IMPROVED MUTATION DETECTION FOR HAEMOPHILIA A IN SOUTH AFRICA

Claire Lynne Mitchell

Dissertation submitted to the Faculty of Health Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine

2009

DECLARATION

I, Claire Lynne Mitchell, declare that this dissertation is my own work. It is being submitted for the degree Master of Science in Medicine in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

_____

Claire Lynne Mitchell

_____ day of _____, 2009

DEDICATION

This is dedicated to my husband, Gary, my family, Dad, Mom and Robyn Hetem and to the Mitchells (Dad, Ma, Peadar and Cathy), for all their love and support.

ABSTRACT

Haemophilia A is a common X-linked recessive bleeding disorder, affecting about 1 in 5000 males worldwide. It is caused by a deficiency of functional coagulation Factor VIII (FVIII), resulting in prolonged or abnormal bleeding episodes. The severity of the disease is related to the level of functional FVIII in the plasma. The *FVIII* gene is a large gene, located at Xq28 with a complex genomic organisation. It contains 26 exons spanning 186kb of genomic DNA, and produces a 9kb transcript, resulting in a functional protein of 2332 amino acids.

Over 900 mutations, which span a wide variety of categories, including rearrangements; complete or partial gene deletions; large insertions; duplications; frameshift mutations; splicing defects; nonsense and missense mutations, have been identified in the *FVIII* gene. Most mutations are rare or family specific, except for the intron 22 inversion mutation, which is reported to account for 45-50% of mutations in severe haemophilia A patients in most populations. A second inversion mutation, in intron 1, accounts for approximately 3.8% of haemophilia A patients in the UK. In South Africa, diagnostic mutation testing is currently only available for the intron 22 inversion mutation. Linked marker analysis is used to track high risk alleles in families where the disease-causing mutation is unknown.

This study aims to evaluate an mRNA-based method to identify disease-causing mutations in South African haemophilia A patients and improve the diagnostic service. Blood samples from 120 patients were tested first for the intron 22 and then for intron 1 inversion mutations. Inversion negative patients were analysed further using mRNA.

A mutation has been identified in 73.3% (88/120) of all patients. 30% (36/120) of patients had the intron 22 inversion, 2.5% (3/120) an intron 1 inversion and 40.8% (49/120) of patients had a mutation identified by mRNA analysis. A mutation was not identified in the remaining 26.7% (32/120) due to sample and technical difficulties.

Of the 49 mutations identified through mRNA analysis, 28 patients (57.1%) have a point mutation (17 missense (34.7%), 9 nonsense (18.4%) and 2 splice-site

mutations (4.1%)), 9 patients (18.4%) have a deletion and 7 patients (14.3%) have an insertion. Another 5 patients (10.2%) have a complex mutation (including patients where an exon deletion was detected on mRNA analysis, but no mutation was identified on DNA analysis). One mutation, c.3637insA, was found recurrently in 14% (6/43) of patients from the white population. This single base insertion results in a frameshift mutation with a premature stop codon at amino acid 1221 (only translating about half the normal FVIII protein). This common mutation, together with haplotype analysis, suggests a founder effect for this mutation.

mRNA analysis of the *FVIII* gene is a novel technique in mutation detection for haemophilia A. It decreases the costs involved in sequencing the coding region and it offers improved mutation detection compared to DNA analysis.

Diagnostic testing in South Africa should be extended from the current intron 22 inversion mutation to include DNA analysis for the intron 1 inversion and the founder mutation (c.3637insA) in white patients, followed by mRNA testing, starting with the analysis of the fragments spanning exon 14. mRNA analysis identifies an additional 55.7% of mutations compared to conventional diagnostic testing for the intron 22 inversion alone.

ACKNOWLEDGEMENTS

I would like to extend a special thank-you to the following people and organisations, without whom, I would not have been able to complete this project.

Thank you to Professor Amanda Krause for your dedicated encouragement, support and guidance during this project. I really appreciate all you have taught me. I hope to continue to learn from you for many more years to come. You are a true mentor.

To Doctor Johnny Mahlangu, Sister Bongi, Sister Anne and Kabelo in the Haemophilia clinic at the Johannesburg Hospital, and to Dr Brittain and Darryl at the Mayo Clinic. Thank you for the genuine kindness and compassion you show to the patients and for collecting the patient samples for this project.

The Medical Research Council of South Africa for generous funding of this project

To the staff and students who have been part of the Division of Human Genetics in the past few years. May you be blessed with the happiness and success you deserve.

I would like to thank my wonderful family. You have supported me in striving to do my best and never giving up, no matter what. You inspire me and I am blessed and proud to be part of our family. I love you so much more than words could ever express.

Last but definitely not least, to my fantastic husband, Gary. How do I even begin to thank you? For all the sacrifices you made so that I could follow my dreams, for affording me the opportunity to complete my studies on a full time basis, for the weekends and time you sacrificed, for all the support and encouragement and for being the best husband and friend I could ever have asked for. THANK YOU! I love you with all my heart.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

A	Adenine
Ala	Alanine
APS	Ammonium persulphate
aPTT	Activated partial thromboplastin time
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
bp	Base pairs
C	Cytosine
CFC	Clotting factor concentrates
CSGE	Conformational sensitive gel electrophoresis
Cys	Cysteine
ddH_2O	Deionised distilled water
ddNTPs	Dideoxynucleoside triphosphates
dH_2O	Distilled water
dHPLC	Denaturing high performance liquid chromatography
DGGE	Denaturing gradient gel electrophoresis
dNTPs	Deoxynucleoside triphosphates
EDTA	Ethylene diamine tetra acetic acid
EtBr	Ethidium bromide
FVIII	Coagulation factor VIII
g	Gram
G	Guanine
Gln	Glutamine
Glu	Glutamic acid
Gly	Glycine
His	Histidine
Ile	Isoleucine
kb	Kilobases
kDa	Kilo Daltons
L	Litre
Leu	Leucine
Lys	Lysine
mCi/μl	Millicurie per microlitre
Met	Methionine
μl	Microlitre
mg	Milligram
MHC	Major histocompatibility complex
MIBS	Malmö international brother study

ml	Millilitre
MLPA	Mulitplex ligation-dependent probe amplification
mM	Millimolar
MW	Molecular weight
NHLS	National Health Laboratory Service
PCR	Polymerase chain reaction
Phe	Phenylalanine
Pro	Proline
PT	Prothrombin time
q	Longer arm of a chromosome
RFLP	Restriction fragment length polymorphism
r.p.m	Revolutions per minute
SDS	Sodium dodecyl sulphate
Ser	Serine
SNPs	Single nucleotide polymorphisms
STRs	Short tandem repeats
T	Thymine
Taq	Polymerase from the bacterium *Thermus aquaticus*
TBE	Tris borate EDTA
TE	Tris-EDTA
TEMED	*N-N-N'-N'*-tetramethylenediamine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

1 INTRODUCTION

1.1 INTRODUCTION TO THE CHAPTER

This chapter includes an introduction to haemophilia A, including the general background and history of the disorder, inheritance and prevalence, clinical symptoms and a description of the different severities. The current treatment options for haemophilia and potential side effects from treatment are also discussed.

The role of factor VIII in the blood coagulation pathway is discussed. The *FVIII* gene and the wide variety of mutations occurring within the gene are reviewed. The different methods for DNA testing and the current DNA testing options available in South Africa are explained.

The aim of this project is to optimise and implement an mRNA approach for mutation detection for haemophilia A in South Africa, with the intention of improving the diagnostic service, and assisting female relatives at risk of haemophilia by offering them conclusive carrier testing and prenatal diagnosis.

1.2 HISTORY OF HAEMOPHILIA A

Haemophilia A is one of the most common inherited bleeding disorders, caused by a deficient or non-functional coagulation Factor VIII (FVIII) (reviewed in Bhopale & Nanda, 2003).

Haemophilia encompasses both haemophilia A and haemophilia B. Although the two are clinically similar, they can be differentiated by laboratory testing, as described in section 1.6 below. This study focuses on haemophilia A.

The clinical manifestations of haemophilia have been recognised for many centuries. During the second century, the Talmud stated that male babies with two brothers, who died after excessive bleeding after being circumcised, need not undergo the procedure themselves. However, the first medical report of

haemophilia was only in 1803 by Otto in which he described a family showing symptoms spanning three generations (reviewed in McKusick, 1965).

Although the specific type was unknown, haemophilia was well documented in the European royal family. Queen Victoria of England (1837-1901) passed on a mutation to the royal families of Spain, Germany and Russia (reviewed in McKusick, 1965). The pedigree of the European royal family (reviewed in Stevens, 1999) can be seen in Figure 1.1 below. A paternal age effect may have been responsible for the disease as there was no other family history of the disorder, and Queen Victoria's father was 52 years old when she was born (McKusick, 1965).

1.3 INHERITANCE AND PREVALENCE OF HAEMOPHILIA A

Haemophilia A is a common X-linked recessive bleeding disorder that affects about 1 in 5000 males worldwide, with no ethnic or geographical predisposition (reviewed in Pandey & Mittal, 2001). The high prevalence of haemophilia is due to the X-linked mode of inheritance and the high mutation rate of the *FVIII* gene (Strauss, 1967; Vogel, 1977). Although the prevalence does not vary much between populations, it could be underestimated in developing countries due to poor diagnosis (reviewed in Graw, Brackmann, Oldenburg, *et al.*, 2005).

The majority (approximately 60%) of haemophilia A cases have a family history, however, the remainder are sporadic due to *de novo* mutations (reviewed in Graw, *et al.*, 2005). In 90% of these families, the mutation has arisen in the parents or grandparents (Becker, Schwaab, Möller-Taube, *et al.*, 1996). It has been reported that mutations causing haemophilia A originate 3.6 times more frequently in male germ cells than female germ cells. However, this sex ratio of mutation frequencies differs for each type of mutation. The mutation rate for point mutations is 5-10-fold higher and inversions more than 10-fold higher in male germ cells, while deletions are more than 5-fold higher in female germ cells (reviewed in Oldenburg, Ananyeva & Saenko, 2004b). Therefore a mother of a sporadic case of haemophilia has a high risk of being a carrier of the disorder.

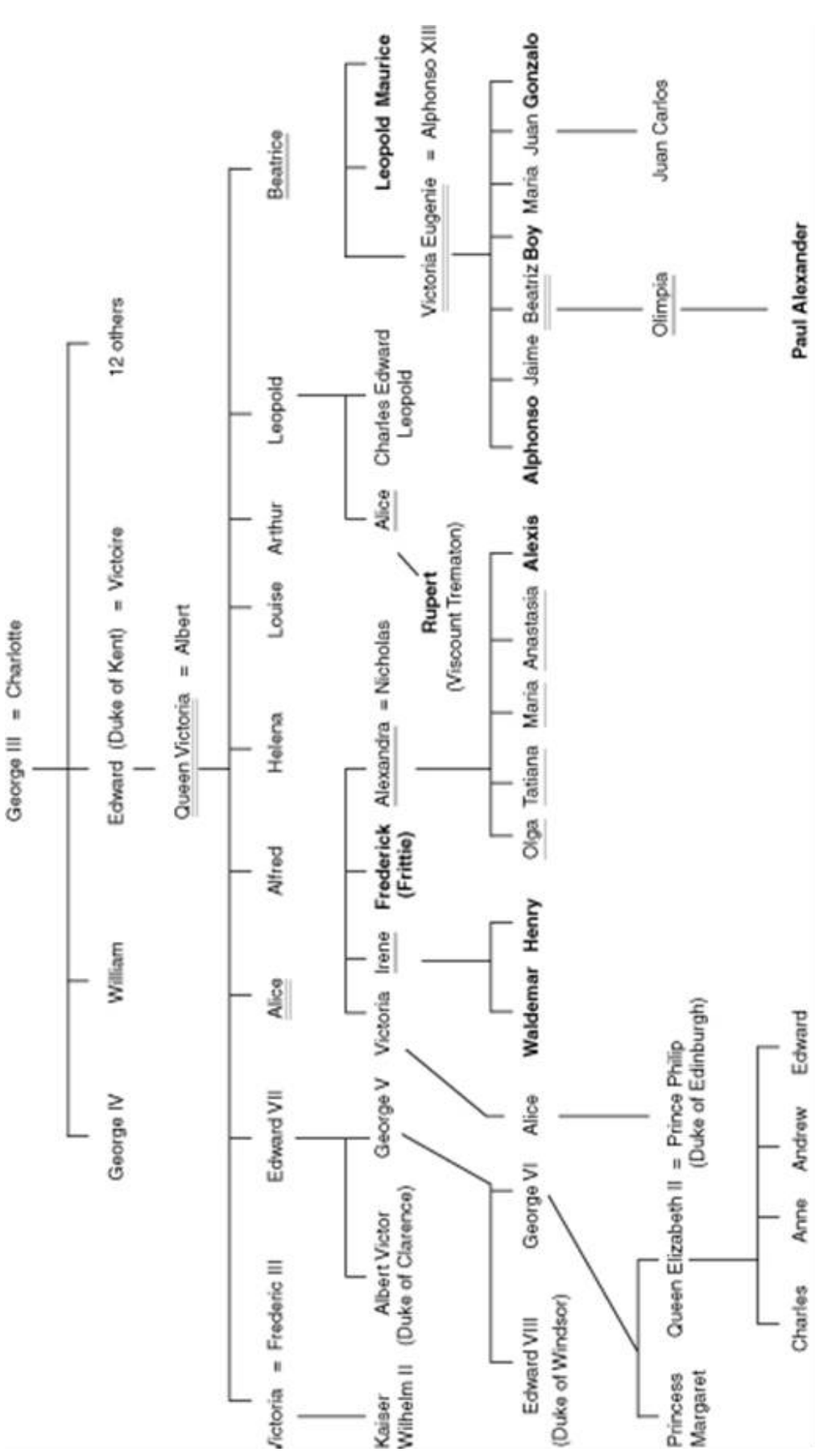

Figure 1.1: A pedigree of the European royal family showing the individuals with haemophilia.
Bold indicates individuals affected with haemophilia, double underlined indicates obligate carriers, single underlined individuals are possible haemophilia carriers (reviewed in Stevens, 1999).

As with all X-linked recessive conditions, there is a 50% chance that a female carrier will pass the defective gene to her child (male or female) (refer to Figure 1.2a). While all female offspring born to a haemophiliac father are obligate carriers, all his sons will be unaffected (refer to Figure 1.2b).

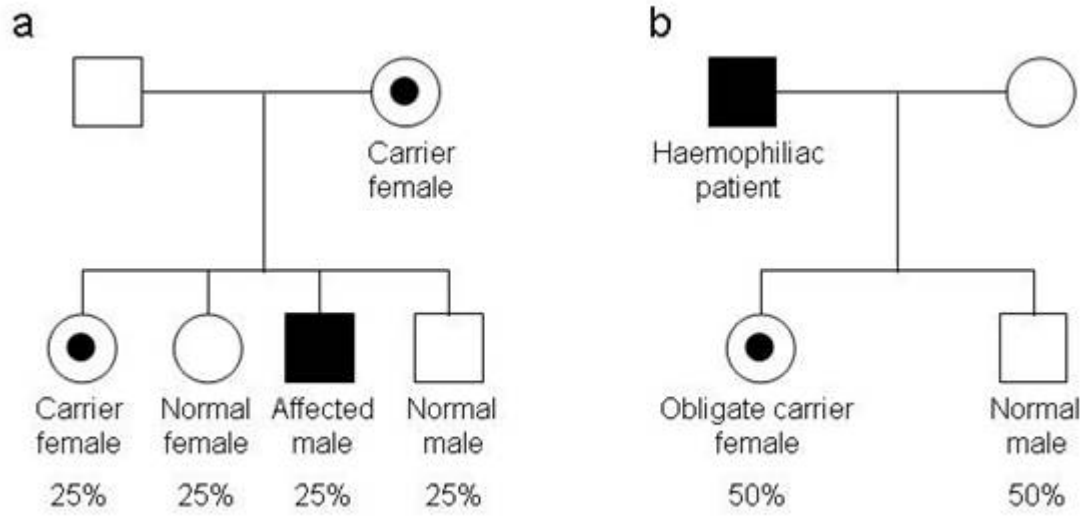

Figure 1.2: Inheritance of Haemophilia A, from an obligate carrier female (a) and an affected male (b).

Although haemophilia A is an X-linked recessive condition, females can present with symptoms in the following instances: skewed X inactivation (e.g. Coleman, Genet, Harper, *et al.*, 1993); inheritance of two mutant *FVIII* genes, one from each parent (e.g. Pola & Svojitka, 1957; Windsor, Lyng, Taylor, *et al.*, 1995); X-autosome translocation disrupting the *FVIII* gene (e.g. Migeon, Axelman, Jan de Beur, *et al.*, 1989) or leading to preferential inactivation of the unaffected X-chromosome; or a female with Turner Syndrome and a mutation in the *FVIII* gene on her only X chromosome.

In developing countries, most patients with haemophilia are in the paediatric age group as they seldom reach adulthood due to inadequate treatment (Evatt & Robillard, 2000). They suffer from the pain of bleeding, the morbidity of orthopaedic deformity and finally death (reviewed in DiMichele, Chuansumrit, London, *et al.,* 2006).

In South Africa in 2007, there were more than 1760 haemophilia patients on the National Haemophilia Register (Mahlangu & Gilham, 2007). In 2003, Hazewinkel, Hoogerwerf, Hesseling, *et al.,* recorded the incidence of the disorder in the Western Cape as 7/100 000. They also describe the mean age of diagnosis for severe patients as 9 months, while moderate haemophiliac patients were diagnosed at 11 months. In addition, they state that, even with therapy, haemophilia causes significant morbidity in South African patients and their families. However, little is known about care of patients outside of the haemophilia treatment centres.

1.4 CLINICAL SYMPTOMS OF HAEMOPHILIA A

Haemophilia A is characterised by a deficiency or decrease in function of factor VIII which results in prolonged bleeding episodes. Typical clinical symptoms include: delayed bleeding and prolonged oozing after injuries, surgeries or tooth extractions, or renewed bleeding after initial bleeding has stopped (Arun & Kessler, 2001).

The clinical presentation and severity of the disease is directly related to the level of functional FVIII activity (reviewed by Kulkarni & Lusher, 2001). The severity is greatly influenced by the mutations in the *FVIII* gene (reviewed in Pandey & Mittal, 2001) and the proportion of severe, moderate and mild haemophilia A cases are about 50%, 10% and 40% respectively (Antonarakis, Rossiter, Young, *et al.*, 1995). However, Rizza, Spooner & Giangrande (2001) suggests a bias towards patients with severe haemophilia A in most studies as they are likely to attend clinic more frequently. The three types of haemophilia A are described in more detail below.

1.4.1 Severe Haemophilia A

Severe haemophilia A patients present with less than 1% factor VIII levels and frequently present with spontaneous bleeds into joints, muscles and internal

organs or abnormal bleeding after minor injuries. These patients are usually diagnosed in the first year of life and have an average of two to five spontaneous bleeding episodes per month (reviewed by Johnson & Thompson, 2005).

1.4.2 Moderate Haemophilia A

Moderately severe patients have between 1% and 5% factor VIII levels and have prolonged bleeding or delayed oozing after minor trauma or surgery, but seldom have spontaneous bleeding episodes. These patients are usually diagnosed before the age of five or six years and the frequency of the bleeding episodes can vary from once a month to once a year (reviewed by Johnson & Thompson, 2005).

1.4.3 Mild Haemophilia A

Mildly affected patients do not have spontaneous bleeding episodes but have abnormal bleeding only after significant trauma, surgery and major injuries. Individuals with mild haemophilia A have 5-35% of residual factor VIII activity and the frequency of bleeding may vary from once a year to once every ten years. These patients are often not diagnosed until later in life (reviewed by Hoyer, 1994; Johnson & Thompson, 2005).

1.5 Blood Coagulation Pathway and Role of Factor VIII

Factor VIII is a plasma glycoprotein cofactor that serves as an essential component in the intrinsic blood coagulation pathway (Kaufman, 1992; Spiegel, Murphy & Stoddard, 2004). The blood coagulation pathway is an intricate process involving a cascade of coagulation factors and proteins which eventually form a blood clot. Essentially, an injury stimulates vasoconstriction and platelets to adhere to the damaged blood vessel wall and to each other at the site of injury (primary haemostasis), mediated by von Willebrand factor. The aggregating platelets release chemicals that initiate blood clotting via the intrinsic or extrinsic

pathways of the coagulation cascade to create a fibrin network to seal the damage in the blood vessel (secondary haemostasis) (summarised by Nathwani & Tuddenham, 1992; Voet & Voet, 1995).

Upon activation, and in the presence of calcium ions and phospholipid surfaces, factor VIII and factor IX form an active complex, which activates factor X. Subsequent stages of the cascade proceed, resulting in the deposits of fibrin, the structural polymer of the blood clot (reviewed in Bowen, 2002). A schematic representation of the coagulation cascade is shown in Figure 1.3 below (Bowen, 2002). If any of the factors involved in the coagulation pathway are deficient or defective, this would result in a condition of abnormal bleeding, as discussed in section 1.6 below.

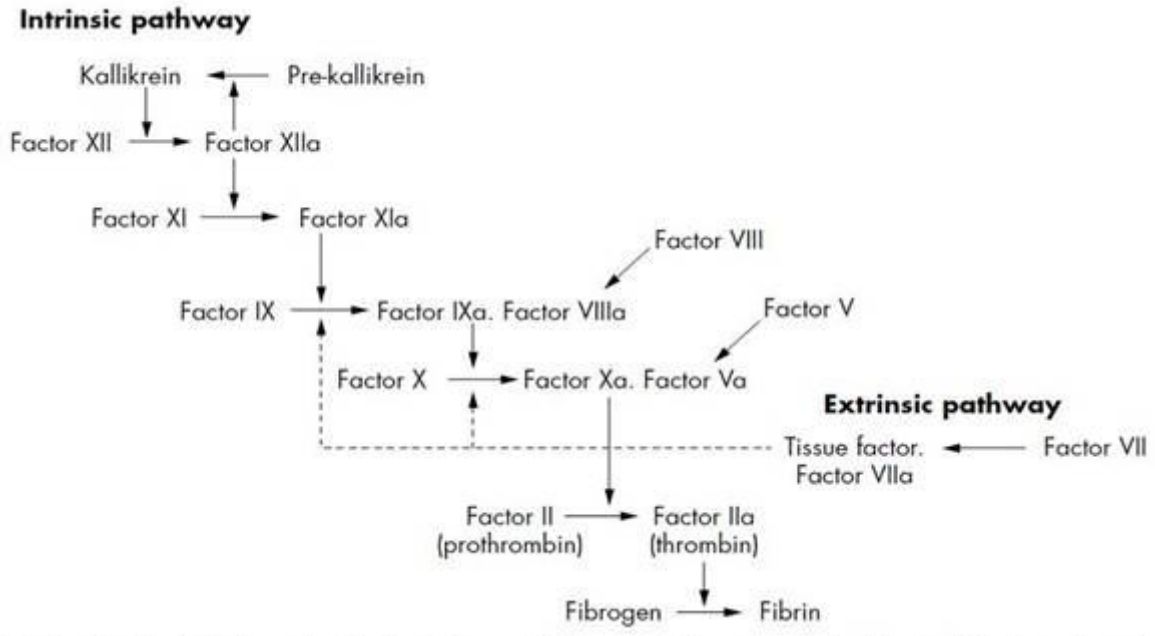

Figure 1.3: A schematic representation of the intrinsic and extrinsic pathways of the coagulation cascade leading to fibrin formation.
Reviewed in Bowen (2002)

1.6 TREATMENT OF HAEMOPHILIA A

1.6.1 Current Treatment Options

As haemophilia A is caused by an absence or functionally abnormal FVIII (reviewed in Bhopale & Nanda, 2003), treatment of haemophilia involves replacement with either plasma-derived or recombinant clotting factor concentrates. Recombinant factor concentrates are preferred over plasma-derived replacements as the infectious agent transmission risk is reduced (reviewed in DiMichele, *et al.,* 2006).

Treatment may be given for several reasons: in an attempt to keep the factor VIII at a normal level to prevent "spontaneous" bleeding episodes, given before surgery to prevent excessive bleeding or after trauma.

Although factor concentrate provides effective haemostasis, there are some contra-indications to the product. Factor concentrates have a short half-life so they need to be intravenously injected 2-3 times per week and patients have a 20-30% chance of developing inhibitors to FVIII. Factor concentrates also involve very high life-long medical expenditure and restrictive global access (reviewed in DiMichele, 2005; Srivastava, 2005). Tuddenham (2005) estimated the life time cost of replacement therapy for one haemophilia patient to be £5 000 000.

Laurian (2005) reviewed the research aiming to develop better replacement therapies with an increased protein stability to facilitate prophylaxis of one injection per week, or less (Pipe & Kaufman, 1997; Ananyeva, Kouiavskaia, Shima, *et al.*, 2001; Lillicrap, 2005); to reduce inhibitor formation (Barrow, Healey, Gailani, *et al.*, 2000); for infused factor concentrate therapies (reviewed in Hoots, 2005) and the possibility of oral instead of intravenous administration (Hemker, Hermans, Muller, *et al.*, 1980). Also, production of the factor concentrates in larger quantities may make haemophilia A treatment cheaper (Brettler, 2005).

Haemophilia A treatment in the future may include gene therapy, cell-based treatments and the development of novel clotting factor molecules, with a variety of potential biological advantages (reviewed in DiMichele, *et al.,* 2006). Research into the potential for gene therapy is discussed in section 1.6.2.

1.6.2 Gene Therapy for Haemophilia A

Gene therapy involves the treatment of inherited disease by the addition, insertion or replacement of a normal gene (Mueller & Young, 1998). Gene therapy raises many ethical issues because current treatments are now much safer than previously and gene therapy carries an unknown risk of long-term adverse effects (reviewed in Laurian, 2005). However it has the potential to provide long-term endogenous production of FVIII, especially if made affordable to the majority of haemophilia patients (reviewed in DiMichele, 2005).

In theory, gene therapy offers potential cost savings with fewer treatments versus a life time cost for one patient's replacement therapy (reviewed in Tuddenham, 2005) and haemophilia gene therapy does not require organ specificity to achieve sufficient protein expression (White & Samulski, 2005). However, inhibitor development is unlikely to change with gene therapy (Brettler, 2005).

Many gene therapy trials for haemophilia A have used vectors like adeno-associated viruses, adenoviruses and retroviruses. Adeno-associated viral vectors show long-term positive results in mouse and dog models but human trials have been disappointing, with low levels of circulating FVIII and only transient efficacy (Snyder, Miao, Meuse, *et al.*, 1999; Scallan, Lillicrap, Jiang, *et al.*, 2003; Giangrande, 2004; Negrier, 2004; reviewed in Laurian, 2005). DiMichele, *et al.* (2006) and Tuddenham (2005) reviewed gene therapy with adenovirus and retrovirus vectors. Adenoviruses have been associated with acute negative effects including hepatotoxicity and thrombocytopenia and even death in a patient with ornithine transcarbamylase deficiency treated with an adenovirus vector (DiMichele, *et al.*, 2006; Tuddenham, 2005). Retrovirus vectors are relatively well tolerated by the immune system but there are concerns about the long-term risk of insertional mutagenesis by integration into the host genome (DiMichele, *et al.*, 2006; Tuddenham, 2005). For example, there is a report of leukaemia development in three children treated for X-linked severe combined immunodeficiency, due to insertional mutagenesis by a retrovirus vector, one of whom subsequently died (Hacein-Bey-Abina, Von Kalle Schmidt, *et al.*, 2003). In two recent reviews (Gabrovsky & Calos, 2008; Mannucci, 2008), no successful gene therapy trials have been reported.

Most gene therapy trials to date have involved adult subjects (White, 2001), although the ultimate goal is to develop treatments for children (reviewed in DiMichele, *et al.,* 2006). Ethicists believe that children and young adults below the age of 25, should not be recruited into gene therapy trials as safety, especially long-term, has not been demonstrated and is not guaranteed (Brettler, 2005).

1.6.3 Haemophilia A Treatment in South Africa

In South Africa, patients attending clinics at a haemophilia treatment centre (in one of the 16 centres in Johannesburg, Cape Town, Port Elizabeth, East London, Durban, Pretoria, Bloemfontein, Potchefstroom, Mthatha and Polokwane) are treated with prophylactic factor replacement therapy. The amount of factor administered at each treatment is dependent of the weight of the patient. The recommendations are as follows: for less than 20kg require 300U, 20-40kg need 800U, 40-54kg requires 1000U, 55-67kg need 1300U and patients between 67kg and 80kg requires 1500U. This treatment regimen facilitates the appropriate management of haemophilia and reduces the associated complications such as musculoskeletal deformities and death (Mahlangu & Gilham, 2007).

