow indicating point of variation) for the SNPs (i) c.961-30 A>G, (ii) c.961-24 C>G, (iii) c.975 C>T; p.Gly325Gly and (d) c.*1057 A>G.

Appendix G: Genotype and Allele Frequencies

Genotype and allele frequencies as identified in the *ANXA2* gene in the pre-eclamptic maternal, fetal and control study cohorts are presented in Tables 1-12. The probability values (*P*-value) that would indicate significant association between maternal and control and fetal and control cohorts as verified in Microsoft® Office Excell (2003) by way of 2X3 contingency tables are indicated for all variants identified in the *ANXA2* gene. In the case of a dash, no result could be obtained. Significant *P*-values ($P<0.05$) identified are highlighted in blue. The variants that showed deviation from the Hardy-Weinberg equilibrium ($P>0.05$) are highlighted in red.

Table 1. Genotype and allele frequencies of the c.-442 C>G and c.-191 G>C variants identified in the *ANXA2* 5'UTR.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-442 C>G	**Genotype Distribution**	**n**	108	66	42	80	47	33	53
		CC	106 (0.981)	66 (1.000)	40 (0.952)	77 (0.963)	44 (0.936)	33 (1.000)	53 (1.000)
		CG	2 (0.019)	0	2 (0.048)	3 (0.037)	3 (0.064)	0	0
		GG	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	0.991	1.000	0.976	0.981	0.968	1.000	1.000
		G	0.009	0	0.024	0.019	0.032	0	0
		***P* value**	0.320	-	0.110	0.156	0.064	-	-
	HWE	***P* value**	0.9226	-	0.8744	0.8643	0.8212	-	-
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-191 G>C	**Genotype Distribution**	**n**	108	66	42	80	47	33	53
		GG	103 (0.954)	64 (0.970)	39 (0.930)	79 (0.988)	47 (1.000)	32 (0.970)	52 (0.981)
		GC	5 (0.046)	2 (0.030)	3 (0.070)	0 (0.012)	0	1 (0.030)	1 (0.019)
		CC	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**G**	0.977	0.985	0.964	1.000	1.000	0.985	0.991
		C	0.023	0.015	0.036	0	0	0.015	0.009
		***P* value**	0.392	0.694	0.210	0.211	0.345	0.734	-
	HWE	***P* value**	0.8055	0.9005	0.8103	-	-	0.9296	0.9447

Table 2. Genotype and allele frequencies of the c.-189_-188GCCGG and c.-135 C>G variants identified in the *ANXA2* 5'UTR

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-189_-188GCCGG	Genotype Distribution	n	108	66	42	80	47	33	53
		1_1	103 (0.954)	64 (0.970)	39 (0.930)	79 (0.988)	47 (1.000)	32 (0.970)	52 (0.981)
		1_Ins	5 (0.046)	2 (0.030)	3 (0.070)	1 (0.012)	0	1 (0.030)	1 (0.019)
		Ins_Ins	0	0	0	0	0	0	0
		P value	-	-	-	-	-	-	-
	Allelic Distribution	1	0.977	0.985	0.964	0.994	1.000	0.985	0.991
		Ins	0.023	0.015	0.036	0.006	0	0.015	0.009
		P value	0.392	0.694	0.210	0.769	0.345	0.734	-
	HWE	*P* value	0.8055	0.9005	0.8103	0.9551	-	0.9296	0.9447
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-135 C>G	Genotype Distribution	n	108	66	42	80	47	33	53
		CC	106 (0.981)	65 (0.985)	40 (0.952)	77 (0.963)	44 (0.936)	33 (1.000)	52 (0.981)
		CG	2 (0.019)	1 (0.015)	2 (0.048)	3 (0.037)	3 (0.064)	0	1 (0.019)
		GG	0	0	0	0	0	0	0
		P value	-	-	-	-	-	-	-
	Allelic Distribution	C	0.991	0.992	0.976	0.981	0.968	1.000	0.991
		G	0.009	0.008	0.024	0.019	0.032	0	0.009
		P value	0.988	0.876	0.430	0.541	0.257	0.429	-
	HWE	*P* value	0.9226	0.9506	0.8744	0.8643	0.8212	-	0.9447

Table 3. Genotype and allele frequencies of the c.-92 G>C and c.-31 T>C variants identified in the *ANXA2* 5'UTR.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-92 A>T	**Genotype Distribution**	**n**	107	66	41	73	41	32	51
		AA	72 (0.673)	45 (0.682)	27 (0.659)	51 (0.699)	32 (0.780)	19 (0.594)	43 (0.843)
		AT	35 (0.327)	21 (0.318)	14 (0.341)	22 (0.301)	9 (0.220)	13 (0.406)	8 (0.157)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**A**	0.836	0.841	0.829	0.849	0.890	0.797	0.922
		T	0.164	0.159	0.171	0.151	0.110	0.203	0.078
		***P* value**	0.039	0.063	0.055	0.086	0.466	0.019	-
	HWE	***P* value**	0.0431	0.1243	0.1874	0.1296	0.4299	0.1493	0.5433
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-31 T>C	**Genotype Distribution**	**n**	107	66	41	73	41	32	51
		TT	28 (0.261)	17 (0.258)	11 (0.268)	37 (0.507)	23 (0.561)	14 (0.438)	21 (0.412)
		TC	25 (0.234)	17 (0.258)	8 (0.195)	5 (0.068)	4 (0.098)	1 (0.031)	15 (0.294)
		CC	54 (0.505)	32 (0.485)	22 (0.537)	31 (0.425)	14 (0.341)	17 (0.531)	15 (0.294)
		***P* value**	0.039	0.088	0.062	0.003	0.065	0.007	-
	Allelic Distribution	**T**	0.379	0.386	0.366	0.541	0.610	0.453	0.559
		C	0.621	0.614	0.634	0.459	0.390	0.547	0.441
		***P* value**	0.003	0.009	0.009	0.783	0.486	0.185	-
	HWE	***P* value**	0.0000	0.0002	0.0002	0.0000	0.0000	0.0000	0.0040

Table 4. Genotype and allele frequencies of the c.-12+75 C>A and c.-11-43 G>A variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-12+75 C>A	**Genotype Distribution**	**n**	107	66	41	73	41	32	51
		CC	88 (0.822)	55 (0.833)	33 (0.805)	64 (0.877)	36 (0.878)	28 (0.875)	44 (0.863)
		CA	16 (0.150)	10 (0.152)	6 (0.146)	9 (0.123)	5 (0.122)	4 (0.125)	6 (0.117)
		AA	3 (0.028)	2 (0.049)	2 (0.049)	0	0	0	1 (0.020)
		***P* value**	0.812	0.821	0.661	0.485	0.666	0.726	-
	Allelic Distribution	**C**	0.897	0.896	0.878	0.938	0.939	0.938	0.922
		A	0.103	0.104	0.122	0.062	0.061	0.063	0.078
		***P* value**	0.490	0.495	0.323	0.607	0.646	0.700	-
	HWE	***P* value**	0.0501	0.0976	0.0426	0.5746	0.6776	0.7061	0.1837
			Total Patients						
			Maternal			**Fetal**			**Controls**
			Total	**Mixed Ancestry**	**Black**	**Total**	**Mixed Ancestry**	**Black**	**Total**
c.-11-43 G>A	**Genotype Distribution**	**n**	118	71	47	94	52	38	54
		GG	12 (0.102)	6 (0.085)	6 (0.128)	18 (0.191)	12 (0.231)	6 (0.158)	7 (0.130)
		GA	63 (0.534)	40 (0.563)	23 (0.489)	42 (0.447)	23 (0.442)	19 (0.500)	22 (0.407)
		AA	43 (0.364)	25 (0.352)	18 (0.383)	34 (0.362)	17 (0.327)	13 (0.342)	25 (0.463)
		***P* value**	0.306	0.218	0.685	0.410	0.244	0.511	-
	Allelic Distribution	**G**	0.369	0.366	0.372	0.415	0.452	0.408	0.333
		A	0.631	0.634	0.628	0.585	0.548	0.592	0.667
		***P* value**	0.526	0.590	0.562	0.165	0.077	0.301	-
	HWE	***P* value**	0.1104	0.0718	0.7474	0.4396	0.8286	0.4398	0.5403