Treatment products available for patients without inhibitors include: plasma-derived, solvent detergent-treated FVIII (such as Haemosolvate Factor VIII concentrate from National Bioproducts Institute), desmopressin for mild/moderate haemophilia patients (such as 1-deamino-8-D-arginine vasopressin (DDAVP) from Stimate®), antifibrinolytic and adjunctive agents (Mahlangu & Gilham, 2007). Recombinant human FVIII, Kegonate (FVIII), was registered with the South African regulatory authority in October 2008 (Dr Mahlangu, personal communication, 2009).

In South Africa the 2008/2009 cost for 1000U of Haemosolvate factor VIII is R1986.19 (2008/2009 price list obtained from National Bioproducts Institute). As per current guidelines, an average adult male with severe haemophilia A, weighing between 67kg and 80kg would require 1500U per treatment (at R3007.82). Discovery Health, the biggest medical aid insurer in South Africa, only covers a maximum of R14 000 per month in their chronic illness benefits for 2009

(https://www.discovery.co.za/contentSources/logged_out/pdfs/2007_cib_formulary.pdf). This means that a patient is only covered for just over 4 treatments per month.

For patients with FVIII inhibitors, treatment products available are recombinant FVIIa (for example, NovaSeven® from Novo Nordisk) or an activated prothrombin complex concentrate known as FEIBA® (Factor Eight Inhibitor Bypassing Agent by Adcock Ingram Critical Care) together with antifibrinolytic and adjunctive agents (Mahlangu & Gilham, 2007).

Hazewinkel, *et al.* (2003) reported the majority of haemophilia patients (73%) in the Western Cape are treated with factor “on demand” after a bleed. Some patients (20%) with recurrent severe bleeds in joints are given clotting factor three times a week as “periodic prophylaxis” until the joint has improved and a few patients (7%) receive “continuous prophylaxis” with Factor from shortly after birth.

1.7 FACTOR VIII INHIBITORS

Approximately 20-35% of severe haemophilia A patients, and occasionally patients with a milder phenotype, develop antibodies to exogenous FVIII derived from FVIII replacement therapy (Lusher, Arkin, Abildgaard, *et al.*, 1993; Hay, 1998; Darby, Keeling, Spooner, *et al.* 2004). These antibodies, or inhibitors, bind to infused FVIII and reduce its half-life and neutralise the coagulation activity (Mannucci & Tuddenham, 2001).

Inhibitor formation is a complex process and is reviewed by Astermark (2006). Several factors (genetic, environmental and treatment-related) are thought to influence inhibitor development. These are described in more detail below.

1.7.1 Genetic Factors

Genetic factors are believed to be involved in inhibitor development as siblings have a higher risk (50%) of developing inhibitors than members of an extended haemophiliac family (9%) (Gill, 1999). The Malmö International Brother Study

(MIBS) found that inhibitor concordance was 78% in siblings, significantly higher than in unrelated patients (Astermark, Berntorp, White, *et al.*, 2001). Interestingly, a study by Scharrer, Bray & Neutzling (1999) found that the inhibitor incidence was 51.9% in African-Americans compared to 25.8% in Caucasians, suggesting patients with African heritage are at higher risk of developing inhibitors.

Genetic factors which influence inhibitor development with FVIII replacement therapy include: the type of causative *FVIII* gene mutation (Oldenburg, Schroder, Brackmann, *et al.*, 2004a); the MHC class I and II phenotype (and polymorphisms in the genes encoding cytokines and immune regulatory molecules (reviewed in Astermark, 2006).

1.7.2 Environmental and Treatment Related Factors

In monogenic conditions, it is expected that monozygotic twins will present with the same phenotype. However, this is not the case for inhibitor development in haemophilia A, indicating that non-genetic factors influence inhibitor risk (Astermark, *et al.*, 2001). Some of these environmental factors include age when treatment was initiated, immune system challenges, type of factor concentrate used and mode of administration.

1.8 LABORATORY DIAGNOSIS OF HAEMOPHILIA A

There are other disorders that are clinically similar to haemophilia A, for example haemophilia B (factor IX deficiency) or von Willebrand disease (due to a deficiency of von Willebrand factor). It is important to accurately diagnose and differentiate between the specific types of bleeding disorders, to determine the therapeutic intervention required and for accurate genetic counselling. A diagnosis is made using the following tests.

1.8.1 Coagulation Screening Test

This involves a platelet count, bleeding time or a platelet function analysis, activated partial thromboplastin time (aPTT, to measure the intrinsic pathway) and prothrombin time (PT, to measure the extrinsic or tissue-factor pathway) (reviewed by Dahlbäck, 2000; Johnson & Thompson, 2005). The diagnostic interpretation of the coagulation screening test results as suggested by the World Federation of Haemophilia (2005) can be found in Table 1.1.

Table 1.1: Interpretation of the coagulation screening test for different bleeding disorders.

Possible condition	Platelet count	Bleeding time	PT	aPTT
Normal	Normal	Normal	Normal	Normal
Haemophilia A or B	Normal	Normal	Normal	Prolonged
Von Willebrand disease	Normal or reduced	Normal or prolonged	Normal	Normal or prolonged
Platelet defect	Normal or reduced	Normal or prolonged	Normal	Normal

1.8.2 Coagulation Factor Assays

Since haemophilia A and B are indistinguishable on coagulation screening tests, their diagnosis must be confirmed using coagulation factor assays. These assays measure the amount of a specific coagulation factor, for example factor VIII, factor IX or von Willebrand factor. The normal range for factor VIII clotting activity is 50-150%. In haemophilia A, the factor VIII activity is less than 35% with normal factor IX and von Willebrand factor levels. Severe haemophilia A is diagnosed with less than 1% factor VIII activity, while moderate and mild patients have a factor VIII clotting activity of 1-5% and 5-35%, respectively. About 10% of female haemophilia A carriers have a factor VIII clotting activity less than 35%, indicating that coagulation factor assays are unreliable in detecting the majority of female carriers (reviewed by Johnson & Thompson, 2005). The FVIII:vWF ratio is often decreased in female haemophilia A carriers and was routinely used for carrier

detection, however, the accuracy of this test is limited and variable among individual laboratories (Rimoin, Connor, Pyeritz, *et al.*, 2007).

1.8.3 Molecular Genetic Testing

Coagulation screening tests and coagulation factor assays are generally sufficient to make a clinical diagnosis of haemophilia A. However, the detection of carrier females for haemophilia A (as with most X-linked recessive conditions) is relatively difficult because most females are asymptomatic and the information provided by a coagulation assay can be equivocal due to random X-chromosome inactivation (reviewed in Pandey & Mittal, 2001). DNA based analysis (using either direct mutation detection or indirect/linked marker analysis, discussed in section 1.11.1 and 1.11.2 respectively) is able to accurately detect carriers and provide prenatal testing for at risk individuals with 95% confidence (reviewed in Pandey & Mittal, 2001).

Although DNA testing is a more costly diagnostic tool, the benefits include: a definitive diagnosis, accurate carrier testing, prenatal diagnostic options and identifying symptomatic carriers that may require treatment (Sorenson, Jennings-Grant & Newman, 2003; Ludlam, Pasi, Bolton-Maggs, *et al.*, 2005). Knowledge of the disease-causing mutation may be useful in predicting the disease severity, the likelihood of inhibitor formation and even selecting the best candidates for different gene therapies in the future (reviewed in DiMichele, *et al.,* 2006).

1.9 THE FACTOR VIII GENE AND PROTEIN

1.9.1 The *FVIII* Gene

The factor VIII gene (*FVIII*) is a large gene with a complex genomic organisation (reviewed in Pandey & Mittal, 2001). The gene was cloned between 1982 and 1984 by Gitschier, Wood, Goralka, *et al.* (1984), and is located in the most distal band of the long arm of the X chromosome (Xq28) (Poustka, Dietrich, Langenstein, *et al.* 1991; Freije & Schlessinger, 1992) (refer to Figure 1.4,

obtained from http://www.genecards.org/pics/loc/GC0XM153717.F8.png). The Human Genome Organisation (HUGO) database (http://www.genenames.org/) uses the *F8* gene symbol, however, the Haemophilia A Mutation, Structure, Test and Resource Site (HAMSTeRS) (http://europium.csc.mrc.ac.uk/Main/main.htm) uses *FVIII*. The *FVIII* symbol will be used in this dissertation.

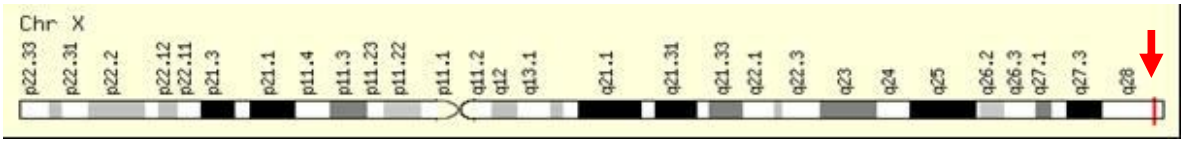

Figure 1.4: A schematic representation of the position of the *FVIII* gene (arrow) on the X chromosome.

The *FVIII* gene spans 186kb of genomic DNA and consists of 26 exons encoding an mRNA transcript of approximately 9kb (Toole, Knopf, Wozney, *et al.*, 1984; Wood, Capon, Simonsen, *et al.*, 1984). It encodes a 2351 amino acid polypeptide (Vehar, Keyt, Eaton, *et al.*, 1984), from which 19 amino acids are cleaved off the N-terminal, to produce the mature, functional protein of 2332 amino acids (reviewed in Graw, *et al.*, 2005). The genomic organisation of the *FVIII* gene is represented schematically in Figure 1.5.

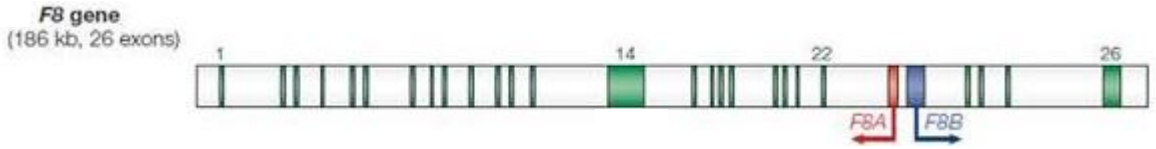

Figure 1.5: A schematic representation of the *FVIII* gene.
The relative positions and sizes of the exons (green) and introns (white) of the *FVIII* gene, and the F8A and F8B genes in intron 22 (in red and blue respectively) with their transcription orientation are shown (adapted from Graw, *et al.*, 2005).

The *FVIII* gene is unusual in that within intron 22 (IVS22) a CpG island exists that is associated with two additional transcripts, F8A (Levinson, Kenwrick, Lakich, *et al.*, 1990) and F8B (Levinson, Kenwrick, Gamel, *et al.*, 1992) as can be seen in Figure 1.5. F8B is a 2.7kb transcript and is transcribed in the same direction as the

FVIII gene, using a private exon and FVIII exons 23 to 26. The function of the F8B transcript and its potential translated product remains unknown (reviewed in Graw, *et al.*, 2005). The F8A gene spans 2kb, contains no introns and is transcribed in the opposite orientation to the *FVIII* gene. It encodes a 40-kDa Huntingtin-associated protein (Peters & Ross, 2001), which is thought to be involved in the aberrant nuclear localisation of the huntingtin protein in Huntington Disease (reviewed in Graw, *et al.*, 2005). Two homologous copies of F8A have been found approximately 400kb telomeric to the FVIII gene (Levinson, *et al.*, 1990). The significance of the F8A transcript and its homologues will be described in more detail in section 1.10.1.1 below.

Although most factor VIII protein synthesis takes place in the liver, other tissues do contain detectable levels of FVIII mRNA (Wion, Kelly, Summerfield, *et al.*, 1985).

1.9.2 Activation of the FVIII Protein

FVIII is a large glycoprotein with three distinct domain types, in the arrangement (NH_2)-A_1-A_2-B-A_3-C_1-C_2-(COOH) (Vehar, *et al.*, 1984; Lenting, van Mourik & Mertens, 1998; O'Brien, Mastri & Fay, 2000), as can be seen in Figure 1.6. Activation of FVIII involves proteolysis by thrombin or FXa, which cleaves the FVIII protein at Arg1689 (between the B and A_3 domain), producing a heavy (A_1-A_2-B) and light (A_3-C_1-C_2) polypeptide chain (Eaton, Rodriguez & Vehar, 1986). The heavy and light chain heterodimers are covalently bound to a copper ion, although the role of the copper in the structure and function of FVIII is unclear (reviewed in Bhopale & Nanda, 2003; Fang, Wang & Wang, 2007).

The heavy chain is cleaved further at Arg740 (between A_2 and B domain) releasing the B domain from the polypeptide and at Arg372 (between the A_1 and A_2 domains) to produce the A_1 and A_2 polypeptides of 50 and 43kDa in size, respectively. The light chain is 73kDa (Eaton, *et al.*, 1986). The presence of the 50, 43 and 73kDa polypeptides correlates with weak thrombin activation (Fulcher, Roberts & Zimmerman, 1983; Eaton, *et al.*, 1986).

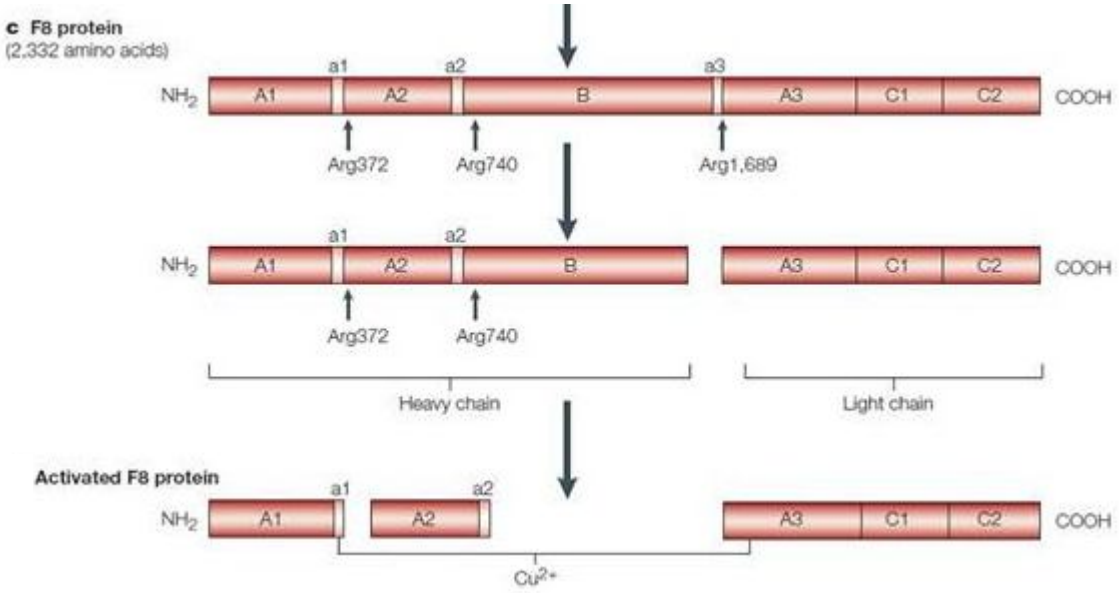

Figure 1.6: A schematic representation of the FVIII protein.
The translated native/precursor protein is shown, which after cleavage by thrombin, becomes activated into the mature FVIII protein (adapted from Graw, *et al.*, 2005).

The inactive precursor form of FVIII binds to and circulates with von Willebrand factor (vWF). Cleavage at the Arg1689 residue releases FVIII from vWF (Pittman & Kaufman, 1988) and permits the interaction of vWF with a phopsholipid surface (Walker, Chavin & Fay, 1987; Brandstetter, Bauer, Huber, *et al.*, 1995) and FVIII with platelets (reviewed in Fang, *et al.*, 2007) (refer to the blood coagulation pathway in section 1.6).

1.10 MUTATIONS IN THE *FVIII* GENE

Over 900 mutations causing haemophilia A have been identified in the *FVIII* gene. These were obtained from the results of genetic analysis of over 2500 DNA samples from haemophilia A patients examined for molecular defects using a variety of methods. These results were compiled into a database of *FVIII* gene mutations and published by Tuddenham, Cooper, Gitschier, *et al.* (1991). This database is summarised, continually updated and available online at the Haemophilia A Mutation, Structure, Test and Resource Site (HAMSTeRS) at

http://europium.csc.mrc.ac.uk/Main/main.htm (Tuddenham, Schwaab, Seehafer, *et al.*, 1994).

Haemophilia A disease-causing mutations span a wide variety of categories, including rearrangements; complete or partial gene deletions; large insertions; duplications; frameshift mutations; splicing defects; nonsense and missense mutations (Tuddenham, *et al.*, 1994). In a review, Antonarakis, *et al. (*1995) analysed data from an international consortium and found that point mutations account for 46% of *FVIII* mutations, inversions for 42% and deletions account for 8% of severe haemophilia A cases.

Many of these mutations are rare or family specific, except for two relatively common recurrent inversion mutations occurring in intron 22 and intron 1 of the *FVIII* gene. These two inversion mutations, which occur fairly frequently, are described in section 1.10.1 below.

1.10.1 Gene Rearrangements

Large gene rearrangements are responsible for 40-50% of all severe Haemophilia A cases (Antonarakis, *et al.,* 1995).

1.10.1.1 Intron 22 Inversion

While studying the mRNA from haemophilia A patients, Naylor, Brinke, Hassock, *et al.*, (1993) found that approximately half of the patients in their cohort did not have amplification across intron 22. It was later determined that this non-amplification was due to an inversion mutation. As discussed in section 1.9.1, intron 22 of the *FVIII* gene contains an additional transcript, the *F8A* gene (*int22h-1*). Two additional copies of the *F8A* gene, a distal or telomeric copy (*int22h-3*) and a proximal or centromeric copy (*int22h-2*), are located approximately 400kb and 500kb telomeric to the *FVIII* gene (Lakich, Kazazian, Antonarakis, *et al.*, 1993; Naylor, *et al.*, 1993; Naylor, Buck, Green, *et al.*, 1995). The intron 22 inversion mutation results from an intrachromosomal crossover event between the intragenic copy (*int22h-1*) and one of the two extragenic copies (*int22h-2* or *int22h-3*) (Lakich, *et al.*, 1993; Naylor, *et al.*, 1993; Bagnall, Waseem, Green, *et al.*,

2002). As a result, the *FVIII* gene is separated into two parts in opposite orientations, causing the mRNA to be interrupted so that synthesis of the FVIII protein is incomplete (Maugard, Tuffery, Aguilar-Martinez, *et al.*, 1998).

This intron 22 inversion mutation accounts for approximately 45-50% of severe haemophilia A patients' mutations (Lakich, *et al.*, 1993; Naylor, *et al.*, 1993; Bagnall, *et al.*, 2002). In a study of 2093 severe haemophilia A patients, Antonarakis, *et al.* (1995) found that the distal copy (*int22h-3*) is involved in 35% of cases, the proximal copy (*int22h-2*) in 7%, while just over 1% of cases involved an additional abnormal pattern. Although most inversion mutations involve *int22h-1* and either *int22h-2* or *int22h-3*, rare, atypical inversions have been reported (Lakich, *et al.*, 1993; Naylor, Nicholson, Goodeve, *et al.*, 1996; Yamazaki, Mohri, Inaba, *et al.*, 1997). These atypical inversions are thought to be due to a rearrangement between *int22h*-1 and a third, truncated copy of the *F8A* gene within the *FVIII* gene (reviewed in Bowen, 2002).

For many years, it was believed that *int22h-2* and *int22h-3* were orientated in the same direction as each other, but in opposite orientation to *int22h-1*, because either copy can be involved in the intron 22 inversion mutation (Lakich, *et al.*, 1993). Sequencing of the human genome, especially the X chromosome, revealed that *int22h-2* and *int22h-3* are actually in opposite orientations to each other, and that *int22h-2* (the proximal copy) is in the same orientation as *int22h-1* (Ross, Grafham, Coffey, *et al.*, 2005).

It has subsequently been found that *int22h-2* and *int22h-3* are part of an imperfect palindrome (as illustrated in Figure 1.7a, adapted from Bagnall, Giannelli & Green, 2006). Bagnall, *et al.* (2006) noted that if all X chromosomes had the arrangement suggested by Ross, *et al.* (2005), recombinations between *int22h*-1 and *int22h*-2 would result in deletions or duplications due to intra-chromosomal rearrangements. Bagnall, *et al.* (2006) proposed that the arms of this palindrome undergo intra-chromosomal recombination, thereby changing the orientation of *int22h-2* and *int22h-3* relative to *int22h-1* (refer to Figure 1.7b and c). They further suggested that this *int22h-2* and *int22h-3* rearrangement is polymorphic.

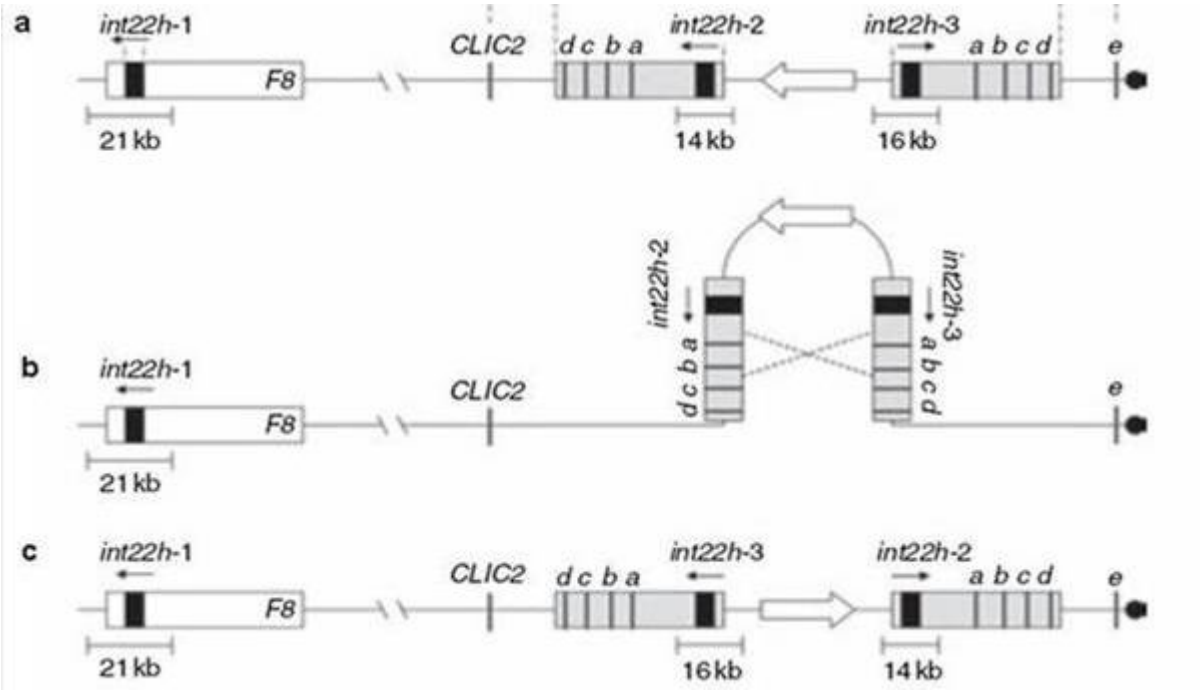

Figure 1.7: The proposed mechanism causing the palindrome inversion polymorphism between *int22h*-2 and *int22h*-3.

(a) shows the common arrangement, while (b) shows the proposed intra-chromosomal recombination resulting in the inverted palindrome configuration (c). Obtained from Bagnall, *et al.* (2006)

Bagnall, Giannelli, Green, *et al.* (2005) also illustrated the possible results of recombinational events between the different *int22h* copies, with respect to the *int22h-2* and *int22h-3* palindromic polymorphism. The distal (type I) and proximal (type II) intron 22 inversion mutations resulting from intra-chromosomal recombinations are illustrated in Figure 1.8A and B below. Inter-chromosomal rearrangements between mis-aligned *int22h* copies may cause deletions or duplications, as illustrated in Figure 1.8C and D below.

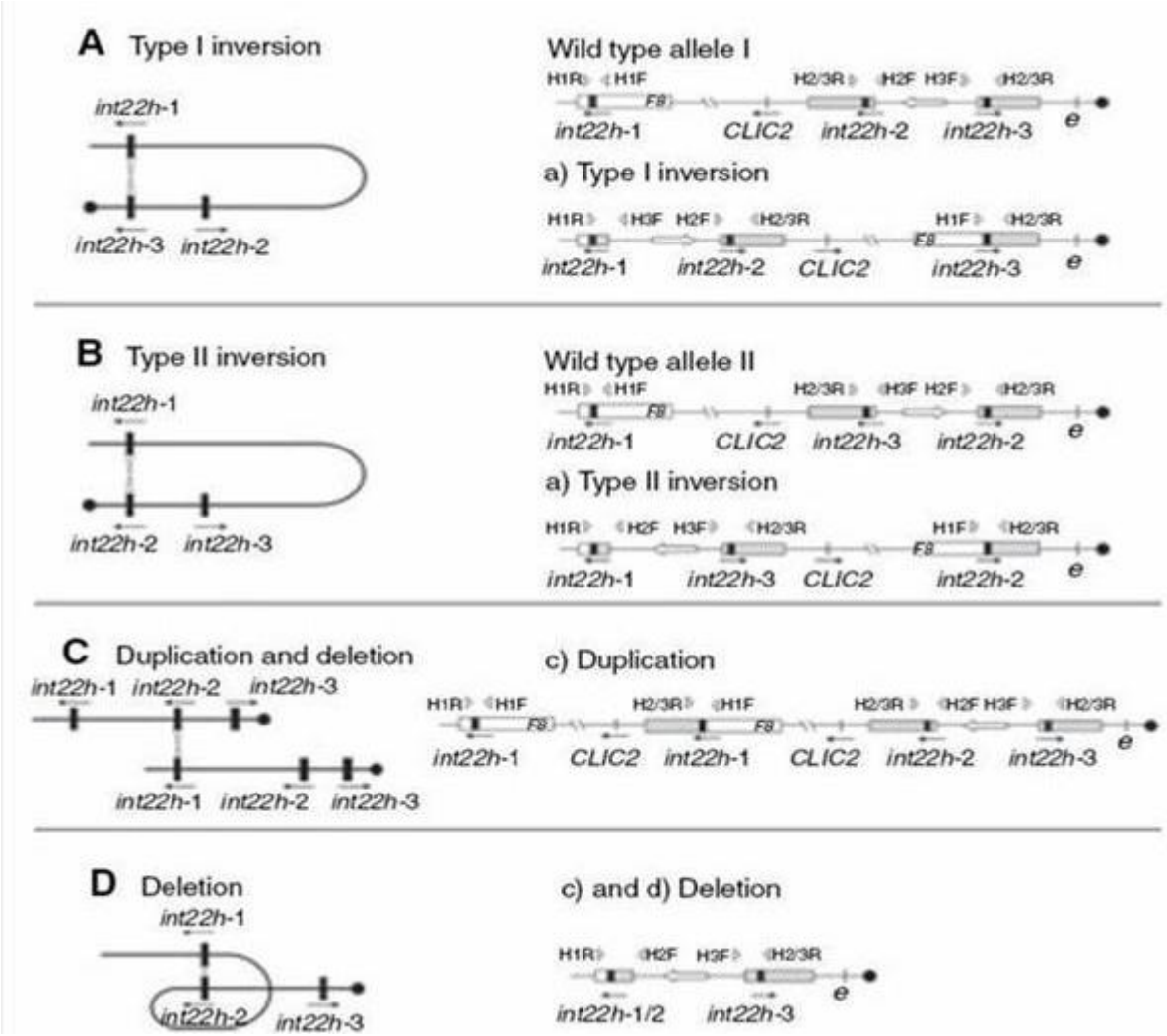

Figure 1.8: An illustration of the possible recombinational events between the different *int22h* copies.

(A) and (B) show the proposed intra-chromosomal rearrangements causing *int22h*-related inversions. (C) and (D) show the inter-chromosomal or intra-chromosomal recombinations between mis-aligned *int22h* repeats causing deletions or duplications. The left panel shows the mode of rearrangement while the right panel shows the structure of wild type and mutant alleles. Obtained from Bagnall, *et al.* (2005).

Family studies have found that the intron 22 inversion occurs almost exclusively during spermatogenesis from a normal male (Rossiter, Young, Kimberland, *et al.*, 1994), usually the patient's maternal grandfather (Antonarakis, *et al.*, 1995).

In South Africa, Dangerfield, Manga, Field, *et al.* (1997) found the intron 22 inversion mutation to account for 43% (6/14) of black and 32% (13/41) of white haemophilia A cases. All six of the black patients had the type I type inversion,

while in the white group, 69% (9/13) were of the type I type and 31% (4/13) were type II inversions. The frequency of the intron 22 inversion mutation in the South African white population group is one of the lowest reported (Antonarakis, *et al.*, 1995). The type II inversion has subsequently been identified in black South African patients and these results will be presented in section 3.2.1.2.