Table 5. Genotype and allele frequencies for the c.-11-13 A>T and c.48+67 C>T variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.-11-13 A>T	Genotype Distribution	n	118	71	47	94	52	38	54
		AA	12 (0.102)	6 (0.085)	6 (0.128)	18 (0.191)	12 (0.231)	6 (0.158)	7 (0.130)
		AT	63 (0.534)	40 (0.563)	23 (0.489)	42 (0.447)	23 (0.442)	19 (0.500)	22 (0.407)
		TT	43 (0.364)	25 (0.352)	18 (0.383)	34 (0.362)	17 (0.327)	13 (0.342)	25 (0.463)
		P value	0.306	0.218	0.685	0.410	0.244	0.511	-
	Allelic Distribution	A	0.369	0.366	0.372	0.415	0.452	0.408	0.333
		T	0.631	0.634	0.628	0.585	0.548	0.592	0.667
		P value	0.526	0.590	0.562	0.165	0.077	0.301	-
	HWE	*P* value	0.1104	0.0718	0.7474	0.4396	0.8286	0.4398	0.5403
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.48+67 C>T	Genotype Distribution	n	118	71	47	94	56	38	54
		CC	117 (0.991)	70 (0.986)	47 (1.000)	94 (1.000)	56 (1.000)	38 (1.000)	54 (1.000)
		CT	1 (0.009)	1 (0.014)	0	0	0	0	0
		TT	0	0	0	0	0	0	0
		P value	-	-	-	-	-	-	-
	Allelic Distribution	C	0.996	0.993	1.000	1.000	1.000	1.000	1.000
		T	0.004	0.007	0	0	0	0	0
		P value	0.498	0.382	-	-	-	-	-
	HWE	*P* value	0.9631	0.9523	-	-	-	-	-

Table 6. Genotype and allele frequencies of the c.244-76 C>G and c.244-42 G>C variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.244-76 C>G	**Genotype Distribution**	**n**	119	71	48	77	43	34	52
		CC	103 (0.865)	66 (0.930)	37 (0.771)	69 (0.896)	41 (0.953)	28 (0.824)	49 (0.942)
		CG	16 (0.135)	5 (0.070)	11 (0.229)	7 (0.091)	1 (0.023)	6 (0.176)	3 (0.058)
		GG	0	0	0	1 (0.013)	1 (0.023)	0	0
		***P* value**	-	-	-	0.154	0.392	-	-
	Allelic Distribution	**C**	0.933	0.965	0.885	0.942	0.965	0.912	0.971
		G	0.067	0.035	0.115	0.058	0.035	0.088	0.029
		***P* value**	0.154	0.781	0.018	0.268	0.813	0.087	-
	HWE	***P* value**	0.4317	0.7584	0.3699	0.1269	0.0000	0.5726	0.8340
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.244-42 G>C	**Genotype Distribution**	**n**	119	71	48	77	43	34	52
		GG	115 (0.966)	69 (0.942)	46 (0.958)	73 (0.948)	42 (0.947)	31 (0.912)	51 (0.981)
		GC	4 (0.034)	2 (0.028)	2 (0.042)	4 (0.052)	1 (0.023)	3 (0.088)	1 (0.019)
		CC	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**G**	0.983	0.986	0.979	0.974	0.988	0.956	0.990
		C	0.017	0.014	0.021	0.026	0.012	0.044	0.010
		***P* value**	0.610	0.752	0.514	0.350	0.892	0.142	-
	HWE	***P* value**	0.8521	0.9042	0.8828	0.8150	0.9385	0.7878	0.9442

Table 7. Genotype and allele frequencies of the c.499-17 G>A and c.528+38 C>T variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.449-17 G>A	**Genotype Distribution**	**n**	117	71	46	87	50	37	53
		GG	87 (0.744)	47 (0.662)	40 (0.870)	62 (0.713)	30 (0.600)	32 (0.865)	34 (0.642)
		GA	28 (0.239)	23 (0.324)	5 (0.109)	20 (0.230)	15 (0.300)	5 (0.135)	19 (0.358)
		AA	2 (0.017)	1 (0.014)	1 (0.021)	5 (0.057)	5 (0.100)	0	0
		***P* value**	0.191	0.646	0.010	0.073	0.060	-	-
	Allelic Distribution	**G**	0.863	0.824	0.924	0.828	0.750	0.932	0.821
		A	0.137	0.176	0.076	0.172	0.250	0.068	0.179
		***P* value**	0.3309	0.948	0.032	0.884	0.216	0.030	-
	HWE	***P* value**	0.8829	0.3259	0.1238	0.0697	0.1573	0.6594	0.1119
			Total Patients						
			Maternal			**Fetal**			**Controls**
			Total	**Mixed Ancestry**	**Black**	**Total**	**Mixed Ancestry**	**Black**	**Total**
c.528+38 C>T	**Genotype Distribution**	**n**	117	71	46	87	50	37	53
		CC	109 (0.932)	66 (0.930)	43 (0.935)	85 (0.977)	48 (0.960)	37 (1.000)	52 (0.981)
		CT	8 (0.068)	5 (0.040)	3 (0.065)	2 (0.023)	2 (0.040)	0	1 (0.019)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	0.966	0.965	0.967	0.989	0.980	1.000	0.991
		T	0.034	0.035	0.033	0.011	0.020	0	0.009
		***P* value**	0.188	0.191	0.248	0.870	0.527	0.402	-
	HWE	***P* value**	0.7018	0.7584	0.8192	0.9136	0.8853	-	0.9447

Table 8. Genotype and allele frequencies of the c.589-5 C>T and c.682+49 C>T variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.589-5 C>T	**Genotype Distribution**	**n**	120	72	48	89	51	38	52
		CC	115 (0.958)	69 (0.958)	46 (0.958)	82 (0.921)	47 (0.922)	35 (0.921)	52 (1.000)
		CT	5 (0.042)	3 (0.042)	2 (0.042)	7 (0.079)	4 (0.078)	3 (0.079)	0
		CC	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	0.979	0.979	0.979	0.961	0.961	0.961	1.000
		T	0.021	0.021	0.021	0.039	0.039	0.039	0
		***P* value**	0.138	0.139	0.139	0.041	0.041	0.041	-
	HWE	***P* value**	0.8157	0.8567	0.8828	0.6994	0.7707	0.8000	-
			Total Patients						
			Maternal			**Fetal**			**Controls**
			Total	**Mixed Ancestry**	**Black**	**Total**	**Mixed Ancestry**	**Black**	**Total**
c.682+49 C>T	**Genotype Distribution**	**n**	120	72	48	89	51	38	52
		CC	115 (0.958)	69 (0.958)	46 (0.958)	85 (0.955)	50 (0.980)	35 (0.921)	48 (0.923)
		CT	5 (0.0417)	3 (0.042)	2 (0.042)	4 (0.045)	1 (0.020)	3 (0.079)	4 (0.077)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	0.979	0.979	0.979	0.978	0.990	0.961	0.962
		T	0.021	0.021	0.021	0.022	0.010	0.039	0.038
		***P* value**	0.347	0.408	0.465	0.435	0.181	0.972	-
	HWE	***P* value**	0.8157	0.8567	0.8828	0.8283	0.9436	0.8000	0.7730

Table 9. Genotype and allele frequencies of the c.683-56 G>A and c.961-30 A>G variants identified in the *ANXA2* non-coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.683-56 G>A	Genotype Distribution	n	120	72	48	86	50	36	53
		GG	67 (0.558)	37 (0.514)	30 (0.625)	53 (0.616)	29 (0.580)	24 (0.667)	27 (0.509)
		GA	49 (0.408)	32 (0.444)	17 (0.354)	29 (0.337)	18 (0.360)	11 (0.305)	24 (0.453)
		AA	4 (0.034)	3 (0.042)	1 (0.021)	4 (0.047)	3 (0.060)	1 (0.028)	2 (0.038)
		P value	0.837	0.991	0.486	0.395	0.594	0.338	-
	Allelic Distribution	G	0.763	0.736	0.802	0.785	0.760	0.819	0.736
		A	0.238	0.264	0.198	0.215	0.240	0.181	0.264
		P value	0.596	0.996	0.266	0.348	0.690	0.194	-
	HWE	*P* value	0.1628	0.2218	0.4235	0.9896	0.9259	0.8449	0.2301
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.961-30 A>G	Genotype Distribution	n	118	71	47	90	52	38	53
		AA	117 (0.992)	71 (1.000)	46 (0.979)	90 (1.000)	52 (1.000)	36 (0.947)	51 (0.962)
		AG	1 (0.008)	0	1 (0.021)	0	0	2 (0.053)	2 (0.038)
		GG	0	0	0	0	0	0	0
		P value	-	-	-	-	-	-	-
	Allelic Distribution	A	0.996	1.000	0.989	1.000	1.000	0.974	0.981
		G	0.004	0	0.011	0	0	0.026	0.019
		P value	0.180	0.100	0.633	0.064	0.159	0.735	-
	HWE	*P* value	0.9631	-	0.9412	-	-	0.8677	0.8887