1.10.1.2 Intron 1 Inversion

An intrachromosomal inversion between a 1kb sequence in intron 1 (*int1h-1*) of the *FVIII* gene and a second sequence (*int1h-2*), repeated in the reverse orientation approximately 140kb telomeric to the *FVIII* gene occurs in about 4.8% of families with severe haemophilia A in the United Kingdom (Bagnall, *et al.*, 2002). This mutation was found in 3.0% of Chinese patients (Liang, Yan, Xioa, *et al.*, 2008); 2.7% of Indian patients (Ahmed, Ivaskevicius, Kannan, *et al.*, 2005), and 15% of severe haemophilia A patients in the Netherlands (Boekhorst, Verbruggen, Lavergne, *et al.*, 2005).

1.10.2 Single Base Substitutions

Single base mutations or point mutations include missense mutations (which result in a single amino acid change), nonsense mutations (change an amino acid into a stop codon) and mRNA splice site or splice junction mutations (which change the true splice site or create a new one, resulting in exon skipping or the insertion of intronic sequence into the mRNA sequence) (reviewed in Bowen, 2002). In August 2007, there were 615 different single base substitutions reported, of which, 75% were missense mutations, 15% were nonsense mutations and 10% splice junction alterations (HAMSTeRS database).

1.10.3 Deletions

In a 2003 review of the molecular pathology of haemophilia A on the HAMSTeRS mutation database, the database lists over 270 deletion mutations in the *FVIII* gene. These deletions may be micro-deletions of 1bp or more (causing a frame

shift mutation), partial gene deletions or deletions of 210kb (deleting the entire gene). Most of these mutations result in frameshift mutations which are associated with severe haemophilia A (HAMSTeRS database; reviewed in Bowen, 2002).

1.10.4 Insertions

The 2003 review of the mutation database describes over 57 different insertions varying in size from 1bp up to 2.1kb as well as a 3.8kb retrotransposon or LINE element (Kazazian, Wong, Youssoufian, *et al.*, 1988). Single base insertions are most common. All insertions are associated with severe disease as they either result in a frameshift mutation or gross insertions which adversely affect the gene product (HAMSTeRS database; reviewed in Bowen, 2002).

1.11 Molecular Genetics Testing for Haemophilia A

DNA testing for haemophilia A, using either direct mutation detection (section 1.11.1) or indirect or linked marker analysis (section 1.11.2), is used for confirmation of diagnosis (in affected patients), carrier testing (in at risk female relatives) and prenatal testing. Direct mutation detection provides a more accurate diagnosis of haemophilia A than coagulation screening tests and coagulation factor assays, particularly in carriers and for prenatal diagnosis. However, the highly heterogeneous nature of mutations and the complexity of the *FVIII* gene can make direct mutation detection difficult (reviewed in Pandey & Mittal, 2001). Indirect detection or linked marker analysis may be more useful in these cases.

DNA mutation results must be interpreted with caution, as there is a potential risk of somatic mosaicism in families with sporadic haemophilia A of approximately 10% which could cause uncertainty regarding recurrence risks in mothers who appear to be mutation negative (Leuer, Oldenburg, Lavergne, *et al.*, 2001; reviewed in Peyvandi, Jayandharan, Chandy, *et al.*, 2006).

DNA testing is important for prenatal diagnosis, as it influences obstetric decisions, delivery options, and early care of haemophilia, even if the family decide not to terminate the pregnancy (Baty, Drayna, Leonard, *et al.*, 1986).

1.11.1 Direct Mutation Detection

Direct detection (detecting the disease-causing mutation) is used in haemophilia A with high accuracy and a high detection rate in haemophilia A families. It is able to detect mutations in familial and sporadic cases, even if the proband is unavailable (reviewed in Peyvandi, *et al.* 2006).

Currently in South Africa, direct mutation analysis is only available for the intron 22 inversion mutation in the *FVIII* gene (refer to section 1.10.1.1). Detection of this mutation involves the use of Southern blot analysis as described in Lakich, *et al.* (1993). Genomic DNA is digested with the *Bcl* I restriction enzyme and the fragments of interest are detected using the p482.6 probe, labelled with ^{32}P (Oldenburg, Rost, El-Maarri, *et al.*, 2000). All three copies of *int22h* are distinguished as different sized fragments as follows: *int22h-1* (*FVIII* gene fragment) is 21.5kb; *int22h-2* (proximal copy) is 14kb; while *int22h-3* (distal copy) is 16kb (Naylor, *et al.*, 1995). The type of inversion can be determined by the band patterns observed by Southern blotting (reviewed by Goodeve, 1998), as shown in Figure 1.9.

Bagnall, *et al.* (2006) recently developed a single-tube, long-range PCR assay to test for the intron 22 inversion mutation. Primers were specifically designed to amplify each copy of *int22h* (*int22h*-1, *int22h*-2 and *int22h*-3) in 10kb, 11kb and 12kb fragments respectively, which are resolved on a 0.5% agarose gel. This assay is able to distinguish between the different types of inversions in patients and carrier females, and is able to identify deletions and duplications resulting from recombinations between *int22h* sequences.

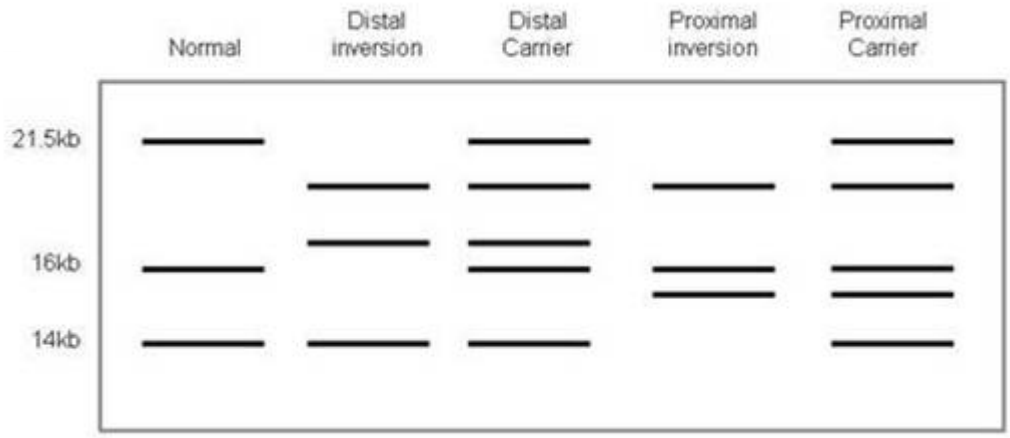

Figure 1.9: A diagrammatic representation of the intron 22 inversion mutation, showing schematic results obtained from Southern blot analysis after digestion with *Bcl* I, with the expected sizes for both types of inversion mutation.

1.11.1.1 Mutation Detection Strategies

Many centres perform mutation pre-screening that rely on heteroduplex formation and the subsequent detection of mismatched heteroduplexes (reviewed by Keeney, Mitchell & Goodeve, 2003). The three main heteroduplex detection methods include denaturing gradient gel electrophoresis (DGGE) (Higuchi, Antonarakis, Kasch, *et al.*, 1991); conformational sensitive gel electrophoresis (CSGE) (Ganguly, Rock, & Prockop, 1993) and denaturing high performance liquid chromatography (dHPLC) (Oefner & Underhill, 1995). Another method of pre-screening includes protein truncation testing (PTT) which detects truncating mutations at the protein level (Maugard, *et al.* 1998); however, missense mutations would not be detected using this method.

Aberrant products obtained from these pre-screening techniques are then resolved by DNA sequencing, the gold standard for mutation detection (reviewed by Keeney, *et al.*, 2003).

Multiplex ligation-dependent probe amplification (MLPA) is a relatively new technique that detects copy number quantification, such as deletion or duplication mutations. As mutations and/or polymorphisms very close to the probe ligation site may result in a false positive MLPA result, all apparent single exon

deletions/duplications require confirmation using a different technique. Point mutations would not be detected using MLPA (http://www.mlpa.com/WebForms/WebFormProductDetails.aspx?Tag=tz2fAPIAupLo0TRBqtiGHA||&ProductOID=kZ3txtB8eJg|)

Another method of mutation detection involves the analysis of mRNA. This method has proved successful in previous studies, for example Naylor, *et al.* (1992) detected mutations in all 28 subjects studied, and the study also resulted in the discovery of the intron 22 inversion mutation. A mRNA-based approach to assess the transcribed *FVIII* gene, will therefore detect a wide range of mutations such as large rearrangements, deletions, insertions, mutations affecting splicing and point mutations at a reduced cost due to the large size of the *FVIII* gene at a DNA level.

1.11.2 Indirect Detection or Linked Markers

Linked marker analysis uses two types of markers, namely: those that can be detected using a restriction enzyme (or RFLP markers) and short tandem repeat sequences (STRs). STR markers result from repeat differences in a stretch of di- or tri-nucleotide repeats and are distributed randomly throughout the whole genome (reviewed in Pandey & Mittal, 2001).

Certain polymorphic markers within or close to the *FVIII* gene (to reduce the chance of recombination between the gene and the marker) are used to track the inheritance of the high-risk X chromosome in families affected with haemophilia A (reviewed in Pandey & Mittal, 2001). Linkage needs DNA from an affected proband to determine the high-risk X-chromosome, and the mother of the proband must be informative (heterozygous) for at least one of the polymorphic linked markers tested.

Linkage analysis is rapid, relatively inexpensive to perform and 99% reliable in familial cases of haemophilia A. However, linkage analysis can only exclude a female relative from being a carrier in sporadic cases (which account for up to 40% of haemophilia A cases) (reviewed in Peyvandi, *et al.* 2006) as a high risk allele cannot detect where the mutation arose in a sporadic case. Also some

families may not be informative for any of the linked markers tested, and depending on the marker, there may be a chance of recombination between the gene and the marker (reviewed in Pandey & Mittal, 2001).

In South Africa, the Diagnostic Molecular Genetics Laboratory at the Division of Human Genetics, National Health Laboratory Service (NHLS) and School of Pathology, the University of the Witwatersrand in Johannesburg uses three STR markers for haemophilia A linked marker analysis, namely two intragenic markers (in introns 13 and 22 of the *FVIII* gene) (Lalloz, McVey, Pattinson, *et al.*, 1991; Lalloz, Schwaab, McVey, *et al.*, 1994) and one extragenic marker (p39), 500kb 3' to the *FVIII* gene (Wehnert, Reiner & Caskey, 1993).

A study by Dangerfield, *et al.* (1997) found the combination of the two intragenic markers (intron 13 and intron 22) to be 100% informative in black families (5 families) and 67% (24 families) informative in the white group. When the p39 extragenic marker was included, the linked marker informativity increased to 91% of white South African patients (McKibbin, 2002).

Further, the findings of Dangerfield, *et al.* (1997) suggested the presence of a founder haemophilia A haplotype in the Afrikaner population of alleles 20 and 26 repeats for the intron 13 and intron 22 markers respectively, thereby suggesting a possible common founder mutation. However, no common founder mutation was identified by sequencing the coding region of the *FVIII* gene (Dangerfield & Manga, unpublished data). The high incidence of other disorders such as porphyria variegata (Dean, 1971) and familial hypercholesterolemia (Jenkins, Nicholls, Gordon, *et al.*, 1980) in the Afrikaner population has been attributed to founder effect.

1.12 AIMS OF THE PROJECT

1.12.1 Motivation for the Project

Currently, the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand,

Johannesburg is only able to offer direct mutation detection for the intron 22 inversion mutation for haemophilia A. As this mutation only accounts for 32-43% of mutations in patients with haemophilia A in South Africa (section 1.10.1.1), the disease-causing mutation in the remaining patients remains unidentified. This impacts on carrier testing and prenatal diagnosis options for family members at risk of haemophilia A, especially if the family is uninformative on linked marker analysis.

Knowledge of the disease-causing mutation may have clinical significance for appropriate patient management, as it may indicate disease severity, may help predict patients at risk of developing inhibitors to FVIII replacement therapy (section 1.7) and in the future, could identify prospective candidates for gene therapy (section 1.6.2).

mRNA analysis for haemophilia A has proved successful in previous studies of haemophilia A mutation detection, for example Naylor, *et al.* (1992) detected mutations in all 28 subjects studied, and the study also resulted in the discovery of the intron 22 inversion mutation. mRNA analysis has also proved useful in mutation detection in other diseases such as Duchenne and Becker Muscular Dystrophy (for example, Roest, Bout, van der Tuijn, *et al.*, 1996), breast and ovarian cancers (examples include Spearman, Sweet, Zhou, *et al.*, 2008; Rhei, Bogomolniy, Federici, *et al.*, 1998), and neurofibromatosis (an example described in Pros, Gómes, Martín, *et al.*, 2008).

As DNA-based analysis has not been as useful as was expected (Dangerfield & Manga, unpublished data) and is very expensive because of the large size of the *FVIII* gene, it is hoped that an mRNA-based approach to assess the transcribed *FVIII* gene, will detect a wide range of mutations such as large rearrangements, deletions, insertions, mutations affecting splicing and point mutations at a reduced cost.

1.12.2 The Specific Aims of the Project

1. To develop and assess an mRNA-based mutation detection strategy for *FVIII* gene mutations.

2. To improve the molecular genetic diagnostic service available to haemophilia A patients and their families.

3. To identify the potential founder mutation in the Afrikaans population.

2 SUBJECTS AND METHODS

2.1 INTRODUCTION TO THE CHAPTER

This chapter describes the subjects obtained for the project and the methods used to analyse the samples. Fresh blood from which mRNA and DNA was extracted for analysis was collected from patients with haemophilia A. The patients were analysed for the intron 22 inversion mutation. The mRNA was reverse transcribed into more stable cDNA, amplified by PCR and sequenced. All sequences were compared to a reference sequence to detect mutations. These mutations were compared to an online mutation database to determine whether they were previously reported mutations or polymorphisms. An attempt was made to confirm all mutations on the patients' DNA sample.

2.2 SUBJECTS

Ethics approval for this study was obtained from the Human Research Ethics Committee (Medical) at the University of the Witwatersrand, Johannesburg, with the protocol number of M040801. A copy of the ethics clearance certificate can be found in Appendix A1.

Data were collected and analysed from 120 haemophilia A patients. The majority of the subjects used in this study were ascertained through the Haemophilia Clinic at the Johannesburg Hospital. The distribution of where the samples were collected from is as follows:

- 110 patients from the haemophilia clinic at the Johannesburg Hospital
- 4 patients from the haemophilia clinic at Chris Hani Baragwanath Hospital
- 4 patients through private practice (Dr D. Brittain, Mayo Clinic, Floracliff)
- 2 patients from Groote Schuur Hospital, Cape Town.

A blood sample of 5-10ml fresh blood was collected from these patients after informed consent was obtained from either the patients or the parents of the

patients if the patient, was under 16 years of age. The consent and assent forms can be found in Appendices A2 and A3 respectively.

Samples were collected from all South African population groups. Due to the small sample size of the Indian and mixed ancestry patients, they were grouped together as “others” for ease of analysis, as, separately, the sample numbers were too small to draw any conclusions. The population breakdown of the patient cohort was as follows:

56 white patients
56 black patients
4 Indian patients
4 mixed ancestry patients.

As suggested by Rizza, *et al.* (2001) there is a bias towards sample collection from patients with severe haemophilia as they are likely to attend clinics more frequently. This is evident in the distribution of the disease severity in these patients. The severity of haemophilia A in this cohort is:

97 severe disease
13 moderate disease
5 mild disease and
5 unknown severity.

In a pilot study, 34 intron 22 inversion negative patients from the diagnostic cohort were tested for the intron 1 inversion mutation, and consisted of 25 white patients, seven black patients and two patients from the Indian and mixed ancestry populations. These patients will be referred to as the diagnostic haemophilia A cohort (refer to section 3.2.2).

Screening for the c.3637insA mutation was undertaken in all 30 of the white families from the diagnostic cohort who tested negative for the intron 22 inversion mutation, however, mutation results were only obtained for 23 of these patients. These patients will be referred to as the white haemophilia A families in the diagnostic database (refer to section 3.2.6.2).

2.3 METHODS

The composition of all reagents and solutions used in the techniques below can be found in Appendix D, unless otherwise stated.

2.3.1 DNA Extraction

DNA was extracted from 3-10ml of patient blood, collected in EDTA Vacutainers, using a modified protocol of the salting-out method, as described by Miller, Dyk, Pelesky, *et al.* in 1988. The detailed protocol can be found in Appendix B1. DNA samples were stored at 4°C.

2.3.2 RNA Extraction

RNA was extracted within 24 hours of sample collection from 1ml of fresh patient blood, collected in EDTA Vacutainers, using the QiAamp RNA Blood Mini kit (QIAGEN) using the protocol supplied with the kit. A detailed description of the extraction process can be found in Appendix B2. All the RNA samples were stored at -70°C. These samples were found to be stable for up to one year.

An mRNA extraction was attempted on a single blood sample stabilised with TRIzol (Invitrogen Life Technologies). The protocol for stabilising RNA in blood using TRIzol was supplied by Dr Marius Coetzee (Department of Haematology and Cell Biology, University of the Free State, Bloemfontein) and can be found in Appendix B2. The TRIzol extraction protocol can also be found in Appendix B2. The mRNA extracted from the TRIzol stabilised blood amplified for all the cDNA fragments, and therefore offers a viable option for sample transport of samples from other centres around the country (discussed further in section 4.3.2.1.1).

2.3.3 Detection of the Intron 22 Inversion Mutation

The detection of the intron 22 inversion is performed routinely in the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg using Southern Blot analysis as described by Levinson, *et al.* (1990). Inversion testing using Southern blot analysis was performed as part of the study for all 120 patients.

Southern blotting studies the organisation of genes within genomes. Genomic DNA is digested with one or more restriction enzymes and the resulting fragments are separated, according to size, by agarose gel electrophoresis. The DNA is denatured *in situ* and transferred from the gel to a nylon membrane and the relative positions of the DNA fragments are preserved. The DNA, attached to the membrane, is hybridised to a labelled oligonucleotide probe and the bands complementary to the probe are located by the autoradiography (reviewed by Sambrook & Russell, 2001).

2.3.3.1 Digestion of Genomic DNA

To detect the intron 22 inversion, approximately 10µg of genomic DNA was digested with 30U *Bcl* I restriction enzyme (Roche) in a reaction containing 1x reaction buffer M (Roche) and 0.05M spermidine, with a final volume made up to 50µl with ddH_2O. This reaction was incubated overnight at 50°C. 5µl of the digested products were checked on a trial gel to ensure adequate DNA concentration and sufficient digestion of the DNA. Once this was confirmed, the remaining digested sample (45µl) was mixed with 5µl Ficoll dye and loaded on a 0.6% agarose gel without ethidium bromide. A 1kb marker and a MWMII marker (Roche) were used as the size standards. The gel was run at 1.6V/cm for two nights to ensure adequate separation of the products of interest. The progress of the gel was checked after the first night by staining the gel in a solution of 100ml 1x TBE buffer containing 10µl ethidium bromide. After sufficient separation of the products (between 16kb and 20kb) was obtained, the gel was prepared for Southern transfer.

2.3.3.2 Southern Transfer

The gel was soaked in depurinating solution for 15 minutes, rinsed in ddH_2O and soaked in depurinating solution for another 15 minutes. The gel was then soaked in denaturing solution for 30 minutes, followed by neutralising solution for 30 minutes. The gel was finally soaked in 20x SSC for a minimum of 10 minutes, or until ready to blot. The DNA in the gel was transferred using upward capillary action onto a positively charged nylon membrane (Hybond^{TM}-N, Amersham) for 2 to 3 nights. The membrane was baked at 80°C for 1 hour to immobilise the DNA on the membrane.

2.3.3.3 Southern Hybridisation

The membrane was soaked in 15-20ml of prehybridisation buffer containing 100µl of heat denatured DNA from fish sperm (Roche) at 65°C for at least 1 hour. The fish sperm DNA is a blocking agent which reduces the levels of background hybridisation on the membrane.

The p482.6 probe, inserted in the pUC19 plasmid vector containing an Ampicillin resistance gene, is grown in *E. coli* bacterial cells. The probe is a 1kb fragment of a 9.6kb insert from intron 22 of the *FVIII* gene, and excised from the vector using an *Eco* RI and *Sst* I restriction enzyme (Roche) double digest.

The p482.6 probe was labelled with ^{32}P-dCTP using the Megaprime DNA labelling kit (Amersham) (refer to Appendix B3 for the detailed labelling protocol). The labelled probe was heat denatured, added to the prehybridisation buffer, then to the membrane. The membrane was incubated in this solution at 65°C in the Maxi14 hybridisation oven (Hybaid) for approximately 72 hours.

After hybridisation, excess and non-specifically bound probe was washed off the membrane by increasing the stringency of the washes. The washes performed were as follows: 2 washes in a 2x SSC, 0.1% SDS solution for 10 minutes each at room temperature; 2 washes of a 1x SSC, 0.1% SDS solution for 20 minutes each at 65°C; 1 wash of a 0.5x SSC, 0.1% SDS solution for 10 minutes at 65°C; 2 washes of a 0.1x SSC, 0.1% SDS solution for 20 minutes each at 65°C; and finally 3 short rinses with a 0.1x SSC solution.

The membrane was exposed to Hyperfilm MP (Amersham) and stored in an X-ray cassette and kept at -80°C for 3-21 days before the X-ray film was developed in the XP400 automated X-ray film processor (AXIM).

2.3.3.4 Stripping of Labelled Southern Blots

If the labelling process was unsuccessful, the blots were stripped of the radioactive label which involved washing the membrane in denaturing solution for 10 minutes, followed by 2 washes of special neutralising solution for 10 minutes each and the membranes were stored in 2x SSC. Freshly labelled probe was hybridised to the stripped blot as described in section 2.3.3.3 above.

2.3.4 Detection of the Intron 1 Inversion Mutation

The detection of the intron 1 inversion was performed using a PCR protocol as described by Bagnall, *et al.* (2002), which involves the amplification of 1μl genomic DNA, 1x Amplitaq polymerase buffer (Perkin Elmer), 1.25mM $MgCl_2$ (Perkin Elmer), 0.125mM dNTP mix (Promega), 5% DMSO (Merck), 20pmol/μl of primers 9F, 9R and int1h-2F (IDT), 2.5U Amplitaq polymerase (Perkin Elmer) and made up to a final volume of 25μl with ddH_2O. Primer sequences can be found in Appendix C, Table C1.

The reaction was placed in an Applied Biosystems 2720 PCR machine with the following conditions: 94°C for 5 minutes, followed by 30 cycles of 94°C for 30 seconds, 63°C for 30 seconds and 72°C for 2 minutes. The reaction was completed with a final extension of 72°C for 5 minutes. Once the PCR was complete, 15μl of the PCR product was mixed with Ficoll dye and detected on a 1.5% agarose gel using the GeneSnap software (SynGene) and the G:box gel documentation system (Syngene).

The expected PCR product size is 1.5kb in a normal individual and 1.0kb in a gene with an intron 1 inversion. The detailed protocol for the intron 1 inversion PCR can be found in Appendix E1.

2.3.5 Reverse Transcription

RNA was reverse transcribed into cDNA in order to create a stable working stock of patient material, using the ImPromII™ Reverse Transcription System (Promega). The detailed protocol for reverse transcription can be found in Appendix B4.

Three different primer options can be used for cDNA synthesis, namely random hexamer, gene specific and oligo(dT)$_{15}$ primers. The random hexamer primer amplifies all RNA through random primer binding, the gene specific primer amplifies only the gene of interest, while the oligo(dT)$_{15}$ primer amplifies all mRNA. The oligo(dT)$_{15}$ primer was used to amplify the *FVIII* mRNA to allow the entire *FVIII* mRNA to be reverse transcribed in one piece. This method would exclude non-amplification due to potential polymorphisms or mutations in the gene specific primer.

2.3.6 Polymerase Chain Reaction for *FVIII* cDNA

2.3.6.1 Primer Selection and Design

The primer sequences used to amplify the *FVIII* cDNA fragments were obtained from Dr Peter Green, Kings College London (personal e-mail, 2004). They were checked against the September 2004 *FVIII* mRNA sequence download from www.genome.ucsc.edu (accession number: NM_000132) to confirm that the sequences and orientations of the primers were correct and the sequences were modified where needed.

As it was initially intended to perform protein truncation testing (PTT) on the overlapping *FVIII* fragments, special PTT primers were ordered and used during the optimisation phase of each fragment. However, it was later decided not to perform PTT due to the costs and direct sequencing was performed on the individual fragments instead (refer to section 4.3.1.1). For PTT, the forward and reverse primers needed certain additions that would allow for *in vitro* translation of the amplified products. The forward primers required a T7 promoter, a Kozac

(Eukaryotic translation initiation sequence) sequence and a start codon. A stop codon followed by a polyA tail needed to be incorporated into the reverse primers.

As primers were ordered for different applications, they were given different prefixes to avoid confusion. The cDNA primers had a prefix of HPA-RNA-, PTT primers with the prefix HPA-PTT- and internal sequencing primers with the prefix HPA-PTT-seq-. Sequences for all the primers used can be found in Table C2 in Appendix C. Some primers were used for other applications other than for what they were originally ordered.

2.3.6.2 PCR Optimisation and Final Protocols

Due to the large size of the *FVIII* transcript (approximately 9kb), the cDNA was initially divided into six overlapping segments as described in Naylor, Green, Rizza, *et al.* (1992), with each fragment spanning several exons. During optimisation of the PCR amplification conditions for these fragments, it was decided to amplify three fragments (namely fragment 1, 4 and 6) in smaller parts (fragment 1a & 1b, fragment 4a & 4b and fragment 6a & 6b) using the internal sequencing primers to reduce the size of the fragment thereby allowing for improved amplification of these regions. Thus there were nine fragments in total. These fragments and the exons covered by each fragment are shown schematically in Figure 2.1 and in more detail in Table C3 in Appendix C.

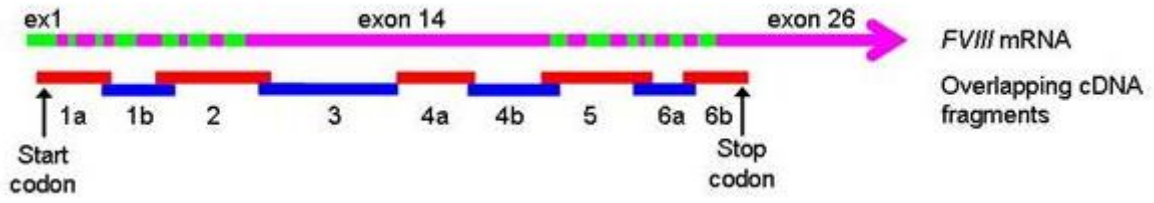

Figure 2.1: A schematic representation of the relative position and sizes of the overlapping cDNA fragments amplified in relation to the *FVIII* gene

After numerous attempts at optimisation, the final PCR conditions for each cDNA fragment were determined. Certain fragments, namely fragments 1b, 2, 5 and 6b, required a nested PCR for a better product yield. For this project, a standard PCR

mix for each fragment contained 1-2µl of cDNA or nested PCR products (where applicable) added to a reaction containing a final concentration of 1x buffer 2 of the Expand High Fidelity PCR System (Roche), 1.25mMol dNTPs (Promega), 20pmol of both the forward and reverse primer (IDT), 1U of Expand High Fidelity Taq polymerase (Roche) made up to a final volume of 25µl with ddH_2O

The reaction was placed in an Applied Biosystems 2720 PCR machine with the following conditions: 95°C for 10 minutes, followed by 35 cycles of 95°C for 1:30 minutes, annealing temperature for 1:30 minutes and 72°C for 2:30 minutes. The reaction was completed with a final extension of 72°C for 20 minutes. The annealing temperatures were specific for each fragment. Once the PCR was complete, the entire volume (25µl) of PCR product was mixed with Ficoll dye and checked for amplification on a 1.5% agarose gel using the GeneSnap software (SynGene) and the G:box gel documentation system (Syngene). The bands were excised from the gel and sequenced as described in section 2.3.8 below.

Some of the conditions for the fragments varied slightly from the standard PCR conditions above, to optimise the product yield. The primers used to amplify each fragment, annealing temperatures and expected product sizes can be found in Table 2.1. This table also includes any additional PCR components added to the reaction and the primers used for internal sequencing, where applicable. The primer sequences can be found in Table C2 of Appendix C below. The final versions and detailed protocol sheets for each fragment can be found in Appendices E2 to E10.

Table 2.1: The primers, annealing temperatures and expected PCR product sizes for each cDNA fragment, and where necessary, additional PCR components and internal sequencing primers used.

Fragment	Forward primer	Reverse primer	Annealing temperature	Expected product size	Additional reagents	Internal sequencing primers
1a	HPA-RNA-1a-F	HPA-PTT1-R	54°C	764bp	None	None
1b – first round	HPA-RNA-1a-F	HPA-RNA-1b-R	48°C	-	Betaine	-
1b – nested	HPA-PTT1-seq-F	HPA-PTT-1d-R	48°C	752bp	Betaine	None
2 – first round	HPA-RNA-2a-F	HPA-RNA-2b-R	60°C	-	None	-
2 – nested	HPA-PTT-2c-F	HPA-RNA-2d-R	64°C	1174bp	1 in 40 dilution of 0.1M spermidine	HPA-PTT2-seq-F HPA-PTT2-seq-R
3	HPA-RNA-3c-F	HPA-PTT-3f-R	60°C	1418bp	None	HPA-PTT3-seq-F HPA-PTT3-seq-R
4a	HPA-RNA-4c-F	HPA-PTT4-seq-R	50°C	798bp	Betaine	None
4b	HPA-PTT4-seq-F	HPA-RNA-4f-R-new	50°C	938bp	Betaine	None
5 – first round	HPA-RNA-5a-F	HPA-RNA-5b-R	60°C	-	None	-
5 – nested	HPA-RNA-5c-F	HPA-RNA-5d-R	56°C	1137bp	Betaine	HPA-PTT5-seq-F HPA-PTT5-seq-R
6a	HPA-PTT-6-F	HPA-PTT6-seq-R	54°C	640bp	None	None
6b – first round	HPA-RNA-6a-F	HPARNA-6b-R	48°C	-	Betaine	-
6b – nested	HPA-PTT6-seq-F	HPA-PTT-6d-R	52°C	663bp	Betaine	None

2.3.7 Polymerase Chain Reaction for *FVIII* DNA

The *FVIII* gene spans 186kb of genomic DNA and consists of 26 exons (Toole, *et al.*, 1984; Wood, *et al.*, 1984). Once a mutation was detected in a cDNA fragment, an attempt was made to confirm the mutation at the DNA level by sequencing the specific exon/exons and flanking sequences involved.

Primer sequences for all 26 exons of the *FVIII* gene were obtained from Dr Peter Green, Kings College London (personal e-mail, 2004). The primer sequences were mapped to the genomic DNA of the *FVIII* gene (September 2004 download from www.genome.ucsc.edu, accession number: NC_000023.9) to confirm that the sequences and orientations of the primers were correct. Exon 14 was amplified in four reactions (namely 14a1, 14a2, 14b1 and 14b2) due to the large size of the exon (approximately 3.1kb). The primer sequences for all the exons can be found in Table C4 in Appendix C. They span the exon, exon/intron boundaries and at least 50bp of intronic sequence on either side of the exon.

A semi-standard PCR was used to optimise the individual exons, differing in the primers, annealing temperature and number of cycles used. These conditions were as follows: 1µl genomic DNA added to a reaction containing 1.25mM dNTP mix (Promega), 1x Amplitaq Gold buffer (Perkin Elmer), 2.5mM $MgCl_2$ (Perkin Elmer), 10ρmol of each primer (IDT), 1U Amplitaq Gold polymerase (Perkin Elmer), made up to a final volume of 25µl with ddH_2O. The cycling conditions were 95°C for 5 minutes followed by a variable number of cycles of 95°C for 1 minute, respective annealing temperature for 1 minute, 72°C for 1 minute. A final extension at 72°C for 10 minutes completed the PCR conditions. The corresponding annealing temperatures, number of cycles and expected PCR product sizes for each exon are found in Appendix C, Table C4, while the protocol sheet can be found in Appendix E11.

For one patient, HPA020, primers were specially designed to amplify a region of intron 13 (namely 74701bp to 75082bp of the *FVIII* DNA sequence), encompassing the region where a suspected mutation was thought to occur. The PCR conditions were the same as the DNA PCR conditions and used primers HPA20-int13F and HPA20-int13R (refer to Appendix C, Table C4 for more details).

All 25µl of the PCR products were checked for amplification on a 3% agarose gel. The products excised from the gel and sequenced in an attempt to confirm the mutation at a DNA level (detailed in section 2.3.8)

2.3.8 Sequencing

DNA sequencing can be used as a tool to identify variants and mutations in genes of particular interest. One of two techniques can be used for DNA sequencing, namely the dideoxy or chain termination method of Sanger (Sanger, Nicklen & Coulson, 1977) and the chemical degradation method (Maxam & Gilbert, 1977). Both methods are based on high resolution electrophoresis to resolve their products (reviewed in Wilson & Walker, 1997; Sambrook & Russell, 2001).

The principle of the Sanger method requires an oligonucleotide to anneal to a single stranded DNA template and act as a primer for the reaction. All four dNTPs must be present in the reaction, together with a ddNTP. ddNTPs contain a 3'-H atom instead of the 3'-OH group of the dNTPs, which prevents the phosphodiester bond formation with the succeeding dNTP. There is a chance that a ddNTP will be incorporated into the growing DNA strand instead of a normal dNTP, and termination of chain elongation occurs at that site. The products are resolved on a polyacrylamide gel and as termination will occur at every position along the DNA strand, the order of the bands determines the sequence of the DNA (reviewed by Wilson & Walker, 1997; Sambrook & Russell, 2001).

Prior to automated sequencing technology, each ddNTP was added into a separate reaction and one dNTP was radioactively labelled with either ^{32}P or ^{35}S. The products were resolved in four adjacent lanes of a polyacrylamide gel, and detected by autoradiography. Using automated sequencing, four ddNTPs, each with a different coloured fluorescent label, are added in a single reaction and the products are resolved using an automated DNA sequencing machine (reviewed by Wilson & Walker, 1997; Sambrook & Russell, 2001).

The PCR products obtained from both the cDNA fragments and DNA exons were checked for amplification using agarose gel electrophoresis, and the products were excised from the gel using a clean scalpel and placed in a clean eppendorf tube. The agarose and excess primers and dNTPs were removed using NucleoSpin® Extract II columns (Macherey-Nagel). Cycle sequencing was performed on these cleaned PCR products using the ABI Prism® BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and either the forward or reverse primer. All unincorporated ddNTPs and primers were removed from the cycle sequencing products using either the DyeEx™ 2.0 Spin Kit (Qiagen) or the BigDye® X-Terminator™ Purification Kit (Applied Biosystems) clean up method. The detailed NucleoSpin® Extract II protocol can be found in Appendix B5 while the DyeEx® and X-Terminator™ protocols can be found in Appendix B6.

The cleaned cycle sequencing products were analysed using the ABI3130*xl* genetic analyser (Applied Biosystems). The parameters used for sequencing in the Foundation Data Collection software, v3.0 (Applied Biosystems), were as follows: Instrument Protocol 1 was RapidSeq36_POP7_Z or BDX_RapidSeq36_POP7_Z (depending on the cycle sequencing clean up method and for products larger than 500bp), and the Analysis protocol 1 was 3130POP7_BDTv3-KB-Denovo-v5.2 (POP7 is the polymer used for sequencing and microsatellite analysis). Samples were analysed using the sequence analysis software v5.2 (Applied Biosystems). These results were compared to a normal reference sequence to detect variations using the Seqman application of DNAstar (Lasergene).

Any variations identified in the cDNA fragments were named using standard nomenclature based on the recommendations of den Dunnen & Antonarakis (2000). To determine whether the variation was pathogenic or polymorphic, we accessed the HAMSTeRS mutation and polymorphism databases (http://europium.csc.mrc.ac.za/WebPages/Main/main.htm), a paper by Green, Bagnall, Waseem, *et al.* (2008) and a paper by Viel, Machiah, Warren, *et al.* (2007). Novel changes, possibly affecting splicing, were checked on one of two web-based splice site predictor programs, either NetGene2 server (www.cbs.dtu.dk/services/NetGene2) or Berkeley Drosophila Genome project Splice Site Predictor (www.fruitfly.org/seq_tools/splice.htlm).

2.3.9 Flow Chart of Mutation Analysis Process

Figure 2.2 shows a flow chart demonstrating the basic methods followed to identify the disease-causing mutation in the haemophilia A patients in this project. The different strategies used to determine the type of disease-causing mutation are illustrated in more detail for point mutations (missense and nonsense mutations in Figure 2.3), deletions (Figure 2.4) and insertions (Figure 2.5). The approach used to identify polymorphisms and benign variants can be found in Figure 2.3. As no rearrangement was detected in this study, the approach used to determine the type of rearrangement has not been elucidated.

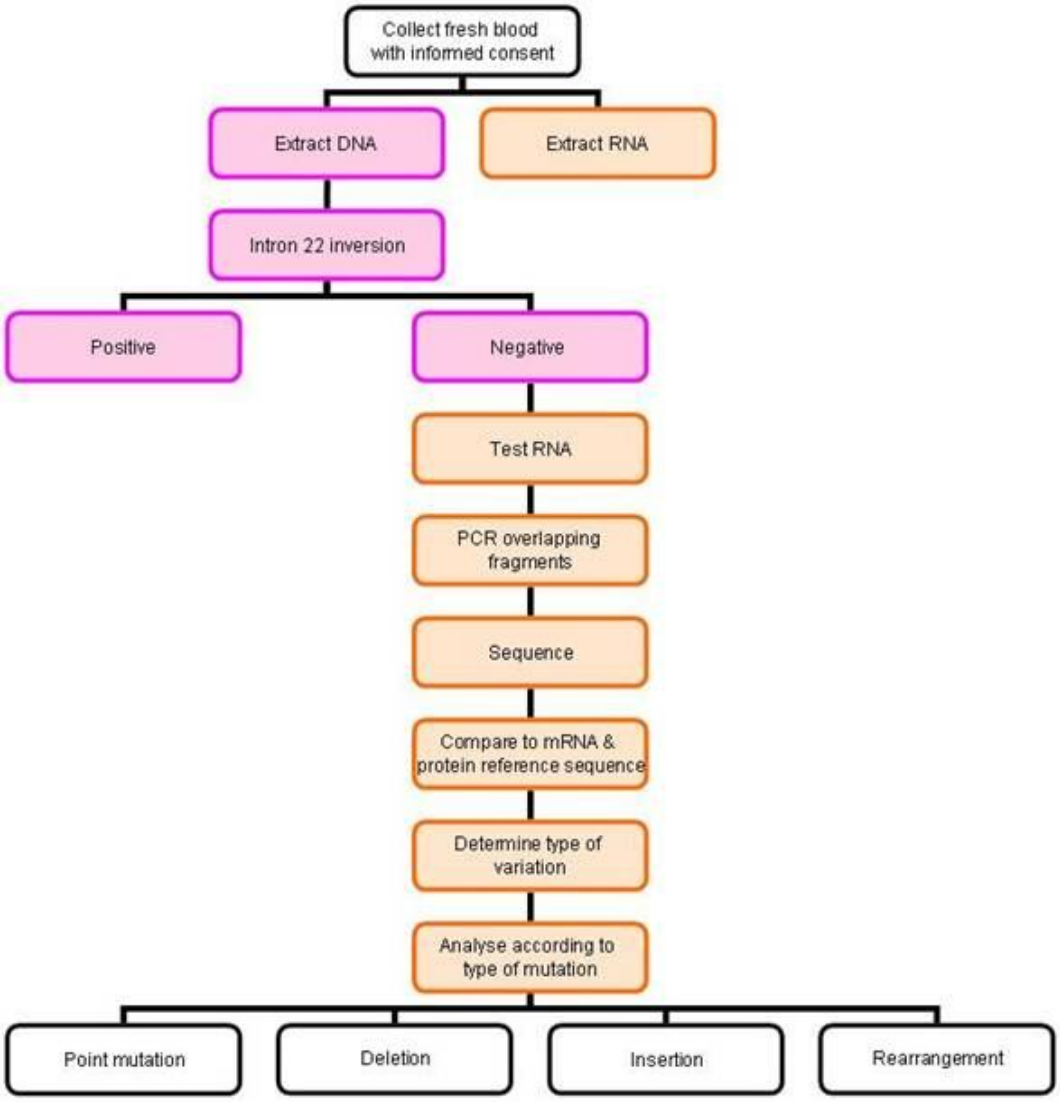

Figure 2.2: Flow chart of the methods followed to determine the disease-causing mutation in the haemophilia A patients in this project.

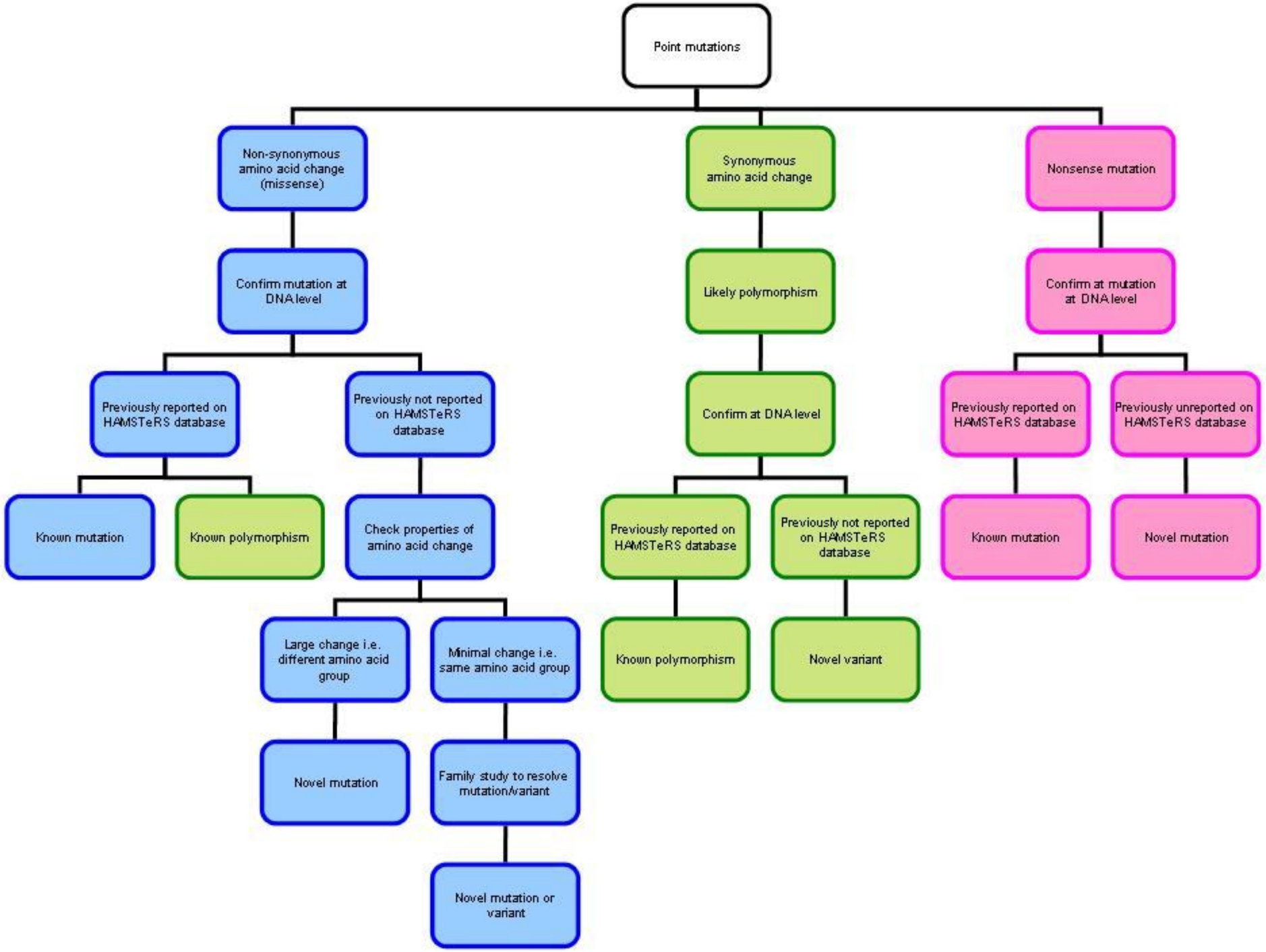

Figure 2.3: Flow chart of the methods followed to identify missense and nonsense mutations in the haemophilia A patients in this project.

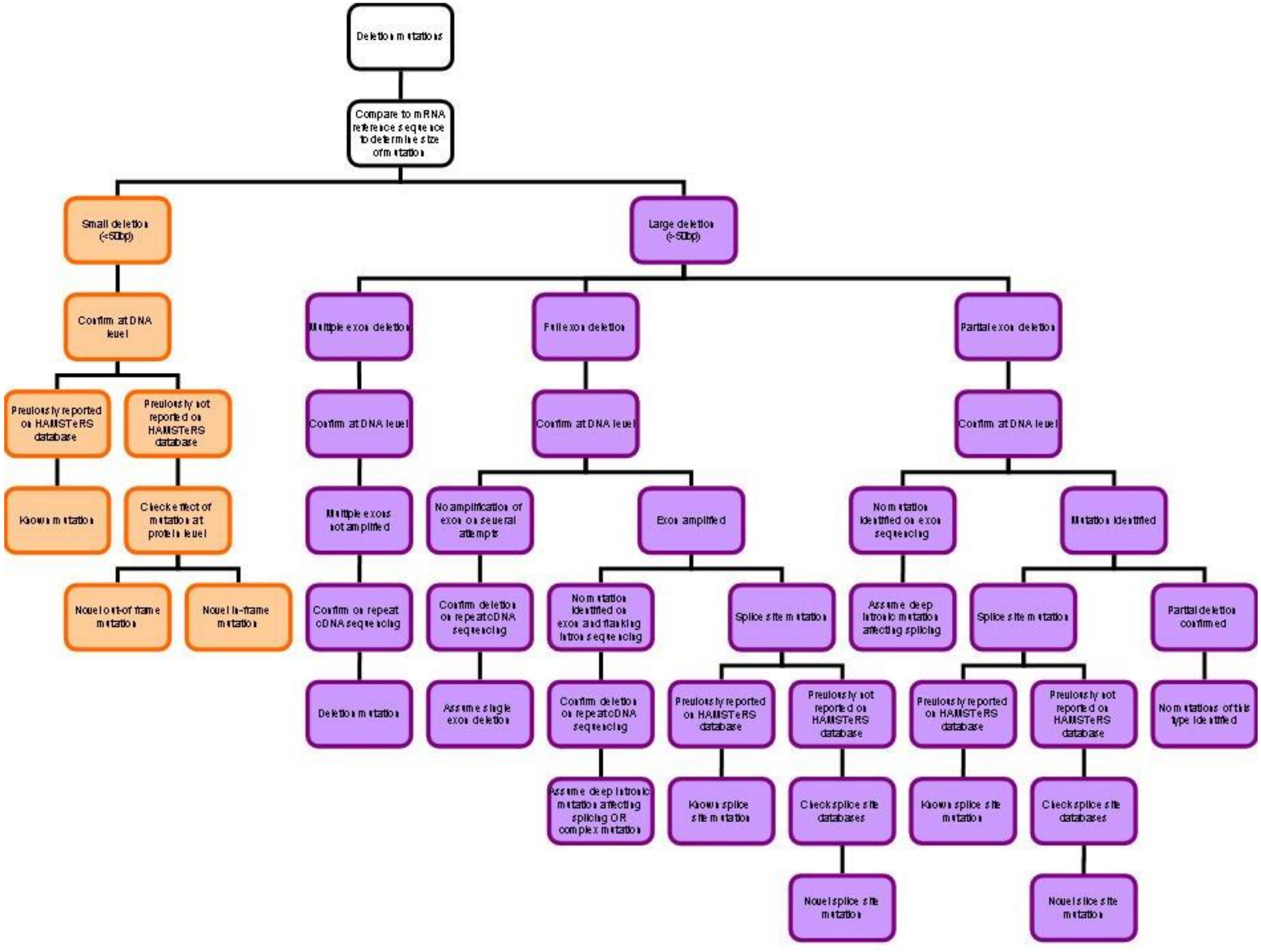

Figure 2.4: Flow chart of the methods followed to identify deletions at the cDNA level in haemophilia A patients in this project.

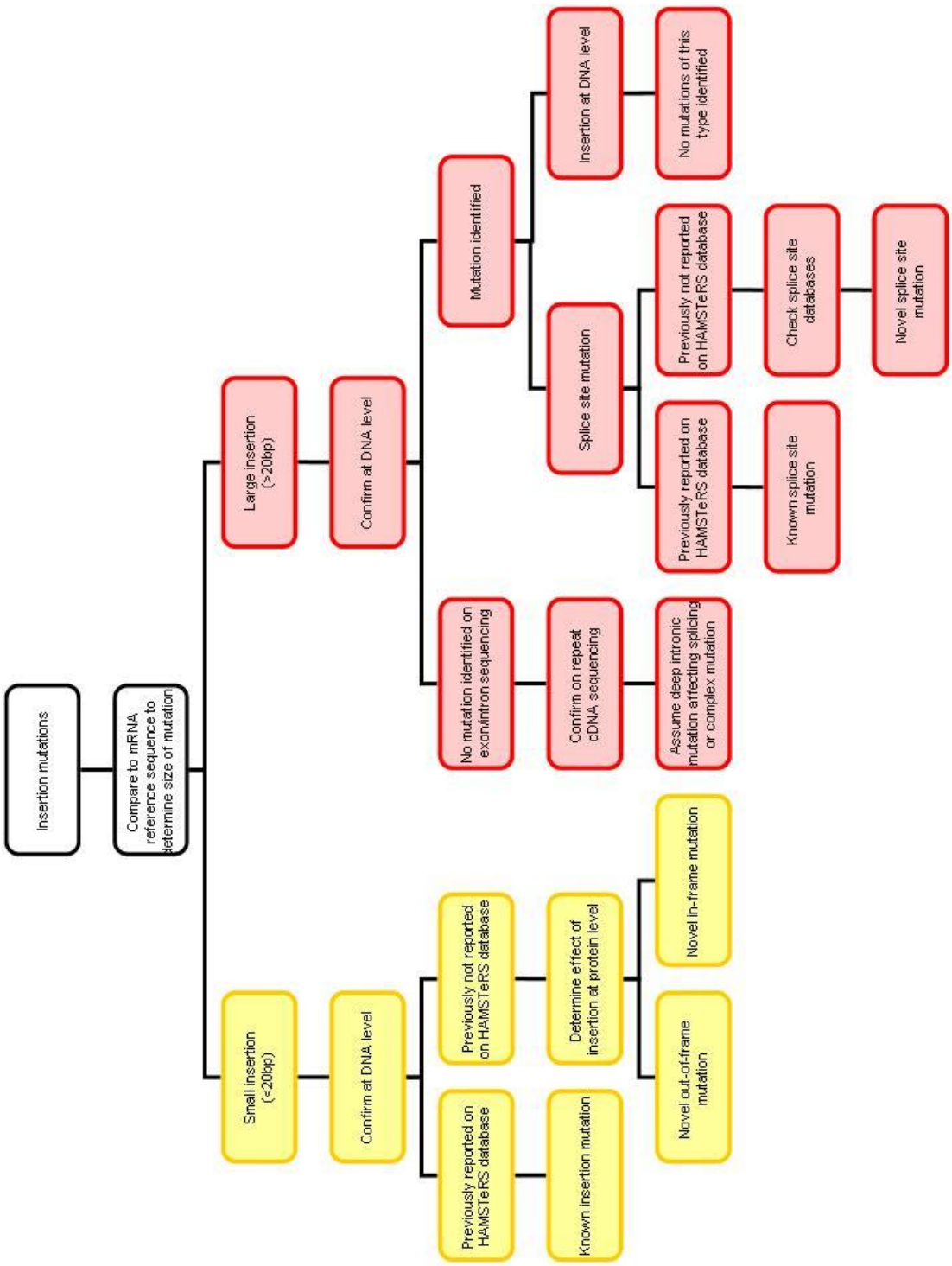

Figure 2.5: Flow chart of the methods followed to identify insertions at the cDNA level in haemophilia A patients in this project.

2.3.10 Haplotype Analysis

Haplotype analysis was undertaken on all patients who tested positive for the c.3637insA mutation using polymorphic linked markers in intron 13 and intron 22 of the *FVIII* gene and p39, an extragenic marker 500kb telomeric to the *FVIII* gene, as described below. The PCR products were sized more accurately using microsatellite analysis (refer to section 2.3.10.4).

2.3.10.1 Diagnostic Protocol for the Intron 13 Linked Marker

The diagnostic protocol for the intron 13 polymorphic STR marker can be found in Appendix E12. Primer 1A is fluorescently labelled with FAM (a blue fluorescent label) and both primer sequences can be found in Appendix C in Table C5. The expected PCR product sizes range from 137bp to 155bp. Intron 13 results were reported as AC repeat numbers. To convert the PCR product size into the number of AC repeats, the following formula was used: AC repeat number = (PCR product size – 108bp) ÷ 2.

2.3.10.2 Diagnostic Protocol for the Intron 22 Linked Marker

The diagnostic protocol for the intron 22 polymorphic STR marker can be found in Appendix E13. Primer 1B is fluorescently labelled with HEX (a green fluorescent label) and both primer sequences can be found in Appendix C in Table C5. The expected PCR product sizes range from 69bp to 87bp. The intron 22 results are also reported as AC repeat numbers. To convert the PCR product size into the number of AC repeats, the following formula was used: AC repeat number = (PCR product size – 32bp) ÷ 2.

2.3.10.3 Diagnostic Protocol for the p39 Linked Marker

The diagnostic protocol for the p39 polymorphic STR marker can be found in Appendix E14. Primer p39F is fluorescently labelled with HEX (a green fluorescent label) and both primer sequences can be found in Appendix C in Table C5. The expected PCR product sizes range from 152bp to 166bp. Unlike intron 13 and intron 22 results, the p39 marker results are reported as PCR product size and not repeat number.

2.3.10.4 Microsatellite Analysis using the ABI3130xl Genetic Analyser

Microsatellite analysis allows for accurate size determination of fluorescently labelled PCR products, which are run on the ABI3130*xl* Genetic Analyser (Applied Biosystems), using the GeneMapper v4.0 software. A detailed description of the protocol used for microsatellite analysis is given below.

To prepare the samples for analysis, 1μl of PCR product was mixed with 9μl of a ROX-HiDi Formamide solution, heat denatured at 94° for 2 minutes and placed on ice until ready to run on the genetic analyser. The sample sheet was created with the applications GeneMapper-Generic or GeneMapper-ABI_3130XL, and the following parameters were selected: sample type was set to standard, the size standard was GS500(-250), the individual microsatellites panel was selected, analysis method was set to Microsatellite Default, and the instrument protocol 1 was microsatellites.

The raw data results were analysed using the GeneMapper v4.0 software (Applied Biosystems). This software compares the peak sizes of the alleles of interest to the scored alleles of the ROX500 size standard to accurately size the peaks of interest. After analysis, each sample was analysed to check that the ROX500 size standard peaks were scored correctly and that the required alleles were called and sized appropriately.

3 RESULTS

3.1 INTRODUCTION TO THE CHAPTER

This chapter describes the results obtained for the analysis of *FVIII* mRNA, together with the intron 22 and intron 1 inversion mutations. Previous studies for other disorders in our laboratory using mRNA have been unsuccessful, but the results obtained from this project show mRNA is a valuable tool for mutation detection, especially in analysis of the *FVIII* gene. Using this fairly novel technique, a modified and improved mutation detection method has been developed for diagnostic use for haemophilia A in South Africa.

3.2 *FVIII* MUTATIONS IDENTIFIED

Of the 120 patients analysed, a causative mutation has been identified in 88 patients (73.3%). Although the disease-causing mutation was not identified in 32 patients (26.7%) this is most likely due to poor mRNA quality, time restrictions of the project or technical problems related to amplification of some mRNA fragments (as discussed further in section 4.3.2.2). Of these 32 patients, 16 of them tested negative for the intron 22 inversion mutation. An intron 22 inversion result was not obtained on the remaining 16 patients, either due to technical reasons involving Southern blot analysis or poor DNA quality. A summary of the patients that were tested for the different mutations in this study can be found in Figure 3.1.

Patients in whom no mutation was identified could have a large deletion (i.e. a full or partial *FVIII* gene deletion) or insertion that would hinder the PCR amplification of the respective fragments. As this possibility cannot be ruled out in the unsuccessful patients, large deletions and rearrangements could be underestimated in this study. The patients in whom no mutation was identified will be excluded from further analysis.

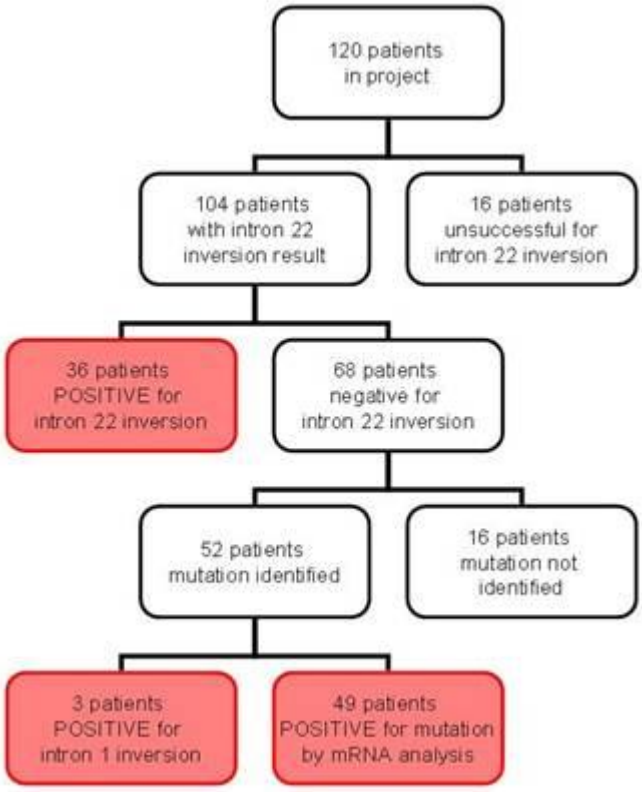

Figure 3.1: A flow chart showing the breakdown of patients tested at each stage of the project and a summary of the mutations identified.