Table 10. Genotype and allele frequencies of the c.222 C>T; p.Ala74Ala and c.292 G>T; p.Val98Leu variants identified in the *ANXA2* coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.222 C>T; p.Ala74Ala	**Genotype Distribution**	**n**	123	78	45	96	58	38	53
		CC	123 (1.000)	78 (1.000)	45 (1.000)	96 (1.000)	58 (1.000)	38 (1.000)	50 (0.943)
		CT	0	0	0	0	0	0	3 (0.057)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	1.000	1.000	1.000	1.000	1.000	1.000	0.972
		T	0	0	0	0	0	0	0.028
		***P* value**	0.008	0.035	0.108	0.019	0.068	0.139	-
	HWE	***P* value**	-	-	-	-	-	-	0.8321
			Total Patients						
			Maternal			**Fetal**			**Controls**
			Total	**Mixed Ancestry**	**Black**	**Total**	**Mixed Ancestry**	**Black**	**Total**
c.292 G>T; p.Val98Leu	**Genotype Distribution**	**n**	119	71	48	77	43	34	52
		GG	114 (0.958)	66 (0.930)	48 (1.000)	75 (0.974)	41 (0.953)	34 (1.000)	50 (0.963)
		GT	5 (0.042)	5 (0.070)	0	2 (0.026)	2 (0.047)	0	2 (0.037)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**G**	0.979	0.965	1.000	0.987	0.977	1.000	0.981
		T	0.021	0.035	0	0.013	0.023	0	0.019
		***P* value**	0.915	0.456	0.172	0.690	0.847	0.250	-
	HWE	***P* value**	0.8149	0.7584	-	0.9081	0.8759	-	0.8876

Table 11. Genotype and allele frequencies for the c.600 C>T; p.Asp110Asp and c.934 G>A; p.Gly312Ser variants identified in *ANXA2* coding regions.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.600 C>T; p.Asp110Asp	**Genotype Distribution**	**n**	120	72	48	89	51	38	52
		CC	120 (1.000)	72 (1.000)	48 (1.000)	89 (1.000)	51 (1.000)	38 (1.000)	51 (0.981)
		CT	0	0	0	0	0	0	1 (0.019)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	1.000	1.000	1.000	1.000	1.000	1.000	0.990
		T	0	0	0	0	0	0	0.010
		***P* value**	0.128	0.238	0.335	0.190	0.321	0.391	-
	HWE	***P* value**	-	-	-	-	-	-	0.9442
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.934 G>A; p.Gly312Ser	**Genotype Distribution**	**n**	119	70	49	89	51	38	52
		GG	116 (0.975)	69 (0.986)	47 (0.959)	86 (0.966)	50 (0.980)	36 (0.947)	48 (0.923)
		GA	3 (0.025)	1 (0.014)	2 (0.041)	3 (0.034)	1 (0.020)	2 (0.053)	4 (0.077)
		AA	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**G**	0.987	0.993	0.980	0.983	0.990	0.974	0.962
		A	0.013	0.007	0.020	0.017	0.010	0.026	0.038
		***P* value**	0.120	0.088	0.450	0.260	0.181	0.654	-
	HWE	***P* value**	0.8892	0.9520	0.8841	0.8715	0.9436	0.8677	0.7730

Table 12. Genotype and allele frequencies for the variant c.975 C>T; p.Gly325Gly identified in *ANXA2* coding region and the variant c.*1057 A>G identified in the *ANXA2* 3'UTR.