3.2.1 Intron 22 Inversion Mutation Results

3.2.1.1 Intron 22 Inversion Frequency

The intron 22 inversion mutation was tested for in all patients, however, Southern blot results were only obtained for 84 patients. Southern blot results were unsuccessful for the remaining 36 patients, although a result can be inferred as negative in an additional 20 patients in whom a pathogenic mutation was identified (no reports of two pathogenic mutations in one patient have been found either in the HAMSTeRS database or in various papers). Of the 104 patients with an intron 22 inversion result (either direct or inferred), the population distribution is as follows: 51 white, 45 black and eight patients from the Indian and mixed ancestry populations.

This intron 22 inversion accounts for 27.5% (14/51) of mutations in white patients, 42.2% (19/45) of mutations in black patients and 37.5% (3/8) of mutations in the

Indian and mixed ancestry patients (graphically represented in Figure 3.2). The frequency in the white population corresponds to previous data obtained by Dangerfield, *et al.* (1997) and remains one of the lowest reported frequencies in the world (Antonarakis, *et al.*, 1995). Green, *et al.* (2008) found the intron 22 inversion to account for 16.6% of all mutations in haemophilia A patients in the United Kingdom, and 38% of patients affected with severe disease.

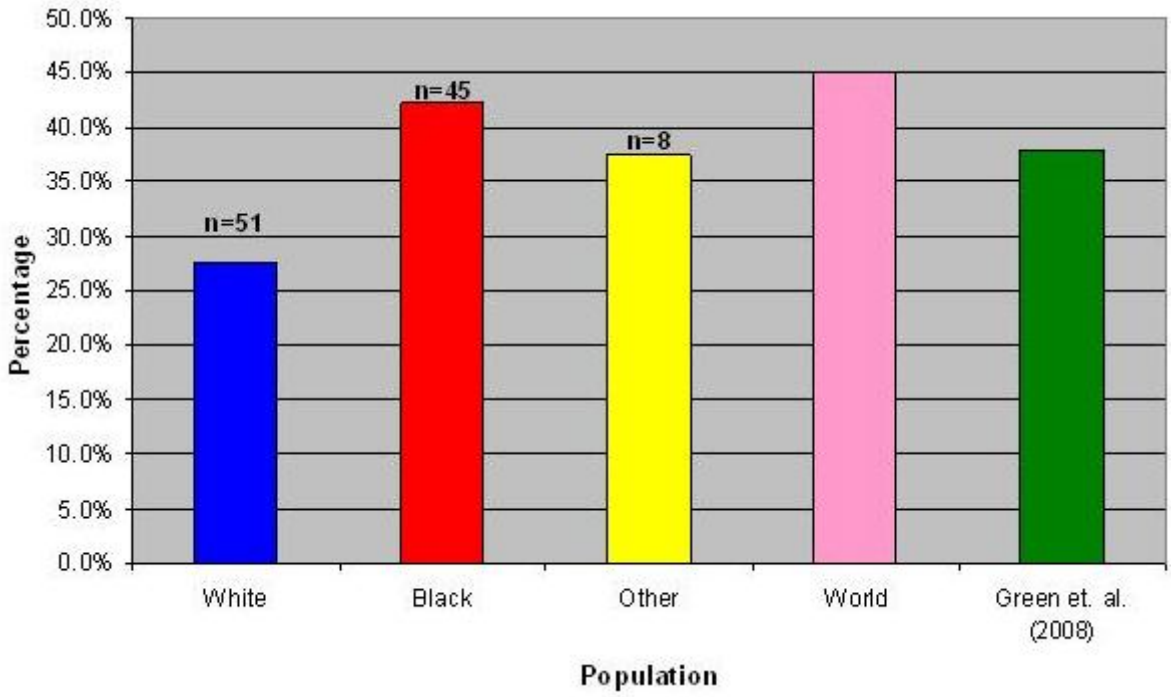

Figure 3.2: The percentages of the intron 22 inversion mutation across the different population groups in South Africa compared to the expected worldwide percentages reported by Antonarakis, *et al.* (1995) and Green, *et al.* (2008).
In this study, patients from the Indian and mixed ancestry populations have been combined as "Other" for ease of analysis, as individually, their sample sizes are too small to draw statistical conclusions.

The frequencies of both the white and black intron 22 positive haemophilia A patients were compared individually to the world frequency, using a binomial distribution. The binomial distribution analyses two independent sample sets (or populations) with two types of data in one population (Clarke, 1997). A probability of less than 0.05 would indicate a significant difference between the observed and expected frequencies. It was found that the probability of such a binomial

distribution in the white patients is p=0.008, which is significantly different from the expected frequency. As for the black patients, p=0.413, which does not deviate significantly from the expected frequency.

3.2.1.2 Type I and Type II Intron 22 Inversion Frequencies

The different types of intron 22 inversions were analysed in the whole South African population as the sample sizes for the individual population groups were too small to draw any conclusions. The type I inversion accounts for 19.2% (20/104 patients), the type II inversion mutation accounts for 13.5% (14/104 patients) and an atypical intron 22 inversion pattern for 1.9% (2/104 patients) of all mutations. The worldwide frequencies reported in Antonarakis, *et al.* (1995) are as follows: the type I inversion accounts for 35%, the type II for 7% and an atypical pattern accounts for just over 1% of all mutations.

Of the intron 22 inversion mutations, the type I inversion accounts for 55.6% (20/36), the type II inversion for 38.9% (14/36) and the atypical pattern for 5.6% (2/36). In the Antonarakis, *et al.* (1995) study, the type I inversion accounted for 82% (740/905) of intron 22 inversion mutations, while 15% (140/905) were the type II inversion and an unusual or aberrant pattern accounting for the remaining 3% (25/905) of intron 22 inversion mutations. A chi-squared statistical comparison of the frequencies of the different types of intron 22 inversions between the South African population and world frequencies as reported by Antonarakis, *et al.* (1995) was undertaken, and showed a significant difference ($p<0.0001$).

A comparison of the relative frequencies of the different types of the intron 22 inversion mutation in the different population groups is shown in Figure 3.3. The type II inversion appears to account for more intron 22 inversion mutations in all the South African population groups studied, compared to those reported by Antonarakis, *et al.* (1995). The breakdown of the frequencies per population results in small sample size for all population groups and could not be compared statistically. However, it is interesting to note that Dangerfield, *et al.* (1997) did not find the type II inversion in the black population, whereas this study found the type II inversion accounting for 42.1% (8/19) of intron 22 inversions in the black population.

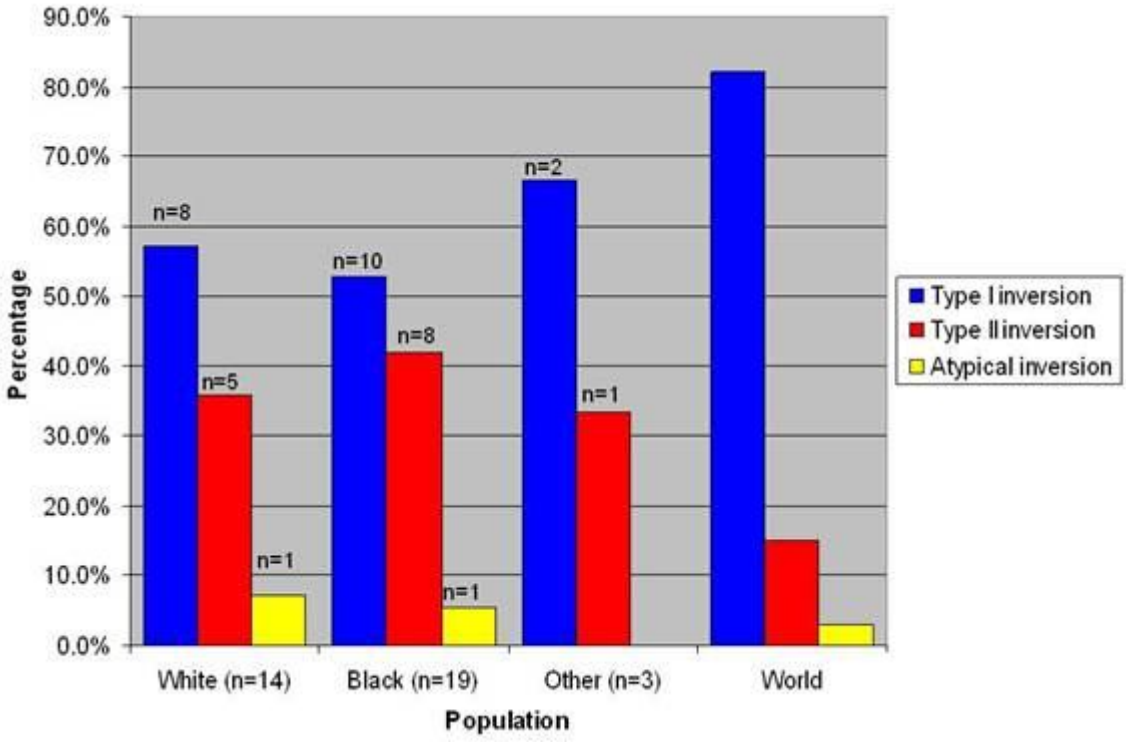

Figure 3.3: The population distribution of the type I, type II and atypical intron 22 inversions in the South African populations compared to the world frequencies reported by Antonarakis, *et al.* (1995).

Patients from the Indian and mixed ancestry populations have been combined as "Other" for ease of analysis, as individually, their sample size is too small to draw statistical conclusions.

3.2.2 Intron 1 Inversion Mutation Results

In a pilot study (not part of this study), 34 intron 22 inversion negative patients from the diagnostic haemophilia A cohort were tested for the intron 1 inversion mutation. All patients tested negative for this mutation.

In this study, screening was undertaken on all patients for the intron 1 inversion mutation using PCR analysis. This mutation was identified in three patients, two of whom were black patients and the third patient was of Indian ancestry. This equates to approximately 5.3% (2/38) of mutations in black patients and 25% (1/4) in Indian patients. It is interesting that this mutation was not found in any white patients in this project (n=49) or the preliminary screen (n=25), whereas it accounts for approximately 4.8% of severe haemophilia A mutations in the United

Kingdom (Bagnall, *et al.*, 2002). A summary of these results can be found in Appendix F in Table F1, highlighted in turquoise.

3.2.3 Mutations Identified by mRNA Analysis

By analysing the mRNA of haemophilia A patients, a wide variety of mutations have been detected in the *FVIII* gene and mutations were identified in 55.7% (49/88) additional patients. mRNA analysis significantly improves the mutation detection rate compared to the testing strategy currently offered by the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg for the intron 22 inversion mutation only.

Sections 3.2.3.1, 3.2.3.2, 3.2.3.3 and 3.2.3.4 present the mutations in haemophilia A patients by type and give an example in each case of the strategy used to identify mutations. A comprehensive breakdown of the mutations identified in all 88 of the patients, in whom a mutation was identified in this project, together with their severity and inhibitor status are summarised in Table F1 (Appendix F). This table includes any polymorphisms and benign variants identified in these patients.

3.2.3.1 Nonsense Mutations

Nonsense mutations were determined as described in the flow diagram in Figure 2.3 and a summary of these mutations are highlighted in pink in Table F1 (Appendix F).

Nonsense mutations in the *FVIII* gene accounted for 18.4% (9/49) of mutations identified using an mRNA approach and 10.2% (9/88) of the total mutations identified in this study. Four nonsense mutations (R427X, R795X, R1696X and W2271X) were previously reported in the HAMSTeRS point mutation database (http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/PointMutationsA.htm for mutations between amino acid 1 and 740, and for mutations in amino acids 741 to 2326, http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/PointMutationsB.htm). One patient

in this study tested positive for a double amino acid change, namely WD393-394XH, the pathogenic change (W393X) was previously reported by Green, *et al.* (2008). The remaining four mutations (Q744X, K1012X, L1229X and Q1629X) were novel.

3.2.3.1.1 An Example of a Previously Reported Nonsense Mutation

In patient HPA184, a G to A transition was identified in fragment 6b on cDNA analysis at position 6879bp of the *FVIII* cDNA. This mutation was confirmed by sequencing exon 25 of the *FVIII* DNA. This mutation caused a tryptophan to be replaced with a stop codon at amino acid position 2271. This is illustrated in Figure 3.4. This nonsense mutation is clearly pathogenic. It was also found in the HAMSTeRS database as having been previously reported by Laprise, Mak, Killoran, *et al.* (1998). The disease-causing mutation in this patient is thus W2271X.

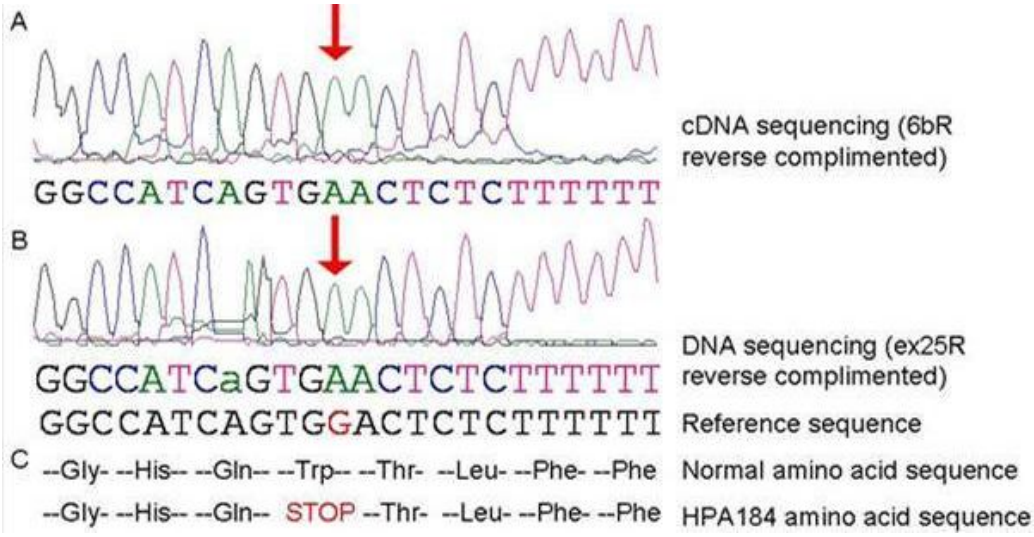

Figure 3.4: Raw data and interpretation of results obtained on patient HPA184 showing the previously reported nonsense W2271X mutation.

The electropherogram data are shown for part of the cDNA fragment 6b (A) and DNA exon 25 (B), with the G to A transition visible below the red arrows. The amino acid translation (C) is given below the electropherogram for both a normal sequence and the patient. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

3.2.3.2 Missense Mutations

Missense mutations were identified in this project using the approach outlined in Figure 2.3, and are indicated in Table F1 (Appendix F), highlighted in green. Missense mutations may be assessed using bioinformatics tools to examine cross species conservation and protein domain changes; however, this analysis has not been assessed in this project. Missense mutations in the *FVIII* gene accounted for 34.7% (17/49) of mutations identified using an mRNA approach and 19.3% (17/88) of the total mutations identified in this study.

A total of 17 missense mutations were identified by mRNA analysis, one of which was found in 2 patients (namely G450E). Of the 16 different missense mutations, six mutations (namely V162M, R372C, A704T, R2209G, W2229C and P2300L) were previously reported on the HAMSTeRS point mutation database (for mutations in amino acids 1 to 740, http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/PointMutationsA.htm, and http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/PointMutationsB.htm for amino acids 741 to 2326), while one mutation, E1829K was previously reported by Green, *et al.* (2008). The following nine mutations were novel: D125V, T395C, G450E, N467T, M614T, H979Q, D1568N, E1766K and W1942R. The G450E mutation was found in two patients, both of whom had the M2238V polymorphism, although they appear to be unrelated. One patient (HPA183) also had the A544A novel variant; however, the second patient (HPA084) was not tested for this variant.

3.2.3.2.1 An Example of a Previously Unreported Missense Mutation

In patient HPA186, an A to T transversion was identified in fragment 1a on cDNA analysis at position 431bp of the *FVIII* cDNA and confirmed by sequencing exon 4 of the *FVIII* DNA. This mutation caused an aspartic acid to be replaced with a valine at amino acid position 125. This is illustrated in Figure 3.5. This specific missense mutation has not been previously reported on the HAMSTeRS point mutation database. However, there were four reports of substitutions of aspartic acid at amino acid position 125, once to tyrosine, and three substitutions to asparagine (HAMSTeRS point mutation database for amino acids 1 to 740,

http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/PointMutationsA.htm). This mutation is likely to be pathogenic as aspartic acid is a polar, negatively charged amino acid and valine is a non-polar hydrophobic amino acid, while both tyrosine and asparagine are polar, hydrophilic amino acids. The disease-causing mutation in this patient is thus D125V.

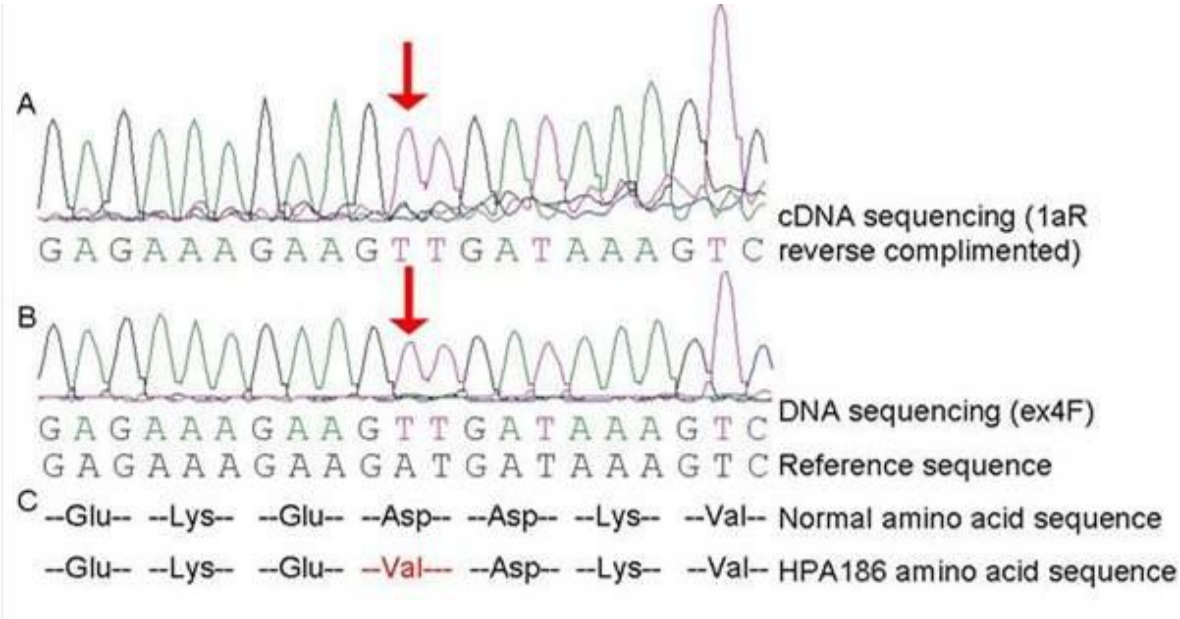

Figure 3.5: Raw data and interpretation of results obtained on patient HPA186 showing a novel D125V missense mutation.

The electropherogram data are shown for part of the cDNA fragment 1a (A) and DNA exon 4 (B), with the A to T transversion visible below the red arrows. The amino acid translation (C) is given below the electropherogram for both a normal sequence and the patient. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

3.2.3.3 Deletion and Insertion Mutations

Deletion mutations were determined using the approach outlined in the flow diagram in Figure 2.4 and insertion mutations were identified using the approach outlined in the flow diagram in Figure 2.5. All deletion and insertion mutations identified in this study are summarised in Table F1 in Appendix F and are highlighted in tan and yellow respectively.

Of the nine deletion mutations detected using *FVIII* mRNA analysis, the following five mutations were previously reported: ΔF652, c.3548delAA, c.3637delA,

c.4321delAGAA and c.4379delA (for large deletions of more than 50bp http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/LargeDeletions.htm and http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/SmallDeletions.htm for small deletions less than 50bp). This project identified three novel deletion mutations, namely: c.3745delCTGA, c.4118delA and c.4920delA (this mutation is described in more detail in section 3.2.3.3.1). The breakpoints of the exon 10-11 deletion in patient HPA202 were unable to be determined due to time constraints of the project and difficulty in amplifying exon 10 and 11 at the DNA level because of the presence of the deletion spanning these two exons. This mutation was confirmed on repeat sequencing of cDNA fragment 2. It cannot therefore be determined if this mutation was previously reported or if it is novel.

Two different insertion mutations were identified in this project in seven patients. One mutation (c.3637insA) was previously reported and one (c.4201insC) was novel (HAMSTeRS insertion mutation database at http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/Insertions.htm). The c.3637insA mutation was identified in 6 white patients which lead to the hypothesis of this being the founder mutation in this population group, and is described in more detail in section 3.2.6.

All the deletion and insertion mutations in this project resulted in a frameshift mutation, except for two in-frame deletion mutations: namely the single amino acid deletion (ΔF652 in patient HPA258) and the deletion of exon 10 and 11 (patient HPA202).

Green, *et al.* (2008) found that 47.7% of the small insertions and deletions identified in their study occurred in a string of six or more adenines. Of the 10 different small insertion or deletion mutations identified in this study, three (30%) occurred in a stretch of six or more adenines. This is most likely due to replication slippage (Becker, *et al.*, 1996).

3.2.3.3.1 An Example of a Novel Deletion Mutation

In patient HPA128, an A deletion was identified in exon 14 on DNA analysis at position 4921bp of the *FVIII* cDNA and confirmed by repeat sequencing of exon 14. This mutation caused a frameshift mutation from the glutamic acid at amino

acid 1621, and a stop codon at 1622. This is illustrated in Figure 3.6. This specific deletion mutation has not been previously reported on the HAMSTeRS deletion mutation database. However, this mutation is likely to be pathogenic as the resulting protein is prematurely truncated at amino acid 1622 (just over half of the 2332 amino acids are translated). The disease-causing mutation in this patient is thus c.4920delA.

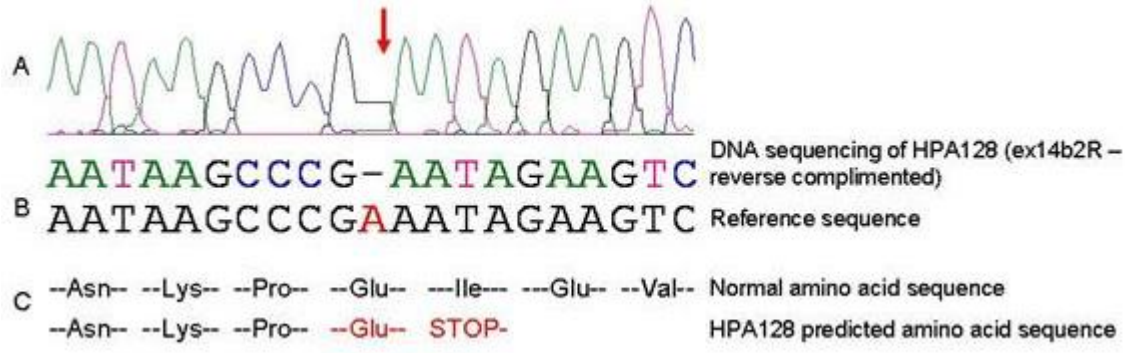

Figure 3.6: Raw data and interpretation of results obtained on patient HPA128 showing the novel deletion mutation, c.4920delA.

The electropherogram data are shown for part of the DNA sequencing of exon 14b2 (A) compared to the reference sequence (B). The amino acid translation (C) is given for both the normal sequence and the patient. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene). The apparent gap in the electropherogram (a) is a program artefact to highlight missing bases.

3.2.3.4 Mutations Affecting Splicing

Mutations affecting splicing were broadly divided into two sections: splice site mutations and complex mutations (refer to the approach outlined in Figure 2.4 and Figure 2.5) and are summarised in Table F1 in Appendix F (highlighted in purple and orange respectively).

Of the two splice site mutations, both were novel mutations (IVS13-1G→C and c.2111C→T which caused a deletion of exon 13, as discussed in section 3.2.3.4.1). Five complex mutations were identified in this project, and were defined as mutations where the exact DNA change was not identified at a DNA level, and are likely to be cryptic splice site or deep intron mutations. Due to the cost and time constraints of this project, sequencing of the entire intronic sequence for

these patients was not performed. We were therefore unable to determine whether these complex mutations were previously reported or novel. These mutations included a deletion of exon 10 (discussed in section 3.2.3.4.2), an insertion of 113bp of part of intron 13 into the *FVIII* mRNA at position 2120bp, a partial deletion of exon 16 (deletion of amino acids 1779-1788, and is discussed in section 3.2.3.4.3), and two deletions of exon 19.

3.2.3.4.1 An example of an Unreported Splice Site Mutation

In patient HPA236, a 209bp deletion was found on analysis of the *FVIII* cDNA fragment 2 (illustrated in Figure 3.7A). This deletion corresponds to a deletion of exon 13 and results in an in-frame deletion of 71 amino acids. DNA sequencing of exon 13 identified a C to T transition at cDNA position 2111bp (3bp from the end of exon 13, just before the exon/intron splice site) (as shown in Figure 3.7B).

This base change appears to substitute a proline at amino acid 687 with a leucine, a previously unreported mutation on the HAMSTeRS mutation database (Figure 3.7C). However, this C to T transition is more likely to affect splicing, based on the mRNA evidence. This mutation was checked on two internet based splice site predictor websites (http://www.fruitfly.org/cgi-bin/seq_tools/splice.pl and http://www.cbs.dtu.dk/services/NetGene2), and both websites predict that this C to T change does not affect splicing. Perhaps this mutation alters the competency of splicing of exon 13 or has a subtle effect on an exon splice site enhancer sequence, rather than affecting the actual splice site, although no normal transcript was seen in this patient.

Repeat sequencing of cDNA fragment 2 confirmed the deletion of exon 13. Therefore, according to the cDNA sequence data, this mutation does alter splicing and is likely to be pathogenic. The disease-causing mutation at a DNA level in this patient is thus 2111C→T which causes the deletion of exon 13 on cDNA, presumably by aberrant splicing.

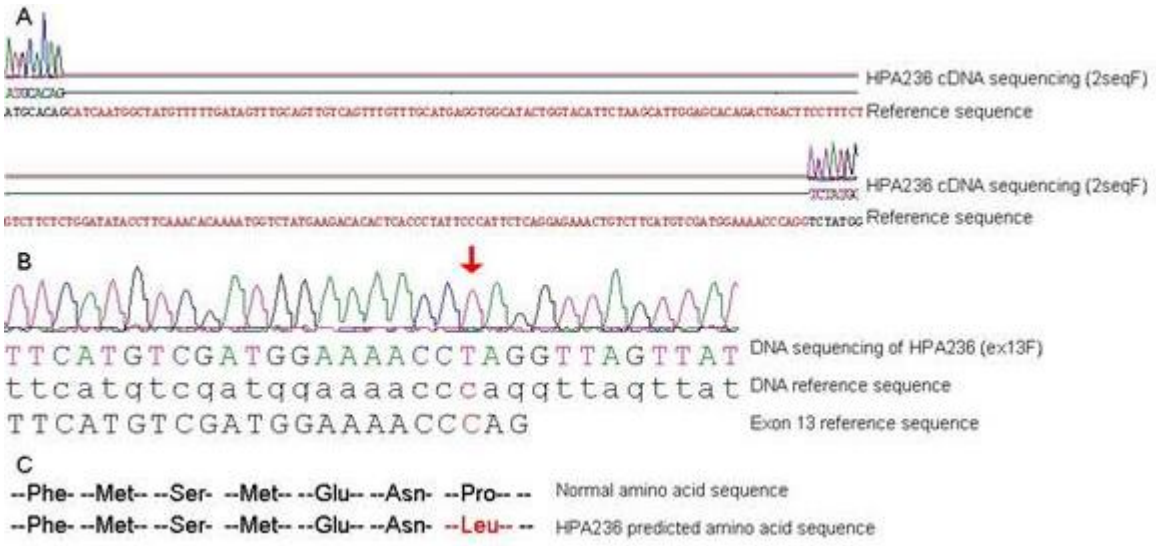

Figure 3.7: Raw data and interpretation of results for patient HPA236 showing an unreported exon 13 deletion resulting from the c.2111C→T splicing mutation.

The electropherogram data are shown for part of cDNA fragment 2 (A) showing a deletion corresponding to a deletion of exon 13. The electropherogram data obtained from DNA sequencing for part of exon 13 (B), with the T to C transition visible below the red arrow. The amino acid translation for a normal sequence and the patient are shown in (C). Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene). The apparent gap in the electropherogram (a) is a program artefact to highlight missing bases.

3.2.3.4.2 An Example of a Complex Mutation

In patient HPA219, a 199bp deletion was identified in fragment 2 on cDNA analysis (from position 1445bp to 1538bp of the *FVIII* cDNA), corresponding to a deletion of exon 10. At the protein level, this deletion caused a frameshift from the isoleucine at amino acid 464, and produced a stop codon at the next amino acid (position 465), and is therefore likely to be the pathogenic mutation in this patient. However, no mutations were identified by DNA sequencing of exon 10 and the intron/exon boundaries (230bp into intron 9 and 126bp into intron 10), thereby eliminating obvious donor and acceptor splice site mutations. This is illustrated in Figure 3.8. Sequencing of the entire intron 9 and intron 10 were not practical within the time and cost confines of this study as these two introns are over 4800bp and 3900bp in size respectively. Thus the specific mutation at the DNA level has not been identified and is assumed to be a deep intronic mutation or a cryptic splice site. The effect of the disease-causing mutation in this patient is a deletion of exon 10, detectable on cDNA analysis, causing a truncated protein.

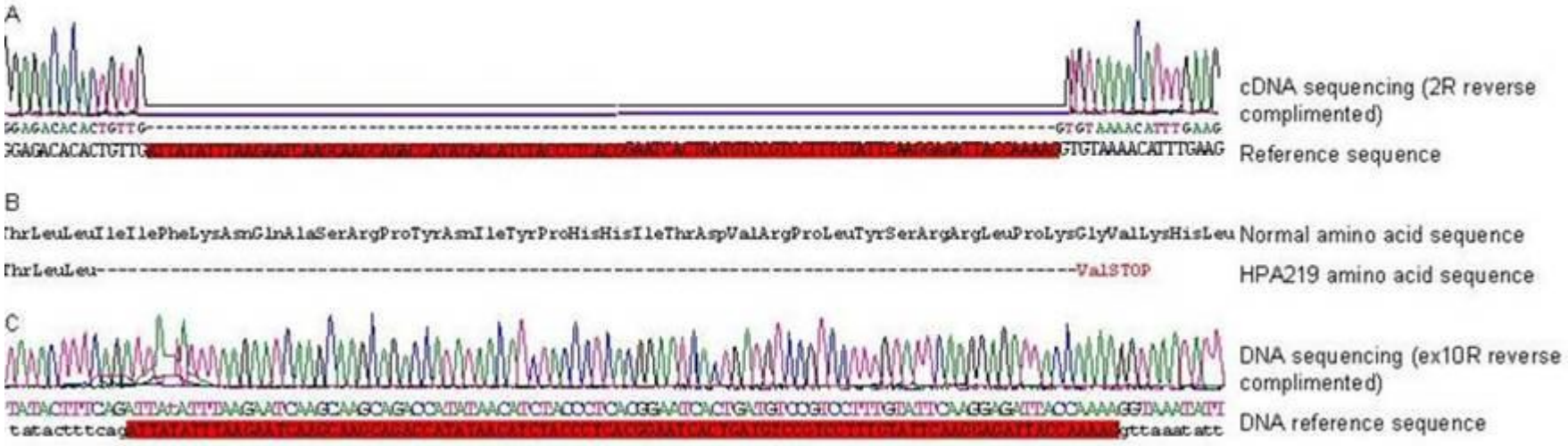

Figure 3.8: Raw data and interpretation of results obtained on patient HPA219, showing an exon 10 deletion on cDNA.

The electropherogram data for part of cDNA fragment 2 (A) shows a 199bp deletion that corresponds to exon 10. The amino acid translation (B) for both a normal sequence and the patient. DNA analysis (C) of exon 10 (highlighted on the DNA reference sequence in red) did not identify any mutations that could affect splicing. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene). The apparent gap in the electropherogram (a) is a program artefact to highlight missing bases.

3.2.3.4.3 An Example of Alternative Splicing Resulting in Two Transcripts

In patient HPA223, a 47bp deletion was identified in cDNA fragment 5 (from 5373bp to 5419bp of the *FVIII* cDNA), which corresponded to a partial deletion of exon 16 (illustrated in Figure 3.9A). At the protein level (Figure 3.9B), this deletion caused a frameshift from the valine at amino acid 1773, and produced a stop codon at amino acid 1793, and is therefore likely to be the pathogenic mutation in this patient. No mutation was identified on DNA sequencing of exons 15, 16 and 17 and their intron/exon boundaries (69bp 5' of exon 15, 77bp 3' of exon 15, 67bp 5' of exon 16, 101bp 3' of exon 16, 60bp 5' of exon 17 and 63bp 3' of exon 17), except for a T deletion 37bp into intron 16 (IVS16+37delT) (shown in Figure 3.9C). This IVS16+37delT could possibly be an unreported variant or a mutation affecting splicing. The http://www.cbs.dtu.dk/services/NetGene2 splice predictor website predicted a splice site at cDNA position 5420bp for both the normal sequence and patient HPA223 (including the IVS16+37delT mutation).

On closer inspection of the electropherogram data of fragment 5 (Figure 3.9A), there appears to be a second sequence 5' to the deletion (this reverse sequence is reverse complemented and reads in the 3' to 5' direction), which corresponds to the normal sequence (highlighted by the blue arrow). This patient therefore seems to have two transcripts, although the normal transcript occurs at a much lower level than the alternate transcript.

Thus the specific mutation at the DNA level has not been identified and is assumed to be a deep intronic mutation creating an alternate splice site. The disease-causing mutation in this patient therefore causes a partial deletion of exon 16, detectable on cDNA analysis. The pathogenicity of the IVS16+37delT is unclear, but could be used as a linked marker.

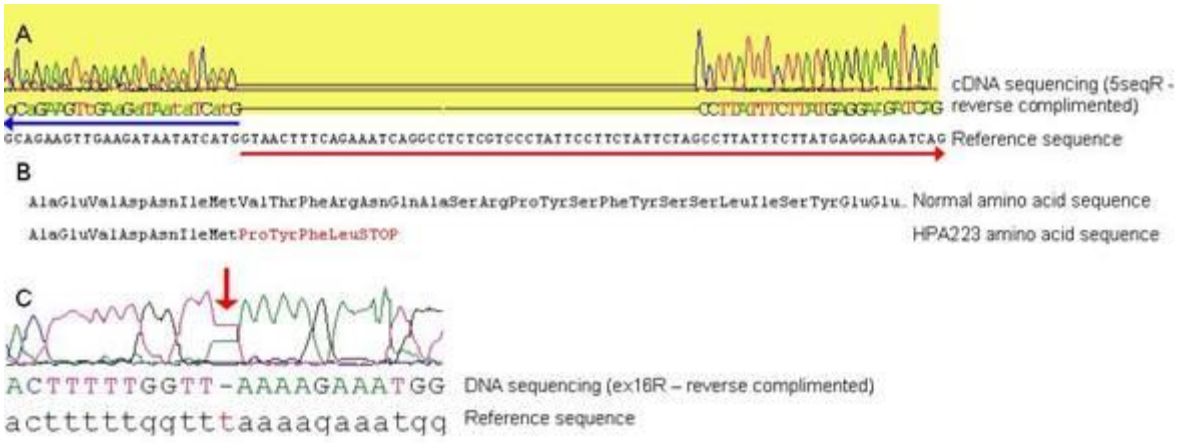

Figure 3.9: The raw data obtained for patient HPA223 showing alternate splicing of part of exon 16 on cDNA analysis

The electropherogram data for part of cDNA fragment 5 (A) showing a deletion that corresponds to part of exon 16 (the red horizontal arrow). The normal amino acid sequence (B) compared to the predicted amino acid sequence in this patient. The electropherogram data for part of intron 16 (117940bp-117962bp) obtained from DNA sequencing of exon 16 (C) showing a T deletion (below the red vertical arrow), the pathogenicity of which is unknown. The alternate transcript is seen in the reverse complemented cDNA sequence (A) as messy or "frameshift" sequence and is highlighted by the blue arrow. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene). The apparent gap in the electropherograms in (a) and (c) are a program artefact to highlight missing bases.

3.2.4 Polymorphisms

Any change that was suspected to be a polymorphism or benign variant was either checked against the HAMSTeRS polymorphism database (available at http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/Polymorphisms.htm) or a paper describing *FVIII* polymorphisms by Viel, *et al.* (2007). In this study, five polymorphisms or benign variants were identified, four of which were previously reported (D1241E, S1269S, P1481P and M2238V). In the South African haemophilia A patient group as a whole, these polymorphisms were found at a frequency of 0.44 (23/52), 0.11 (5/47), 0.02 (1/43) and 0.18 (5/28) respectively. The novel variant, namely A544A, was only found in one patient from the black population (n=28). Table 3.1 shows the population distribution of the different polymorphisms and variants and their frequencies. The frequencies were determined for the specific polymorphism and variant for those patients for whom data were available. Data were not available for each patient for each

polymorphism or variant as complete cDNA sequencing was not undertaken in patients once a mutation was detected due to costs. Also note that the sample sizes of the Indian and mixed population group are too small to draw any significant conclusions.

Table 3.1: Population distribution and frequencies of the different polymorphisms and benign variants found.

Population	A544A	D1241E	S1269S	P1481P	M2238V
White	Not found (0/39)	0.22 (5/23)	0.13 (3/24)	Not found (0/18)	0.10 (1/14)
Black	0.04 (1/28)	0.68 (17/25)	0.10 (2/20)	0.05 (1/22)	0.25 (3/12)
Indian/Mixed ancestry	Not found (0/6)	0.25 (1/4)	Not found (0/3)	Not found (0/3)	0.50 (1/2)

Except for the novel variant, A544A, all polymorphisms are previously reported by Viel, *et al.* (2007).

3.2.5 Implication of Mutations in Family Members

To date, in the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg, linked marker analysis has involved the use of 2 polymorphic STR markers in intron 13 and intron 22 of the *FVIII* gene and one extragenic marker, the p39 marker (as discussed in section 1.11.2). Even if informative, linked marker analysis remains an indirect method of mutation detection. It would only determine if a female is carrying the high risk X chromosome. It does not take into account the possibility of new mutations or germline mosaicism. Knowledge of the disease causing mutation in the proband would allow a female's carrier risk to be refined and determined more accurately through direct mutation detection. Three examples are illustrated in section 3.2.5.1, 3.2.5.2 and 3.2.5.3 below, where mutation analysis has either confirmed or altered the carrier risk determined by linked marker analysis.

3.2.5.1 Confirmed Carrier Status as Determined by Linked Marker Analysis

A novel mutation, c.4201insC, was identified in the patient HPA029. Direct mutation analysis was undertaken on the patient's mother and sister who were at high risk of being carriers based on linked marker analysis (refer to Figure 3.10A). Linked marker analysis could not determine if this was a new mutation in the patient nor exclude the mother from being a germline mosaic for the mutation. Direct mutation detection in the mother and sister confirmed them both to be carriers of the c.4201insC mutation, as shown in Figure 3.10C and D.

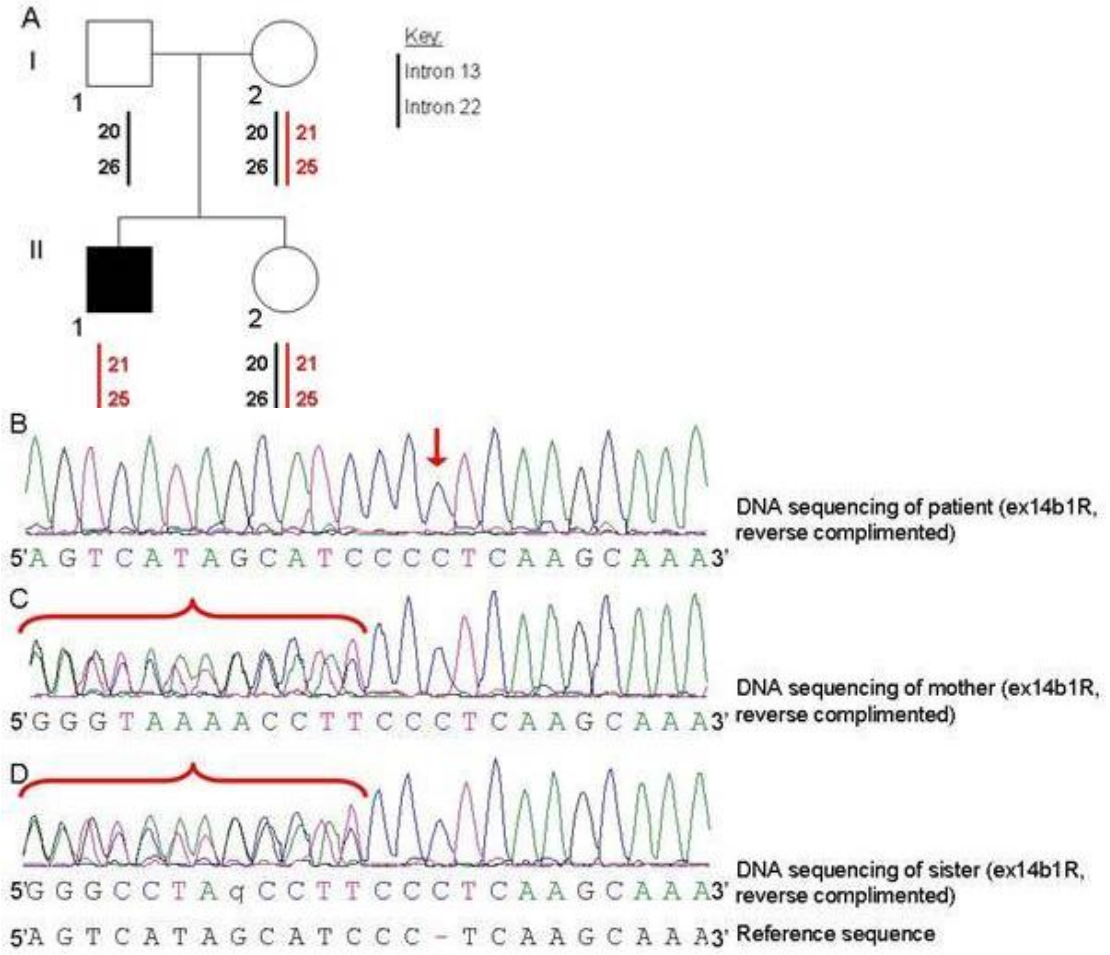

Figure 3.10: An illustration of a family where direct mutation detection confirmed the carrier status obtained from linked marker results.

The pedigree and haplotype results for family HPA029 (A), shows that the high risk chromosome has been inherited by II2. Partial electropherogram data obtained from DNA sequencing of exon 14b1 in family HPA029 in the patient (B) showing the c.4201insC mutation, his mother (C) and sister (D). Although the sequences are reverse complemented, the frameshift in the mother and sister is "before" the point of the mutation. The carrier status of the mother and sister can be confirmed. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

3.2.5.2 Mother with Germline Mosaicism or New Mutation in Patient

In patient HPA184, a previously reported c.6879G→A mutation was identified in the cDNA causing a W2271X mutation (refer to section 3.2.3.1) (Laprise, *et al.*, 1998). Linked marker analysis found that the patient's unaffected brother shared the high-risk X chromosome (refer to Figure 3.11). This would mean that the mothers' carrier risk would be significantly decreased, although she could still be at risk of being a germline mosaic. Direct detection for the c.6879G→A mutation was carried out on the mother and brother in an attempt to refine their status. Both the mother and brother tested negative for the mutation, as can be seen in Figure 3.12. Although the results cannot distinguish if this is a new mutation in the patient or whether the mother is a germline mosaic, future definitive prenatal diagnosis by mutation analysis can be offered to the mother.

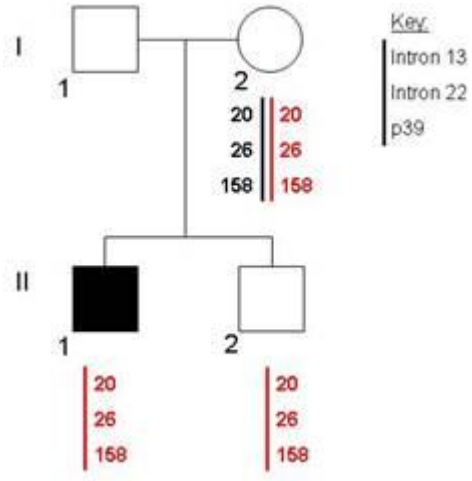

Figure 3.11: The pedigree and haplotype results for family HPA184 before direct mutation detection.

The haplotype results identified the high-risk X chromosome in the affected proband (II1). It has also been inherited by the unaffected brother (II2). This result does not resolve the mother's carrier status although it decreases her carrier risk.

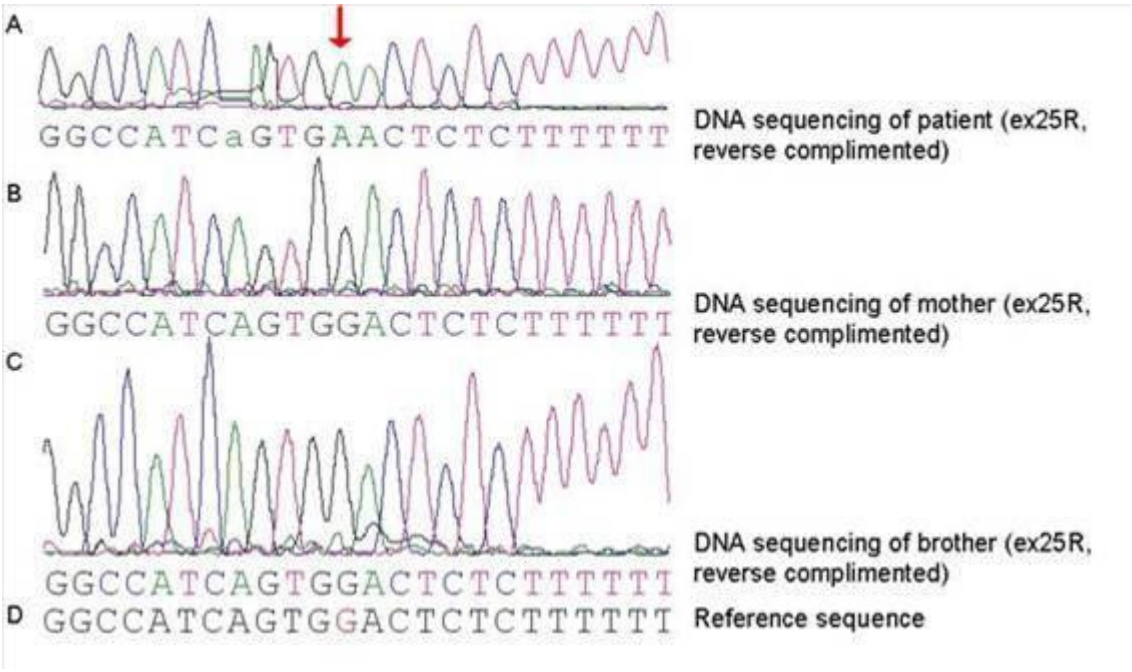

Figure 3.12: Direct mutation detection results for family HPA184 for the c.6879G→A mutation.

Partial electropherogram data for exon 25 from the patient (A), his mother (B) and brother (C) are shown. (D) is the normal reference sequence. The mother and brother tested negative for the c.6879G→A mutation, suggesting a new mutation in the patient or germline mosaicism in the mother. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

3.2.5.3 Inconclusive Linked Markers Resolved on Mutation Analysis.

In family HPA112, linked marker results identified the high risk allele in the patient. His unaffected uncle (II1) shared this high risk allele, as is seen in Figure 3.13. The patient's grandmother (I2) could possibly be a germline mosaic or this could be a new mutation in the patient's mother (II1) or even in the patient. There is also a possibility that his mother could be a germline mosaic. Therefore the carrier risk for the mother and aunt (II2) was difficult to define using these results.

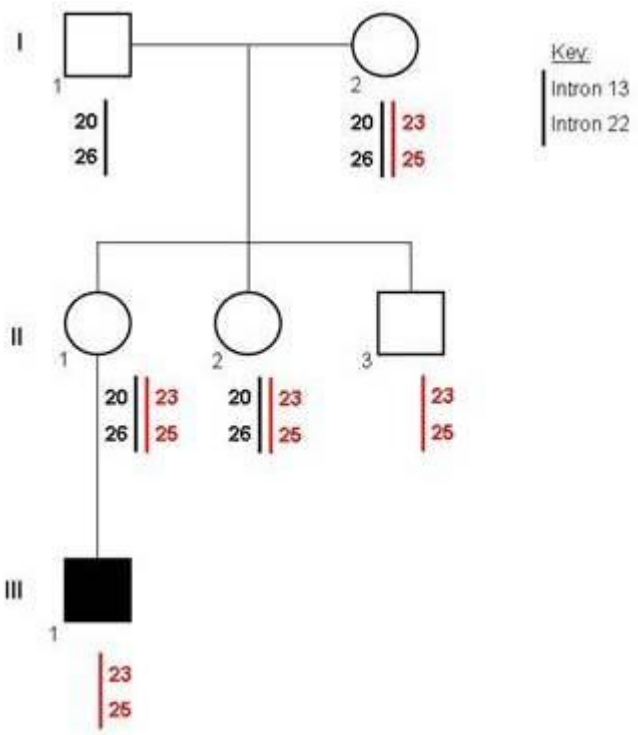

Figure 3.13: A pedigree diagram showing the linked marker results for family HPA112.

The high-risk X chromosome present in the patient is marked in red.

The proband in this family was found to have a novel c.1898T→C mutation in exon 12, resulting in a M614T missense mutation (refer to Figure 3.14). Direct mutation detection for this mutation was able to resolve the haplotype data and refine the carrier risk for the mother, aunt and grandmother. DNA sequencing data for the family can be found in Figure 3.15. It shows that the aunt and grandmother are negative for the disease-causing mutation, thereby reducing their high risk of being a carrier to 0%. It also proves that the unaffected uncle is negative for the M614T mutation. The risk for the mother of the patient being a carrier has been significantly reduced; however, it cannot be determined if she is a germline mosaic or if this is a new mutation in the patient. Either way, definitive prenatal testing is available to her using direct mutation analysis for future pregnancies.

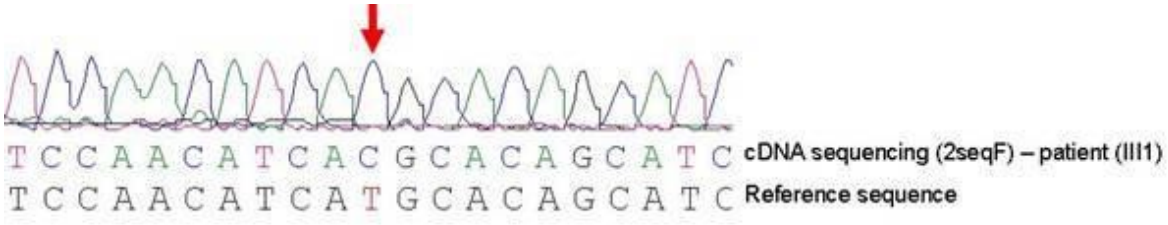

Figure 3.14: Partial electropherogram data for cDNA fragment 2 in the proband of family HPA112 corresponding to the M614T missense mutation.

It shows the c.1898T→C mutation below the red arrow. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

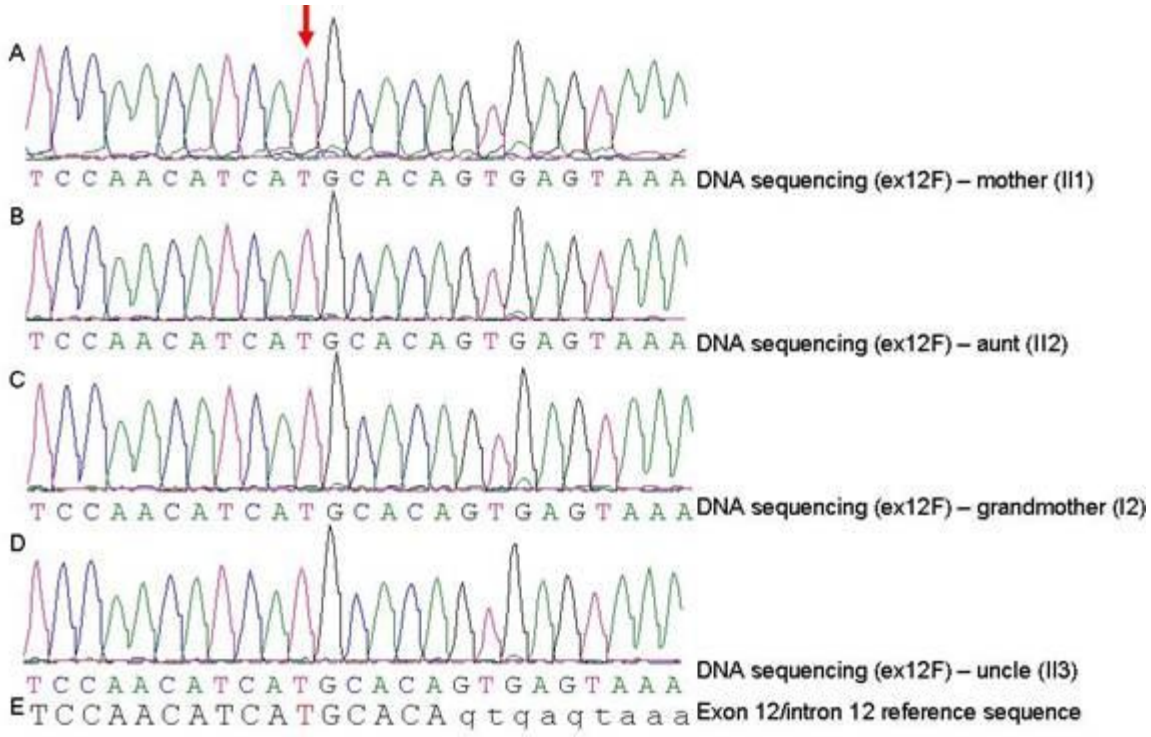

Figure 3.15: The DNA sequencing data for family HPA112 showing the c.1898T→C (M614T) mutation.

The DNA electropherogram for exon 12 showing the absence of the c.1898T→C (M614T) mutation in the patients mother (A), aunt (B), grandmother (C) and unaffected uncle (D). The normal reference sequence is below (E). Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene).

3.2.6 Common Mutation in the White Population

The c.3637insA mutation (described in section 3.2.6.1) was found in 14% of white patients (6/43), four of whom have Afrikaans surnames (66.7%). In most

conditions, especially X-linked recessive disorders such as haemophilia A, it is difficult to determine Afrikaans ancestry just based on surnames, as the mutation is passed down the maternal line while surnames are passed down the paternal lineage. Also Afrikaans ancestry is poorly defined when based on only one generation. This mutation will therefore be referred to as the common white founder mutation. It would be practical in the future to screen all white patients with haemophilia A, who test negative for the intron 22 inversion mutation, for the c.3637insA mutation.

Together with the fact that the intron 22 inversion mutation has one of the lowest reported frequencies in the world (Antonarakis, *et al.*, 1995) and that a potential founder haplotype was proposed by Dangerfield, *et al.* (1997) and McKibbin (2002), these data suggest that the c.3637insA mutation may be the elusive Afrikaans founder mutation. The suspicion of the founder haplotype is based on the many at-risk female relatives in big families who are uninformative on linked marker analysis and that the suspected founder mutation occurs on a common microsatellite haplotype. Further investigations to determine the frequency of this mutation in white haemophilia A patients from the diagnostic cohort and haplotype analysis were undertaken to confirm this hypothesis as described in sections 3.2.6.2 and 3.2.6.4.

3.2.6.1 The c.3637insA Mutation

The insertion of an A nucleotide at cDNA position 3637-3638 (shown in Figure 3.16) causes a frameshift mutation from the isoleucine at amino acid position 1194, with a premature stop codon occurring at amino acid position 1221 (Figure 3.17). The normal FVIII protein is 2351 amino acid in length, therefore in the presence of this mutation, only about half of the *FVIII* gene is translated. According to the HAMSTeRS insertion mutation database (available at http://europium.csc.mrc.ac.uk/WebPages/PublicFiles/Insertions.htm), this mutation has been reported 16 times previously.