			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.975 C>T; p.Gly325Gly	**Genotype Distribution**	**n**	118	71	47	90	52	38	53
		CC	116 (0.983)	69 (0.972)	47 (1.000)	89 (0.989)	51 (0.981)	38 (1.000)	51 (0.962)
		CT	2 (0.017)	2 (0.028)	0	1 (0.011)	1 (0.019)	0	2 (0.038)
		TT	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**C**	0.992	0.986	1.000	0.994	0.990	1.000	0.981
		T	0.008	0.014	0	0.006	0.010	0	0.019
		***P* value**	0.408	0.767	0.181	0.286	0.572	0.229	-
	HWE	***P* value**	0.9260	0.9042	-	0.9577	0.9442	-	0.8887
			Total Patients						
			Maternal			Fetal			Controls
			Total	Mixed Ancestry	Black	Total	Mixed Ancestry	Black	Total
c.*1057 A>G	**Genotype Distribution**	**n**	118	71	47	90	53	37	53
		AA	117 (0.991)	71 (1.000)	46 (0.979)	90 (1.000)	53 (1.000)	37 (1.000)	53 (1.000)
		AG	1 (0.009)	0	1 (0.021)	0	0	0	0
		GG	0	0	0	0	0	0	0
		***P* value**	-	-	-	-	-	-	-
	Allelic Distribution	**A**	0.996	1.000	0.989	1.000	1.000	1.000	1.000
		G	0.004	0	0.011	0	0	0	0
		***P* value**	0.502	-	0.287	-	-	-	-
	HWE	***P* value**	0.9631	-	0.9412	-	-	-	-

Appendix H: Comparison of the Genotype and Allele Frequencies in Different Ethnic Populations

Genotype and allele frequencies of previously documented variants identified in the *ANXA2* gene in this study were compared to the genotype and allele frequencies identified in European, Asia and Sub-Saharan African populations in Table 1. Information on the markers identified and included in the haplotype analysis is presented in Table 2.

Table 1. Genotype and allele frequencies of previously documented SNPs (http://www.ncbi.nlm.nih.gov/SNP/) identified in this study.

			Populations			**This study**	
Region	**Variant (dbSNP)**	**Genotypes & Alleles**	**European**	**Asian**	**Sub-Saharan African**	**Maternal PE Patients**	**Controls**
5'UTR	**rs12904657** **(c.-31 T>C)**	TT TC CC **T** **C**	Not Available	Not Available	Not Available	0.261 0.234 0.505 **0.379** **0.621** *(n=107)*	0.412 0.294 0.294 **0.559** **0.441** *(n=51)*
Coding	**rs1059688** **(c.929 G>T; p.Val98Leu)**	GG GT TT **G** **T**	Not Available	Not Available	Not Available	0.958 0.042 0.000 **0.979** **0.021** *(n=119)*	0.963 0.037 0.000 **0.981** **0.019** *(n=52)*
	rs1244554 **(c.975 C>T; p.Gly325Gly)**	CC CT TT **C** **T**	0.233 0.400 0.367 **0.433** **0.567** *(n=120)*	0.267 0.467 0.267 **0.500** **0.500** *(n=90)*	0.317 0.483 0.200 **0.558** **0.442** *(n=120)*	0.983 0.017 0.000 **0.992** **0.008** *(n=118)*	0.962 0.037 0.000 **0.981** **0.019** *(n=53)*

			Populations			This study	
Region	**Variant (dbSNP)**	**Genotypes & Alleles**	**European**	**Asian**	**Sub-Saharan African**	**Maternal PE Patients**	**Controls**
Non-Coding	**rs12904756 (c.-12+75 C>A)**	CC CA AA **C** **A**	Not Available	Not Available	Not Available	0.822 0.150 0.028 **0.897** **0.103** *(n=107)*	0.863 0.117 0.020 **0.922** **0.078** *(n=51)*
	rs11858864 (c.-11-43 G>A)	GG GA AA **G** **A**	0.094 0.151 0.755 **0.170** **0.830** *(n=106)*	0.000 0.000 1.000 **0.000** **1.000** *(n=88)*	0.113 0.302 0.585 **0.264** **0.736** *(n=106)*	0.102 0.534 0.364 **0.369** **0.631** *(n=118)*	0.130 0.407 0.463 **0.333** **0.667** *(n=54)*
	rs11855679 (c.-11-13 A>T)	AA AT TT **A** **T**	0.857 0.143 0.000 **0.929** **0.071** *(n=98)*	1.000 0.000 0.000 **1.000** **0.000** *(n=90)*	0.600 0.291 0.109 **0.745** ***0.255*** *(n=110)*	0.102 0.534 0.364 **0.369** **0.631** *(n=118)*	0.130 0.407 0.463 **0.333** **0.667** *(n=54)*
	rs3743268 (c.48+67 C>T)	CC CT TT **C** **T**	0.600 0.350 0.050 **0.775** **0.225** *(n=120)*	0.356 0.422 0.222 **0.567** **0.433** *(n=90)*	0.948 0.052 0.000 **0.974** **0.026** *(n=116)*	0.991 0.009 0.000 **0.996** **0.004** *(n=118)*	1.000 0.000 0.000 **1.000** **0.000** *(n=54)*
	rs12898604 (c.449-17 G>A)	GG GA AA **G** **A**	0.583 0.300 0.117 **0.733** **0.267** *(n=120)*	0.289 0.422 0.289 **0.500** **0.500** *(n=90)*	0.767 0.233 0.000 **0.883** **0.117** *(n=120)*	0.744 0.239 0.017 **0.863** **0.137** *(n=117)*	0.642 0.358 0.000 **0.821** **0.179** *(n=53)*
	rs11633619 (c.683-56 G>A)	GG GA AA **G** **A**	0.383 0.400 0.217 **0.583** **0.417** *(n=120)*	0.689 0.267 0.044 **0.822** **0.178** *(n=90)*	0.717 0.250 0.033 **0.842** **0.158** *(n=120)*	0.558 0.408 0.034 **0.763** **0.238** *(n=120)*	0.509 0.453 0.038 **0.736** **0.264** *(n=53)*