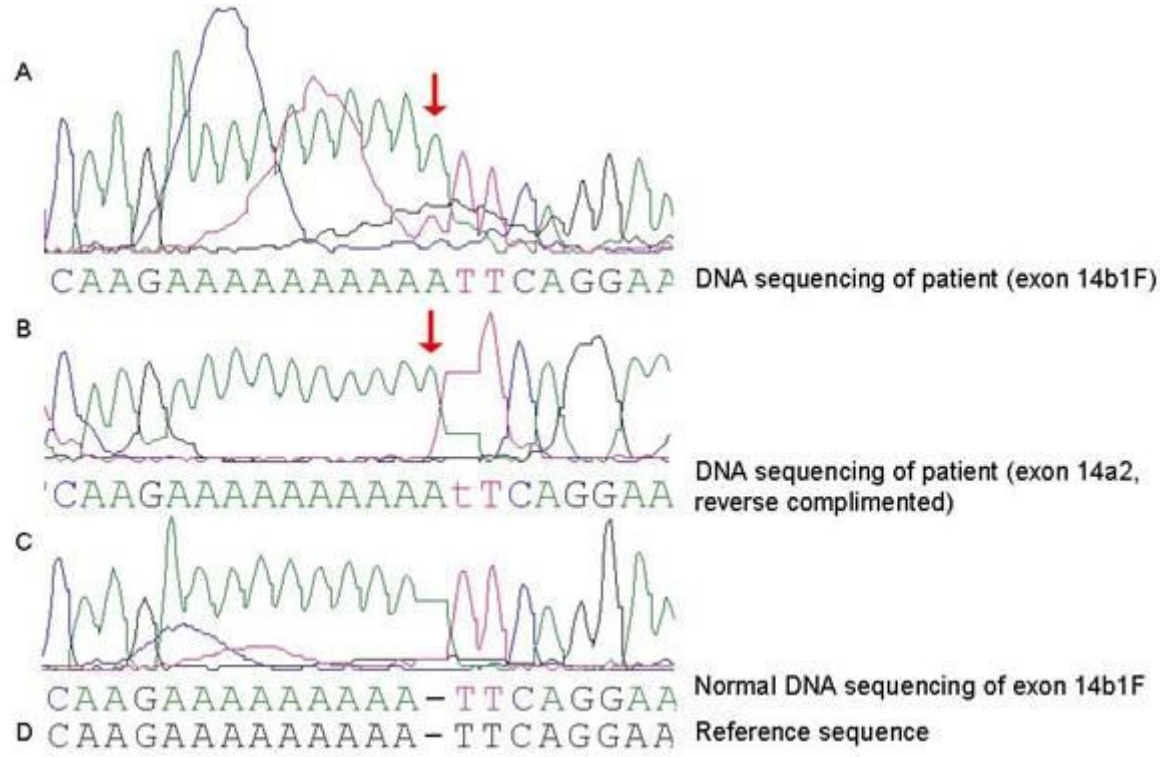

Figure 3.16: The raw electropherogram data obtained for a patient with the c.3637insA mutation.

A partial electropherogram of the forward sequence of exon 14b1 (A) and reverse complemented sequence of exon 14a2 (B) obtained from a patient with the mutation, while (C) is the electropherogram data obtained from a normal individual. The normal reference sequence is seen in (D). The c.3637insA mutation is highlighted by the red arrows. Sequence alignment obtained using the Seqman application of the DNAstar software (LaserGene). The apparent gap in the (c) and (d) electropherograms are a program artefact to highlight missing bases.

Figure 3.17: The amino acid sequence for the normal FVIII protein (A) and c.3637insA mutation (B), terminating at amino acid position 1221.

The difference from the normal sequence is shown in red, with X indicating the premature stop codon.

3.2.6.2 Frequency of c.3637insA Mutation in the White Population

Screening for the c.3637insA mutation was performed on all 30 of the white haemophilia A families in the diagnostic database who previously tested negative for the intron 22 inversion mutation. Results were obtained for 23 unrelated individuals and the mutation was found in three (13%). This frequency is

consistent with the frequency found in the study sample of 43 white patients (of 14%).

3.2.6.3 Pedigree of a Family Positive for c.3637insA Mutation

Most of the patients who tested positive for the c.3637insA mutation have a family history of haemophilia A which can be traced back a number of generations. Family HPA11 is an example of one such family. After the c.3637insA mutation was identified in the proband (IV6), a few other family members whose DNA samples were available in the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg were tested for this mutation. Figure 3.18 shows the pedigree of the members of the HPA11 family. Mutation analysis for the c.3637insA mutation found a second affected individual (IV3) to be positive, individuals II2, III2 and IV7 were found to be carriers and individuals III1 and III3 tested negative for the mutation and are therefore not carriers of haemophilia A.

3.2.6.4 Haplotype Analysis for c.3637insA Mutation in Patients

More conclusive haplotype analysis was required in c.3637insA positive patients to determine if this mutation is indeed a founder mutation or just a common recurrent mutation, especially as it has been reported 16 times in the HAMSTeRS database (section 3.2.6.1). Haplotype analysis was undertaken on all nine patients who tested positive for the c.3637insA mutation (six patients collected for mutation analysis for this project and the three additional patients from the diagnostic white patient screen), using polymorphic microsatellite markers within or close to the *FVIII* gene (section 3.2.6.4.1). Single nucleotide polymorphisms (SNPs) within the *FVIII* gene (section 3.2.6.4.2) were tested for in three patients. If a rare allele had been detected, the remaining patients would have been tested for this SNP.

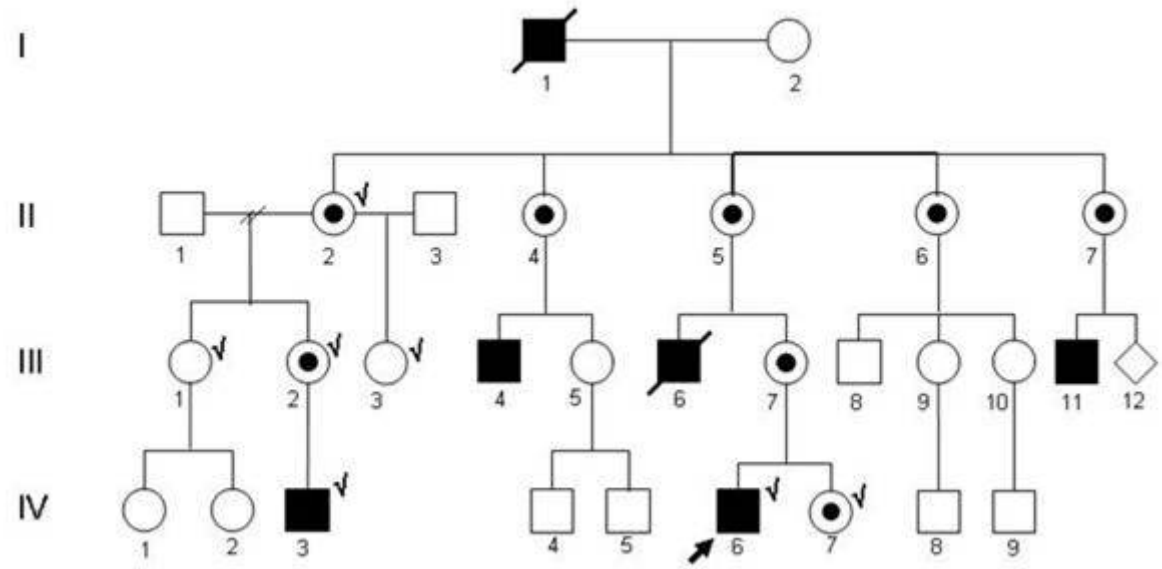

Figure 3.18: A pedigree of family HPA11, positive for the c.3637insA mutation.
Mutation testing for the c.3637insA mutation (for patients marked with a tick) found the proband (IV6) and individual IV3 positive for the mutation. Individuals II2, III2 & IV7 were found to be carriers for the mutation, and individuals III1 & III3 tested negative for the mutation.

3.2.6.4.1 Microsatellite Markers

Haplotype analysis was undertaken using polymorphic microsatellite markers in intron 13 and intron 22 of the *FVIII* gene and the p39 marker, an extragenic marker 500kb telomeric to the *FVIII* gene (described in section 2.3.10). These are the linked markers used in routine analysis undertaken by the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg to track the high risk X chromosome in haemophilia A families.

The results from the microsatellite markers show a common haplotype in 8 of the 9 patients in association with the c.3637insA mutation (refer to Table 3.2). This mutation tracks with the intron13-intron22-p39 haplotype of 20 repeats, 26 repeats, 160bp respectively. The one patient (HPA118) that did not share this common haplotype was found to have a haplotype of 20 repeats, **27** repeats and 160bp respectively.

Table 3.2: The haplotype analysis results for the intron 13, intron 22 and p39 microsatellite markers for patients positive for the c.3637insA mutation.

Patient/family	**Intron 13 (repeat numbers)**	**Intron 22 (repeat numbers)**	**p39 (bp)**
HPA005	20	26	160
HPA011	20	26	160
HPA015	20	26	160
HPA034	20	26	160
HPA050	20	26	160
HPA118	20	27	160
HPA192	20	26	160
HPA230	20	26	160
HPA249	20	26	160

3.2.6.4.2 Single Nucleotide Polymorphism Markers

A number of polymorphisms have been reported on the HAMSTeRS polymorphism database and by Viel, *et al.* (2007) (refer to section 3.2.4). It was decided to investigate 11 of these polymorphisms spanning the entire *FVIII* gene. These polymorphisms and the results obtained for three of the nine c.3637insA patients can be found in Table 3.3.

No rare single nucleotide polymorphisms were found in the three patients and therefore the SNP data does not further define or elucidate the founder haplotype. However, population frequencies for some of these SNPs have been determined in this project (refer to section 3.2.4) and the statistical analysis of the chance of finding this haplotype, together with the microsatellite haplotype in section 3.2.6.4.1, in the random population is calculated in section 3.2.6.4.3.

Table 3.3: Results from the single nucleotide polymorphism screen of the *FVIII* gene in 3 patients positive for the c.3637insA mutation.

Polymorphism	**Protein change**	***FVIII* position**	**Results**		
			HPA011	**HPA050**	**HPA230**
25610G→A	-	intron 2	G	G	G
29567C→T	-	intron 3	C	C	C
61534T→C	-	intron 9	T	T	T
61620G→A	R484H	exon 10	G	G	G
91317A→G	R776G	exon 14	A	A	A
92555C→T	H1188H	exon 14	C	C	C
92714C→G	D1241E	exon 14	C	C	C
92798A→C	S1269S	exon 14	A	A	A
92927G→A	K1312K	exon 14	G	G	G
120776T→A	-	intron 19	T	T	T
162161A→G	M2238V	exon 25	A	A	A

*3.2.6.4.3 Statistical Analysis of the c.*3637insA *Mutation Haplotypes*

Previous studies to determine the allele frequency for each of the microsatellite markers in the white South African population found allele 20 of the intron 13 marker to have a frequency of 0.52 (84/161) (Dangerfield, *et. al*, 1997), allele 26 of the intron 22 marker to have a frequency of 0.58 (93/161) (Dangerfield, *et. al*, 1997) and allele 160 of the p39 marker to have a frequency of 0.17 (12/71) (McKibbin, 2002). The statistical analysis of the chance of finding this haplotype, together with the SNP haplotype in section 3.2.6.4.2, in the random population is calculated using the Fisher's exact test as shown in Table 3.4. The Fisher's exact test is used in the analysis of data for small sample sizes (n<10) and examines the significance of an association between two variables in a 2x2 contingency table (http://udel.edu/~mcdonald/statfishers.html).

As there were no haplotype data for control samples, each marker had to be compared separately. These probabilities indicate that the allele observed at each microsatellite loci occurs far more commonly than expected (p=0.000001 to 0.053) and is therefore likely to be a founder mutation. The SNP haplotype data do not

differ significantly from the expected frequencies from the control samples (p=0.5 to 0.8), however, the sample size for the patient group was extremely small (n=3).

Table 3.4: The statistical probabilities of the common c.3637insA haplotype occurring in the South African white population.

Marker	Patients	Controls	Exact probability[1]	Exact probability[2]
Intron 13	9/9	84/161	0.004	0.004
Intron 22	8/9	93/161	0.061	0.053
p39	9/9	12/71	0.000001	0.000001
D1241E	3/3	18/23 *	0.512	0.5
S1269S	3/3	21/24 *	0.692	0.7
M2238V	3/3	13/14 *	0.824	0.8

[1] The Fisher's exact probability obtained from http://www.langsrud.com/fisher.htm

[2] The Fisher's exact probability determined from http://www.physics.csbsju.edu/cgi-bin/stats/exact

* Allele frequencies obtained from the white patients this study (section 3.2.4)

3.2.7 Mutation Summary

Of the 120 patients analysed, a causative mutation was identified in a total of 88 patients (73.3%). In 32 patients (26.7%) the disease-causing mutation was not identified. This was most likely due to technical problems related to poor mRNA quality, however, large deletions and rearrangements cannot be excluded.

DNA analysis identified the intron 22 inversion in 40.9% (36/88) patients and the intron 1 inversion in 3.4% (3/88) patients, while mRNA analysis identified a mutation in 55.7% (49/88) patients. Of the 49 mutations identified through mRNA analysis, 28 patients (57.1%) have a point mutation (17 missense (34.7%), 9 nonsense (18.4%) and 2 splice-site mutations (4.1%)), 9 patients (18.4%) have a deletion and 7 patients (14.3%) have an insertion. Another 5 patients (10.2%) have a complex mutation where an exon was deleted on mRNA analysis, but no mutation detected on DNA analysis.

In this study, it was found that the intron 22 inversion frequency accounted for 32.5% (14/43) of mutations in the white population, 50% (19/38) of mutations in the black population and 42.8% (3/7) of mutations in the Indian and mixed ancestry populations (as discussed in section 3.2.1.1). An illustration of the breakdown of the different mutation types identified per population group can be found in Figure 3.19. It is likely that the increase in the frequency of insertion and deletion mutations in the white population is due to the presence of the c.3637insA founder mutation (described in section 3.2.6).

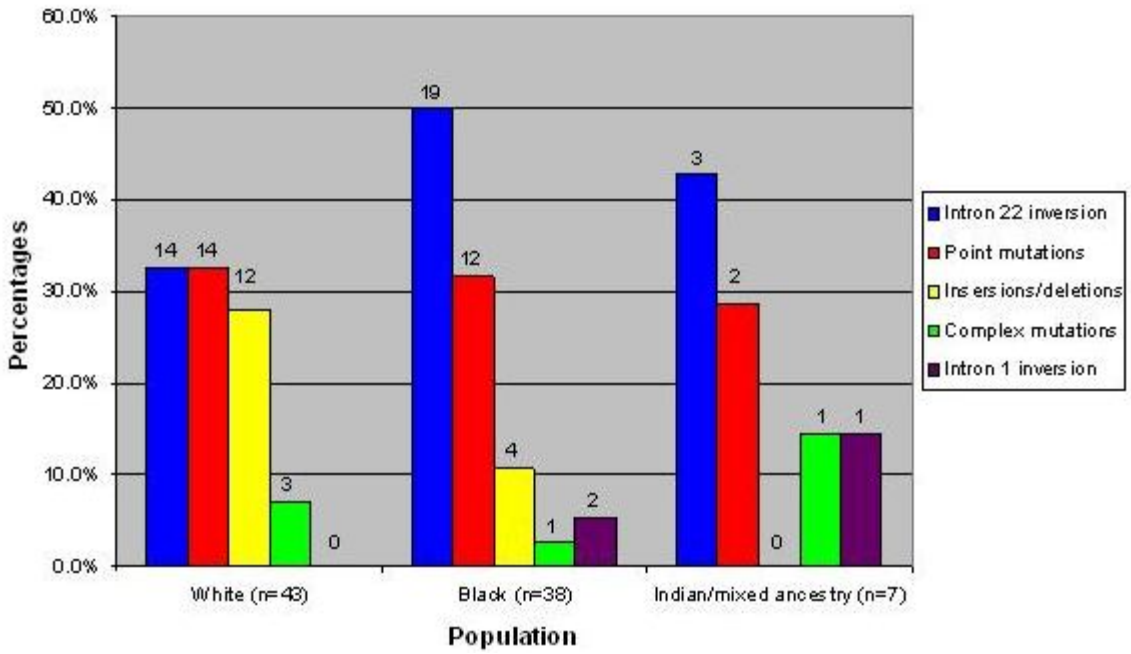

Figure 3.19: A graphical illustration of the different mutation types identified per population group.

The distributions of the different types of mutations, identified by mRNA analysis, in the *FVIII* gene can be found in Figure 3.20. The intron 1 and intron 22 inversion mutations have been excluded from this figure. It is evident that most mutations (24/49 = 48.9%) are found in exon 14 of the *FVIII* gene. This is not surprising as this exon is about 3.1kb in length – almost a third of the mRNA transcript.

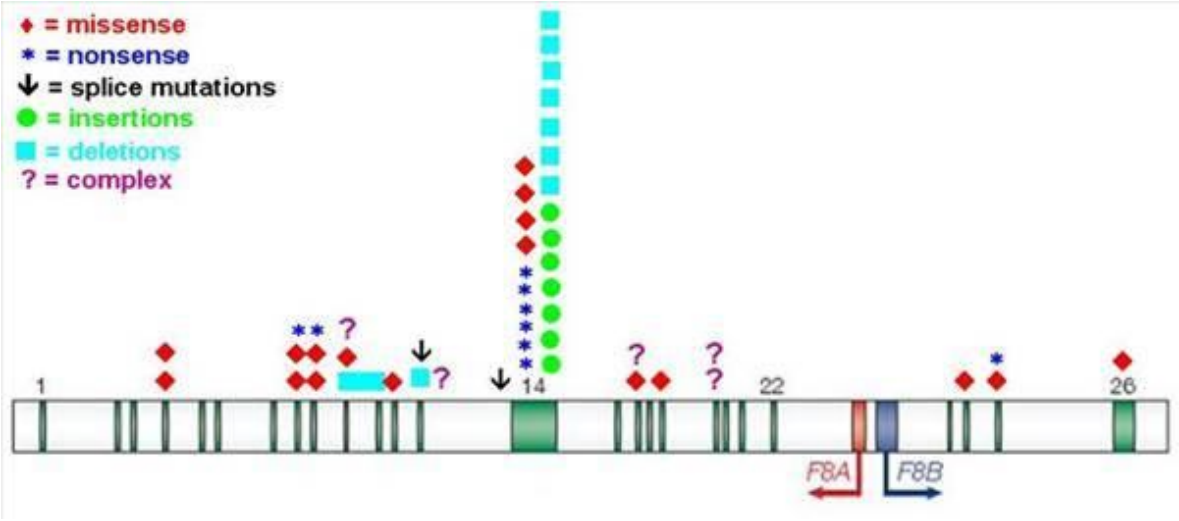

Figure 3.20: A schematic representation of the different types of mutations identified by mRNA analysis in the *FVIII* gene in this study.
The intron 22 and intron 1 inversions have been excluded from this analysis.

A comparison of the frequencies of the different types of mutations in the South African population identified by mRNA analysis versus studies undertaken by Green, *et al.* (2008); Oldenburg, *et al.* (2004b) and Astermark, Oldenburg, Escobar, *et al.* (2005) can be found in Figure 3.21. Green, *et al.* (2008) studied a third of the haemophilia A patients in the UK; Oldenburg, *et al.* (2004b) reviewed all mutations referenced in the HAMSTeRS mutation database and the Astermark, *et al.* (2005) study used data from the Malmö international brothers study. Our study, together with the Green *et al.* (2008) and Oldenburg. *et al.* (2004b) studies determined the frequencies, using mutation data from patients with all severities of haemophilia A. The Astermark, *et al.* (2005) study, however, only assessed patients with severe haemophilia A.

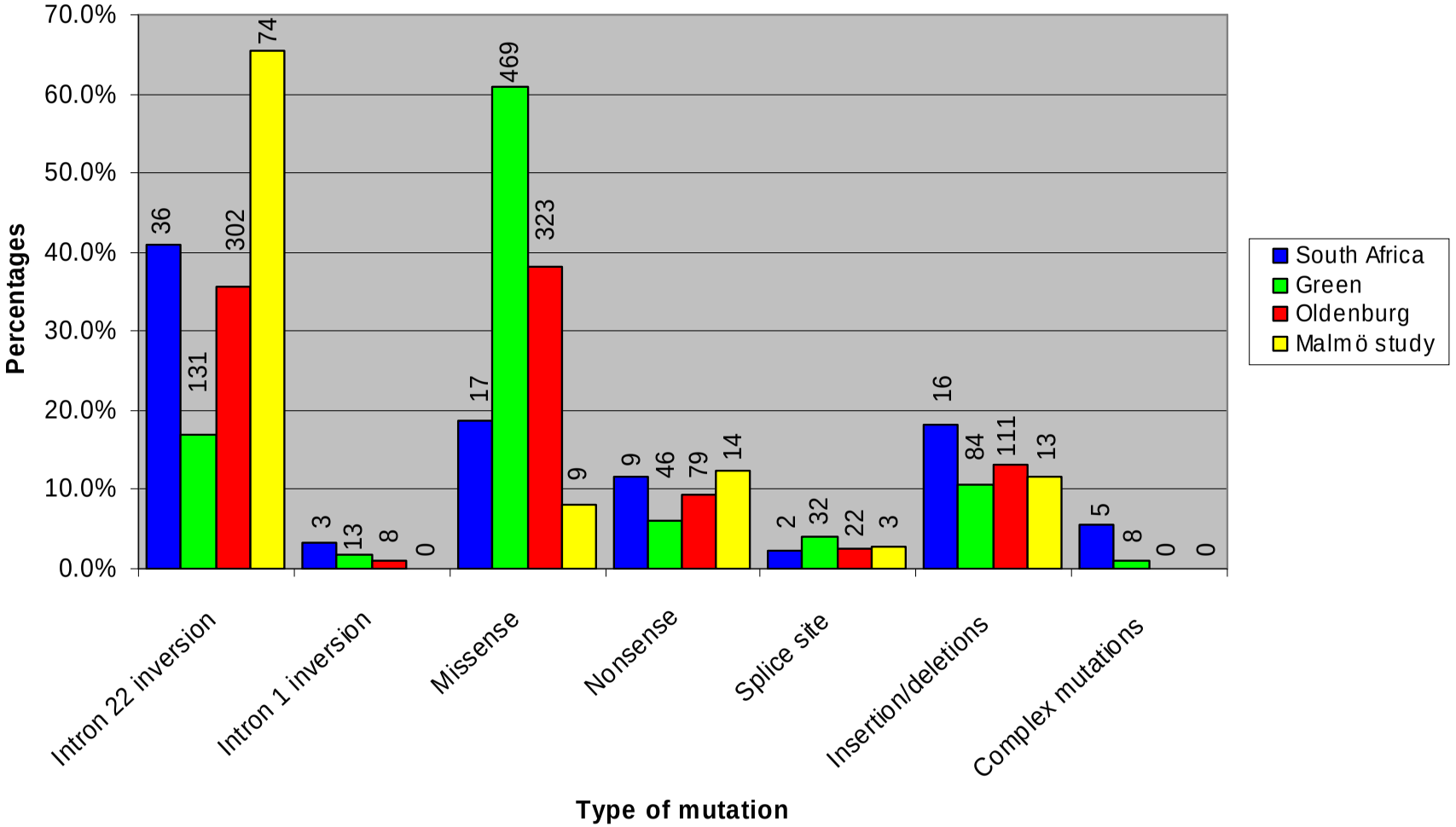

Figure 3.21: A comparison of the percentage of the different types of mutations in the *FVIII* gene in South Africa versus the world percentages.

World percentages were based on studies by Green, *et al.* (2008) (n=769); Oldenburg, *et al.* (2004b) (n=846) and the Malmö study (Astermark, *et al.*, 2005) (n=113). The South African study cohort consisted of 88 patients.

Intron 22 inversion, insertion/deletion and complex mutations appear to occur at a higher frequency in the South African population when compared to the rest of the world, whereas missense mutations appear to be lower in the South African patients. Missense mutations occur less frequently in patients with severe haemophilia A compared to other severities (Astermark, *et al.*, 2005), and the patient cohort of this study has a large number of patients with severe haemophilia A. This bias in severe haemophilia A patients could explain the increased frequency of the intron 22 inversion mutation in the South African population as a whole. While the increase in insertions is not unexpected due to the presence of the c.36937-3638insA founder mutation in the white population (refer to section 3.2.6). A statistical comparison of these study groups is difficult as they were collected under different conditions and therefore has not been performed.

3.2.7.1 Novel and Previously Reported Mutations

The novel and previously described mutations are described for each mutation type as follows: for nonsense mutations (section 3.2.3.1), for missense mutation (section 3.2.3.2), for deletion and insertion mutations (section 3.2.3.3) and for mutations affecting splicing (section 3.2.3.4).

To summarise, of the 49 non intron 22 or intron 1 inversion mutations identified in this project by mRNA analysis, 23 (46.9%) were previously reported (6 of which are the common white mutation, c.3637insA) and 20 (40.8%) were novel mutations not previously reported on the database. As we were unable to identify the exact DNA change resulting in 6 mutations (12.2%), we were unable to determine if they were novel or previously reported.

4 DISCUSSION

4.1 INTRODUCTION TO THE CHAPTER

In this chapter, the results of this project and the implications of these findings will be discussed. It will also discuss aspects of the project such as mRNA analysis (including technical issues, cost effectiveness, advantages and disadvantages), and make suggestions about the implementation of this new technique into the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg. mRNA analysis has proved to be an important technique to identify an increased number of mutations in Haemophilia A patients.

4.2 *FVIII* MUTATIONS IDENTIFIED IN THE PROJECT

4.2.1 Intron 22 Inversion

In this study, the intron 22 inversion mutation was found in 34.6% (36/104) of South African haemophilia A patients. The world frequency of this mutation is 42-50% in severe haemophilia A (Lakich, *et al.*, 1993; Naylor, *et al.*, 1993; Antonarakis, *et. al*, 1995). The patients in this study were not selected for haemophilia A severity, although there was a bias towards patients with severe disease due to sample ascertainment.

Analysis of the intron 22 inversion mutation at a population level, found the mutation in 27.5% (14/51) of white patients, 42.2% (19/45) of black patients and 37.5% (3/8) of the Indian and mixed ancestry patients. The frequency of the intron 22 inversion mutation is significantly lower (p=0.008) in the white population. Although it remains one of the lowest reported frequencies, it is similar to that reported in severe haemophilia A patients in France (28%) (Antonarakis, *et al.*, 1995). The frequency of this mutation in the black population does not deviate (p=0.413) from the expected world frequency of 45%. The frequency of the intron

22 inversion mutation in the black population in this study (42.2%) is similar to that previously reported by Dangerfield, *et al.* (1997) as 43% (6/14) in South Africa.

In the South African population studied in this project, the type I inversion accounts for 55.6% (20/36) of intron 22 inversion mutations, the type II inversion for 38.9% (14/36) and the atypical pattern of the inversion accounts for 5.6% (2/36). In the Antonarakis, *et al.* (1995) study, the type I inversion accounted for 82% (740/905) of intron 22 inversion mutations, while 15% (140/905) were the type II inversion and an atypical pattern account for the remaining 3% (25/905) of intron 22 inversion mutations. The type I inversion is lower in the South African population, while the type II inversion is higher than the reported world frequencies. A chi-squared statistical comparison of these differences was assessed (χ^2=835) and resulted in a probability of <0.0001.

The sample sizes were too small to draw any conclusions as to whether there are differences in the distribution of the different types of the intron 22 inversion mutation between the different South African populations. However, the type II inversion was found in 42.1% (8/19) of the intron 22 inversion mutations in the black population. Previously, the type II inversion was not found in this population (Dangerfield, *et al.*, 1997), but this could have been due to a small sample size, in the previous study.

As the intron 22 inversion mutation is still the commonest mutation in haemophilia A, it is important to test for this mutation first in any mutation analysis.

4.2.2 Intron 1 Inversion

Direct testing for the intron 1 inversion mutation by PCR analysis was undertaken on all patients. This mutation was identified in three patients, two of whom are black patients (5.3% of all mutations) and the third patient is of Indian ancestry (25% of all mutations). A study of haemophilia A patients from India found the intron 1 inversion mutation to account for 2.7% (2/75) of all mutations (Ahmed, *et al.*, 2005).

It is surprising that this mutation was not found in any white patients in this project (n=49) nor in the preliminary screen of 25 white patients from the diagnostic haemophilia A cohort. This mutation accounts for approximately 4.8% (10/209) of severe haemophilia A mutations in the United Kingdom (Bagnall, *et al.*, 2002) and 15% (3/21) of mutations in severe haemophilia A patients in the Netherlands (Boekhorst, *et al.*, 2005).

The intron 1 inversion mutation is an important mutation to screen for in haemophilia A mutation analysis, even though it has not been identified in the white population as yet. As it is a rapid PCR assay, it should be performed as a second line mutation test, after testing for the intron 22 inversion mutation.