Table 2. Markers in accordance with variants in the *ANXA2* gene as imported into Haploview.

Marker	Variant
1	c.-442 C>G
2	c.-191 G>C
3	c.-189_-188insGCCGG
4	c.-135 C>G
5	c.-92 A>T
6	c.-31 T>C
7	c.-12+75 C>A
8	c.-11-43 G>A
9	c.-11-13 A>T
10	c.48+67 C>T
11	c.222 C>T; p.Ala74Ala
12	c.244-76 C>G
13	c.244-42 G>C
14	c.292 G>T; p.Val98Leu
15	c.449-17 G>A
16	c.528+38 C>T
17	c.589-5 C>T
18	c.600 C>T; p.Asp110Asp
19	c.682+49 C>T
20	c.683-56 G>A
21	c.934 G>A; p.Gly312Ser
22	c.961-30 A>G
23	c.961-24 C>G
24	c.975 C>T; p.Gly325Gly
25	c.*1057 A>G

Appendix I: Conference Outputs

Oral Presentation at South African Society of Human Genetics (SASHG), March 2009, Spier, South Africa

Unravelling the PP13-Annexin II Complex in Pre-eclampsia

De Jager JM[1], Bosman M[1,] Gebhardt GS[2], Meiri H[3], Marai S[3], Hillermann R[1]

[1]Department of Genetics, [2]Department of Obstetrics and Gynaecology, Stellenbosch University, South Africa; [3] Diagnostic Technologies Limited (DTL), Yokneam, Israel.

Objectives: Reduced levels of PP13 early in pregnancy are associated with a higher incidence of pre-eclampsia in later gestation. PP13 and Annexin II have been co-localised to the brush border membrane of syncytiotrophoblasts, and form a complex which is transported to the maternal circulation. It is speculated that genetic alteration in the gene encoding Annexin II (*ANXA2*) could underlie the reduced PP13 levels.

Methods: We report the screening of 6 of the 13 *ANXA2* exons in a cohort comprising 77 pre-eclampsia maternal-fetal pairs and 50 healthy control individuals. The various conformations identified on Multiphor SSCP/HD gels were subjected to automated sequencing, and subsequently to REA, to confirm the genotypes in the remainder of the cohort.

Results: Seven novel SNPs were identified, including intronic: IVS7+38 (C/T), IVS6-17 (A/G) [rs12898604], IVS8-5 (C/T), IVS9+49 (C/T), IVS12-30 (A/G), IVS12-24 (C/G) and exonic Ser96Gly (G/A), Thr97Ala (A/G) and Gly15 (C/T) [rs12442554] loci. Significant association was demonstrated at two loci: Ser96Gly was identified in 4% of the controls, but was absent in the maternal cohort (p=0.012), implying a protective effect or reduced pre-eclampsia risk. Marginal association (p=0.03) was demonstrated at intronic IVS8-5 C/T locus in the fetal cohort (4%), but was absent in the controls.

Conclusion: This pilot study provides evidence for genetic alteration in the *ANXA2* gene, which could impact on binding with PP13. Intronic splicing and coding variant analyses and their potential effects on gene expression, protein functionality and interaction with PP13 should be included in follow-up studies.