4.2.3 Mutations Identified using mRNA Analysis

mRNA analysis of haemophilia A patients identified mutations in an additional 55.7% (49/88) patients, thus significantly improving the mutation detection rate compared to the conventional testing strategy currently offered by the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg, for the intron 22 inversion mutation only.

The mutation profile in intron 22 inversion negative patients in the South African population showed a wide variety of different types of mutations: 34.7% (17/49) missense mutations, 18.4% (9/49) nonsense mutations, 4.1% (2/49) splice-site mutations, 18.4% (9/49) deletion mutations and 14.3% (7/49) insertion mutations. The remaining 10.2% (5/49) patients had a complex mutation where an exon was deleted on mRNA, but no mutation was detected at the DNA level.

When compared to other studies (such as Oldenburg, *et al.*, 2004b; Astermark, *et al.*, 2005; Green, *et al.*, 2008), it was found that complex point mutations together with insertion/deletion mutations appear to occur at higher frequencies in the South African population compared to the rest of the world, whereas the frequency of missense mutations in the South African population is lower than expected. However, the fact that insertion/deletion mutations are higher in the South African

population may be explained by the presence of the c.3637insA founder mutation in white patients. In both this study and the Astermark, *et al.* (2005) study, the decrease in frequency of missense mutations, may be explained by the bias of severe haemophilia A patients. Missense mutations are lower in patients with severe haemophilia A than patients with other severities of haemophilia A (Astermark, *et al.*, 2005). Astermark, *et al.* (2005) only studied severe haemophilia A patients, which also explains the high frequency of intron 22 inversion frequency in this patient group.

In summary, a mutation was identified in 76.8% (43/56) of white patients, 67.9% (38/56) of black patients and 87.5% (7/8) of Indian and mixed ancestry patients. mRNA analysis of the *FVIII* gene allows for an increased mutation detection rate in patients. Mutation results have major implications for family members at risk of being carriers of haemophilia A as it can confirm or refute a high-risk result determined by linked marker analysis. There were no patients in whom a mutation was not found after full sequencing analysis of the *FVIII* mRNA, and the patients without results are more likely to be due to technical difficulties rather than the inability of the technique to identify the pathogenic mutation. However, some large deletions, insertions and rearrangements may have been missed in these patients.

4.2.4 Novel and Previously Reported Mutations

Of the 49 mutations identified on mRNA analysis, almost half, 46.9% (23/49) (including six copies of the c.3637insA white founder mutation), were previously reported on the HAMSTeRS mutation database or by Green, *et al.* (2008), while an additional 40.8% (20/49) were novel mutations. As the exact DNA change resulting in 12.2% (6/49) of mutations identified on mRNA analysis could not be identified, we can not comment on whether these mutations are novel. These mutations may be deep intronic mutations affecting splicing or deletions where the breakpoints could not be defined. Screening for the DNA cause of these mutations was only conducted by looking for mutations in the exons and the intron/exon boundaries of the PCR fragments described in 2.3.7. Due to the high costs involved and time constraints, it was not possible to screen the entire flanking

intronic sequences for mutations. Although, the exon 10-11 deletion in patient HPA202 could be confirmed using the haemophilia A MLPA kit available from MRC Holland, this has not been undertaken due to the costs involved in ordering and verifying the kit in this laboratory.

A report by Cutler, Mitchell, Smith, *et al.* (2002) found that 50.6% (41/81) of mutations identified in their intron 22 inversion negative patient cohort were novel. Whereas a study by Green, *et al.* (2008) found that 27.7% (176/636) of mutations resulting from a screen of one third of the UK population to be novel. As about one third of haemophilia A patients are sporadic cases (reviewed in Graw, *et al.*, 2005), the new mutation rate in the *FVIII* gene is high and therefore, many novel mutations would be expected.

4.2.5 Clarification of Linked Marker Results

Linked marker analysis is useful in cases where the mutation has not been identified as it is able to track the inheritance of the high risk allele within a family. However, it may be uninformative in a family and it does not take into account the possibility of a new mutation in the patient, or the risk of germline mosaicism in the mother. Direct mutation detection is optimal for accurate and informative carrier testing as it can confirm (example shown in section 3.2.5.1) or decrease the high risk (described in section 3.2.5.2) determined by linked markers in a female, and possibly reduce the need for testing in other family members or provide definitive testing in other family members. Even in cases where germline mosaicism, or a new mutation, cannot be determined, the mother can be offered conclusive prenatal diagnosis by mutation testing. Direct testing may also be useful in cases where linked marker results are uninformative (examined in section 3.2.5.3) or where key family members are unavailable for linked marker analysis.

4.2.6 Polymorphisms

Five polymorphisms or benign variants were identified in this project, four of which have been previously reported (Viel, *et al.*, 2007). The white population shows

variation at three of these sites, namely D1241E, S1269S and M2238V. All five polymorphisms and variants were present in the black population. The Indian and mixed ancestry population sample size was too small to draw any conclusions.

These polymorphisms and their frequencies may potentially be useful in linkage analysis in determining the high-risk X chromosome, especially in cases where the disease-causing mutation can not be identified and where other markers are uninformative.

The novel variant, A544A was only found in one patient from the black population at a frequency of 0.04 (1/28). It is suggested that more individuals from the black population be screened for this variant to determine if it is indeed a polymorphism in the population or just a rare variant in patient HPA183, before using this variant as a linked marker in future linkage analysis.

4.2.7 Common Founder Mutation in the White Population

The c.3637insA mutation was found at a frequency of 14% (6/43) in patients from the white population. As this was the only mutation that occurred more than once in the white population, haplotype analysis was performed to determine if this mutation was the elusive founder mutation previously proposed (Dangerfield, *et al.*, 1997; McKibbin, 2002), or just a common recurrent mutation as it has been previously reported 16 times in the HAMSTeRS mutation database.

Using haplotype analysis for 3 polymorphic microsatellite markers (intron 13, intron 22 and the p39 extragenic marker), a common haplotype of 20 repeats, 26 repeats and 160bp was found in 8/9 of patients positive for the c.3637insA mutation. The only patient that did not have this haplotype was found to have a haplotype of 20 repeats, 27 repeats and 160bp. It is possible that this single repeat difference for the intron 22 marker may be due to replication slippage, or it may be an independent mutation in this patient. No rare SNP was found in three of these patients to refine the haplotype further, but they also share the same SNP haplotype.

The control group used microsatellite frequencies obtained by Dangerfield, *et al.* (1997) and McKibbin (2002). As haplotype data were not available for the control group, each marker had to be compared separately. The probabilities indicate that the allele observed at each of the microsatellite loci in the c.3637insA patients occurs far more commonly than expected (p=0.000001 to 0.0530). The SNP haplotype data in the patient group did not differ significantly from the expected frequencies from the control samples (p=0.5 to 0.8), which may be explained by the fact that no rare alleles were observed in the patient cohort. The microsatellite data therefore provide strong evidence that the common haplotype identified in the patients is likely to be a founder mutation and not a common recurrent mutation.

It is interesting to note that together the c.3637insA mutation (14%) and the intron 22 inversion mutation (32.5%) account for 46.5% of disease-causing mutations in the white population. This is comparable to the frequency of the intron 22 inversion alone in most world populations (Antonarakis, *et al.*, 1995).

The c.3637insA mutation can be added to the list of other founder mutations in the South African white population, such as porphyria variegata (Dean, 1971), familial hypercholesterolemia (Jenkins, *et al.*, 1980), pseudoxanthoma elasticum (Le Saux, Beck, Sachsinger, *et al.*, 2002); *BRCA1* in breast/ovarian cancer (Reeves, Yawitch, van der Merwe, *et al.*, 2004); Fanconi Anaemia (Tipping, Pearson, Morgan, *et al.*, 2001) and autosomal recessive polycystic kidney disease (unpublished data). These founder effects are all autosomal dominant or recessive disorders, whereas the haemophilia A founder mutation is the first report of an X-linked founder mutation in the white population of South Africa. However, this is not the first case of a founder mutation causing haemophilia A. Santacrose, Santoro, Sessa, *et al.* (2008) report on a founder mutation in the *FVIII* gene in mild and moderate haemophilia A patients in southern Italy.

It is suggested that the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg, include testing for this mutation into their routine mutation testing strategy, especially for white patients that test negative for the intron 22 inversion (refer to section 4.4).

4.3 THE USE OF mRNA FOR MUTATION DETECTION

Analysis of mRNA for haemophilia A mutation detection has proved successful in previous studies. For example, in the identification of the common intron 22 inversion mutation (Naylor, *et al.*, 1992) and the intron 1 inversion mutation (Bagnall, *et al.*, 2002). mRNA analysis has proved useful in finding deep intronic mutations and in assessing the functional impact of such mutations on the gene, and determining the consequences of some mutations (Green, *et al.*, 2008).

mRNA analysis is useful for genes with large intronic sequences or for large genes as it assesses the transcript which is normally smaller in size. Some disorders in which mRNA analysis has proved useful include Duchenne and Becker Muscular Dystrophy (Roest, *et al.*, 1996), breast and ovarian cancers (Rhei, *et al.*, 1998; Spearman, *et al.*, 2008), and neurofibromatosis (Pros, *et al.*, 2008). *FVIII* is also a perfect candidate for mRNA analysis as it is a large gene (186kb of genomic DNA) and DNA analysis was not previously as successful as expected (Dangerfield & Manga, unpublished data). Also, there are a large number of rearrangements in the *FVIII* gene that would not be identified by analysing the DNA alone. The 9kb *FVIII* mRNA transcript can be assessed in nine overlapping fragments, compared to analysis of DNA which would require the 26 exons to be amplified in 29 reactions. Analysis of mRNA would therefore be expected to be more cost effective than DNA analysis, because fewer fragments are amplified and it is able to detect a wider variety of mutations, such as large rearrangements, deletions, insertions and cryptic mutations that would be missed by studying the exonic DNA. mRNA analysis also detects mutations readily detected by DNA analysis such as obvious splice site and point mutations.

Studying the gene at the mRNA level would not be able to detect mutations within the promoter region which would affect the regulation of gene expression rather than the actual protein produced (Mueller & Young, 1998). There is also a report of a patient with severe haemophilia A, who tested negative for the intron 22 and intron 1 inversion mutations and the coding sequences at a DNA level, and in whom there was no detectable mRNA for the *FVIII* gene. This mutation causing haemophilia A in the patient is therefore likely to be due to the absence or rapid degradation of the *F8* mRNA (El-Maarri, Singer, Klein, *et al.*, 2006).

The main aim of the project was to assess an mRNA based approach to develop a more effective mutation detection system that would detect more mutations and reduce the costs of DNA mutation analysis. Initially, this project intended to use PTT as a mutation detection technique, however, it was decided this technique was too costly and limited as it would not detect all types of mutations (as discussed in section 4.3.1.1). PTT only detects mutations that truncate the protein of interest.

4.3.1 Advantages of mRNA Analysis

mRNA has numerous benefits over DNA as a medium for analysis for example, the cost benefits and the mutation detection rate. These are discussed in sections 4.3.1.1 and 4.3.1.2.

4.3.1.1 Cost Effectiveness of Different Mutation Analysis Methods

mRNA analysis for the *FVIII* gene analyses nine overlapping fragments in 24 sequencing reactions (two sequencing reactions each for fragments 1a, 1b, 4a, 4b, 6a and 6b, and four sequencing reactions each for fragments 2, 3 and 5) (refer to section 2.3.6). DNA analysis of the 26 exons of the *FVIII* gene, requires 58 sequencing reactions (two sequencing reactions for each exon, except for exon 14 which requires eight sequencing reactions to cover this large exon) (refer to section 2.3.7). mRNA analysis therefore drastically reduces the sequencing cost to approximately half, compared to DNA mutation analysis for haemophilia A, as seen in Table 4.1, and also reduces the analytical work required, thereby reducing the time required for analysis. Further, full mRNA analysis may not be required in all patients. A few examples where fewer mRNA sequencing reactions may be required are illustrated below:

1. If a patient's PCR product differs from the expected size, it is likely to indicate an insertion or deletion at the RNA level and only this fragment would be sequenced initially. However, the genomic cause could be an insertion, deletion or point mutation resulting in abnormal splicing.

2. Once a pathogenic mutation has been identified in a patient, further mutation analysis on other fragments may be discontinued.

3. As many mutations were identified in exon 14, analysis of this exon may be done first (either at a DNA or mRNA level) and if a mutation is identified, analysis of the remaining *FVIII* gene may not be required.

Table 4.1: A cost comparison of mRNA versus DNA analysis for the *FVIII* gene per patient.

Analysis	**Technique**	**Approximate cost per patient**	**Total cost per patient**
mRNA	RNA extraction	R95	**R2620**
	Reverse transcription	R65	
	Sequencing (24 reactions)	R2300	
DNA	DNA extraction	R40	**R5710**
	Sequencing (58 reactions)	R5670	

Originally this project intended to use the protein truncation testing technique which selectively identifies certain mutations that result in premature protein truncation (including nonsense mutations, insertions, deletions, gene rearrangements, various splicing mutations) and possibly even detection of mutations that affect protein folding. This technique involves reverse transcription of mRNA, PCR amplification followed by *in vitro* transcription and translation of the PCR products. The protein products are analysed using SDS-PAGE (reviewed by Den Dunnen & van Ommem, 1999). Any truncated proteins would be analysed further, usually by sequencing analysis, to determine the exact nature of the mutation. Maugard, *et al.* (1998) found the protein truncation test identified the disease-causing mutation in 66% of inversion 22 negative, severe haemophilia A patients in France.

After much consideration, it was decided not to use the protein truncation test in this project, mainly for cost reasons. Not only does protein truncation testing require amplification of the nine overlapping cDNA fragments, but it also requires

nine reactions of the *in vitro* transcription/translation kit (containing cell-free extracts of prokaryotic or eukaryotic cells with all the machinery components and polymerases required for coupled transcription and translation). Further, this technique does not detect missense mutations or mutations that may affect protein function rather than protein structure. In these cases, mRNA sequence analysis would still be required in addition to the protein truncation test to identify the disease-causing mutation, thus adding costs. A comparison of the costs is broadly shown in Table 4.2. The costs for the protein synthesis kit and sequencing are based on the suppliers' 2009 prices.

Table 4.2: Comparison of the broad cost estimates per patient for protein truncation testing versus sequencing analysis of *FVIII* mRNA.

Situation	Technique	Approximate costs	Total
PTT identifies mutation	RNA extraction	R95	**R5860**
	Reverse transcription	R65	
	Protein synthesis kit	R611/reaction x 9 = R5500	
	Sequencing	R200	
PTT does not identify mutation	RNA extraction	R95	**R7960**
	Reverse transcription	R65	
	Protein synthesis kit	R611/reaction x 9 = R5500	
	Sequencing	R2300	
mRNA sequencing only	RNA extraction	R95	**R2460**
	Reverse transcription	R65	
	Sequencing	R2300	

4.3.1.2 Detection of Additional Mutations using mRNA Analysis

Another major benefit for mRNA analysis includes the detection of additional mutations (deletions, insertions and deep intronic mutations that affect splicing) that would otherwise have been missed using DNA analysis.

In this study, the disease-causing mutation would have been missed in 12.2% (6/49) of patients if testing was only conducted using DNA analysis. These

mutations include partial, single or multiple exon deletions (discussed in sections 4.3.1.2.1, 4.3.1.2.3 and 4.3.1.2.4) and insertions of intron sequence into the mRNA transcript. Some other mutations may have been misclassified as missense mutations instead of mutations affecting splicing (discussed in section 4.3.1.2.2). Some of these cases are described below.

4.3.1.2.1 Example of a Deep Intron Site Mutation

In patient HPA219 (refer to section 3.2.3.3), a deletion of exon 10 was detected on mRNA analysis; however DNA analysis of exon 10 did not detect the disease causing mutation within exon 10 or the intron/exon boundaries. It is assumed that there is a deep intronic or cryptic splice site mutation that causes exon 10 to be spliced out of the mRNA sequence. This mutation would have been missed using DNA analysis alone.

4.3.1.2.2 Example of a Cryptic Splice Site Mutation

Patient HPA236 was found to have a deletion of exon 13 on mRNA analysis (refer to section 3.2.3.4.1). In this case, although DNA analysis of exon 13 found a c.2111C→T mutation which appears to cause a P687L substitution, mRNA analysis proved that this mutation would actually affect the splicing of exon 13. Here the effect of the mutation would have been misinterpreted as a missense mutation rather than a splice site mutation and the severity perhaps may have been underestimated. This mutation was not predicted by the online splice site predictor programs and emphasises the importance of mRNA analysis in determining the effect of a mutation.

4.3.1.2.3 Example of Alternative Splicing

DNA sequencing of patient HPA223 of exon 16 found no mutation, except for a IVS16+37delT (the pathogenicity of which is unknown). Initial mRNA analysis showed a partial deletion of exon 16. Closer inspection of the sequencing result provided evidence of alternative splicing of exon 16 (refer to section 3.2.3.4.3). This mutation is therefore only able to be detected using mRNA analysis. The IVS16+37delT may be used at least, as an additional linked marker to track the

high-risk X chromosome in this family, even though the pathogenic mutation remains unknown.

4.3.1.2.4 Example of a Multi-exon Deletion

In patient HPA210, exons 10 and 11 were found to be deleted on mRNA analysis. Using DNA analysis the two exons could not be amplified. This was strongly suggestive of a deletion spanning these two exons. If DNA analysis was used in isolation, it may have been assumed that amplification of these two exons was problematic. However, the breakpoints of this mutation have not been defined due to the large size of introns 9 and 11 and time and cost constraints of the project. Once the breakpoints are defined, a PCR for this family may be specifically designed to aid in carrier testing and prenatal diagnosis. MLPA analysis could be used as an alternate mutation detection method, or DNA amplification of exon 10 and 11 using a dosage assay, however, this technique would need careful analysis and at least 5 control samples. Currently, testing of other family members can only be offered using mRNA analysis, which appears to be the best method at present.

4.3.1.3 Detection of Carriers using mRNA

For mutations such as large deletions, rearrangements or complex mutations, direct mutation detection can be offered to family members at risk of haemophilia A by using mRNA, as these mutations are either unable to be detected using DNA, or are easier to detect at an mRNA level. An alternate method of detection would need to be optimised before carrier testing and prenatal diagnosis could be offered to family members. Other possible methods of mutation detection of such mutations in female relatives include: use of linkage analysis which may reveal loss of heterozygosity for the markers in the deleted region; gene dosage analysis and MLPA could be applied to the *FVIII* gene as suggested by Keeney, *et al.* (2005). However, these methods would not be useful if the deletion is not present at the DNA level.

4.3.2 Disadvantages of mRNA Analysis

4.3.2.1 Stability of mRNA

mRNA is less stable than DNA. This is mainly due to the fact that RNA is single stranded and highly prone to relatively rapid degradation. This is important in the transport and storage of patient samples.

4.3.2.1.1 Transport of Patient Samples

The stability of mRNA is a major consideration in transportation of samples for diagnostic testing. In this project, it was found that it is best to extract mRNA within 24 hours of sample collection, which may not always be possible when the sample is being transported from a remote location.

To overcome this limitation, there are two potential options available. The first is the PAXgene product (Qiagen) and the other involves the addition of TRIzol (Invitrogen) to the blood sample immediately after collection. Both possibilities are described below.

4.3.2.1.1.1 PAXgene Blood RNA kit

PAXgene Blood RNA kit is a joint kit that includes blood collection tubes and the extraction kit. It allows stabilisation of RNA for 3 days at 18-25°C or 5 days if stored at 2-8°C. After extraction, the cellular RNA is stable for 6 months at -20°C to -70°C (product catalogue and details are available from the Qiagen website http://www1.qiagen.com/Products/RnaStabilizationPurification/DSP/PaxGeneBloodRnaKit.aspx?r=3710). Although it is reported that this kit is compatible with systems previously optimised using the QIAamp RNA blood mini kit, it was not tested in this project. It would be important to try this option as it would allow an easy alternative to overcome the limitations with mRNA stability and allow transport of samples from other centres around the country to be sent in for diagnostic testing.

It would need to be ensured that haemophilia centres around the country had access to these blood tubes and the shelf-life of the tubes would need to be determined. Also the cost of the tubes is apparently high, and may increase the cost of the test.

4.3.2.1.1.2 TRIzol stabilisation

TRIzol is a chemical that is added to a blood sample to stabilise the RNA and it is used in the extraction process. Unfortunately, TRIzol contains phenol and guanidine isothiocyanate, meaning that special safety precautions are required when working with this solution (further information is available at the Invitrogen website at http://www.invitrogen.com/site/us/en/home/brands/Product-Brand/Trizol.html). The TRIzol method requires some processing in a laboratory at the site of sample collection, and the samples need to be transported frozen, both of which may be problematic when blood needs to be sent from distant locations. In addition, TRIzol requires a different method of extraction to that of the QIAamp RNA blood mini kit (the protocol for mRNA extraction from TRIzol stabilised blood can be found in Appendix B). A sample of TRIzol stabilised blood from a normal individual was obtained from Dr Marius Coetzee (Department of Haematology and Cell Biology, University of the Free State, Bloemfontein) to determine if TRIzol stabilised samples would be able to be used in this project. Amplification of the RNA fragments works well without further optimisation.

4.3.2.1.2 Long Term Storage of Patient Samples

DNA samples from patients are still viable after many years if stored in TE buffer at 4°C, however, during this study it was found that mRNA was only stable for a year at -70°C (also stated in the QIAamp RNA Blood Mini kit manual). The mRNA samples of haemophilia A patients can not be banked to be used as a positive control at a later stage, should other at-risk family members require testing, especially for those patients where the disease-causing mutation is only detectable on mRNA analysis. The long-term stability of cDNA has not been assessed in this project.

4.3.2.2 Failure Rate of mRNA Analysis

In this project, mutations were not detected in 26.7% patients (32/120) using the mRNA approach. Although this is relatively high, this could be due to several factors. mRNA is only stable for a year at -70°C (as discussed in section 4.3.2.1.2)

and some samples took longer than a year to try to identify the disease-causing mutation, and therefore made it difficult to complete analysis in these samples. This should be kept in mind for diagnostic testing and it may not be possible to batch samples for testing. Other factors could involve poor mRNA quality of the initial sample, a deletion of the entire *FVIII* transcript or even large or partial gene deletions. In cases where amplification of one or more fragments failed, it would be difficult to determine if this was due to a full or partial gene deletion, large rearrangements or technical difficulties with the amplification of the cDNA fragment/s. A possible resolution to this limitation in some patients could be to test the patient's DNA, especially in the region were mRNA amplification was unsuccessful. Rearrangements are however, difficult to determine and identify. If the PCR runs are carefully controlled, deletions of the entire fragment should be detectable or at least suspected on repeat runs.

For those patients in whom no mutation was detected, none had amplification of all 9 fragments and were found to be negative on sequencing analysis for these fragments. These patients were therefore unsuccessful for technical reasons. The mutation detection rate is likely to increase if the technical issues are resolved or at least minimised.

Other options to reduce the failure rate of mRNA analysis could be to improve the primer and PCR conditions of the reactions, and perhaps to even optimise the systems so that all the fragments can be amplified under the same conditions or even adding universal sequencing tags to allow for sequencing using a single primer.

4.4 IMPLEMENTATION OF FINDINGS INTO THE DIAGNOSTIC LABORATORY

This project shows that inclusion of mRNA analysis for the *FVIII* gene detected a further 55.7% (49/88) of mutations. Although certain limitations related to mRNA analysis (stability, transportation of samples, and the failure rate of analysis) need to be overcome, mRNA analysis offers a significantly improved method for analysis, and is likely to result in an even higher detection rate in a controlled

diagnostic setting. A step-wise approach to mutation detection may be the most cost effective option, however, it may be more time consuming.

Currently, the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg, only offers direct mutation detection for the intron 22 inversion mutation, which detects mutations in 40.9% (36/88) of patients. However, the frequency of the intron 22 inversion mutation varies across the different population groups: 32.6% (14/43) of the white patients, 50% (19/38) of the black patients and 42.9% (3/7) of patients from the Indian and mixed ancestry patients (refer to section 3.2.1.1). Currently, the intron 22 inversion mutation is detected using Southern blot analysis, which is costly, time consuming and requires the use of hazardous radioactive labelling. It is highly recommended that long range PCR for the intron 22 inversion mutation (Bagnall, *et al.*, 2005) be optimised and implemented into the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg.

Secondly, as the c.3637insA mutation accounts for 14% of mutations in white patients (refer to section 3.2.5 and 4.2.7), it is recommended that testing for this mutation be included into routine diagnostic testing, prior to mRNA analysis. Together, the intron 22 inversion mutation and the c.3637insA mutation account for 46.5% (20/43) of disease-causing mutations in the white patients. This c.3637insA mutation is detected by sequencing analysis of exon 14b1. This sequencing would also detect other mutations in this part of exon 14.

The third recommendation would be to include testing for the intron 1 inversion mutation in patients who test negative for the intron 22 inversion mutation. The intron 1 inversion mutation accounts for 3.4% (3/88) of disease-causing mutations in the South African haemophilia A population. Although this mutation has not been detected in white South African patients, it accounts for 4.8% (10/209) of severe haemophilia A mutations in the United Kingdom (Bagnall, *et al.*, 2002) and 15% (3/21) of mutations in severe haemophilia A patients in the Netherlands (Boekhorst, *et al.*, 2005). As detection is via a single PCR, it is advisable to test for this mutation before mRNA analysis.

Screening of exon 14 of the *FVIII* gene should be considered for patients who test negative for the intron 22 inversion and intron 1 inversion mutations, and for white patients who test negative for the c.3637insA mutation. Excluding the c.3637insA founder mutation, 36.7% (18/49) of non-inversion mutations identified in this project were found to occur in exon 14. This result is not surprising as approximately one third of the *FVIII* transcript is comprised of exon 14. Exon 14 can easily be tested in four amplified reactions using DNA, and three reactions if mRNA is available (fragment 3, 4a and 4b). It is recommended to sequence exon 14 at the DNA level, as it would include the detection of obvious splice site mutations. Fragments 3, 4a and 4b are all within exon 14 and do not cover the exon/intron boundaries. These exon/intron boundaries are covered in fragment 2 and 5.

If the disease-causing mutation has still not been identified in a patient using this approach, mutation detection using the mRNA approach is likely to detect the disease-causing mutation. There are no recommendations for the order of screening of the remaining mRNA fragments.

Further options for mutation detection could involve assessing the use of MLPA. In this project, this technique may only have aided in the detection of mutation in two patients, namely HPA202 with a deletion of exon 10-11 and HPA017 with a deletion of exon 19. If MLPA can be performed using mRNA, it may detect mutations which affect splicing. Such mutations (exon deletions, splice site mutations and complex mutations) were responsible for disease in 16.3% (8/49) of non-inversion patients in this project.

A summary of the recommendations for the step-wise screening of the *FVIII* gene in haemophilia A in a diagnostic setting is outlined in Figure 4.1.

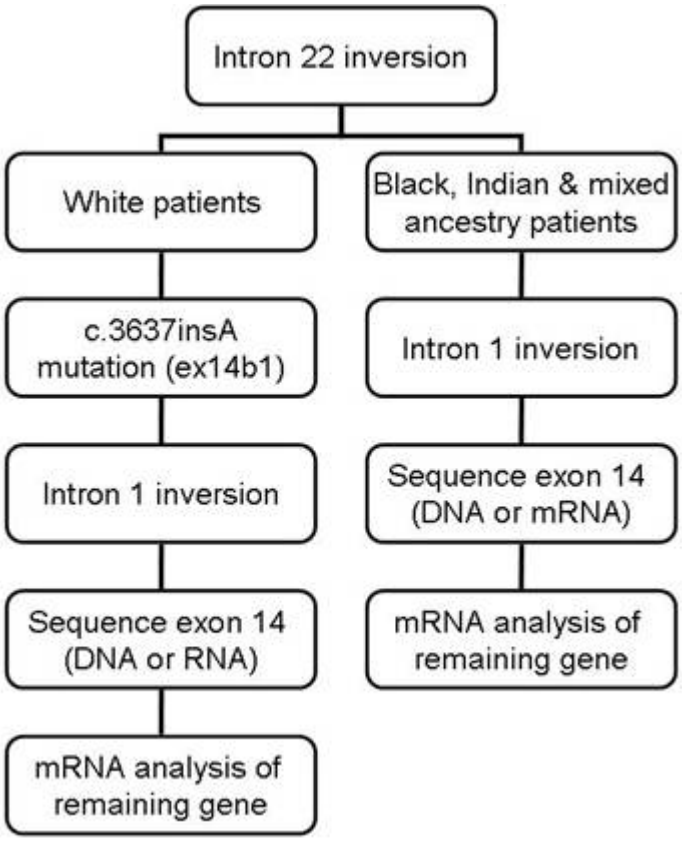

Figure 4.1: A schematic summary of the approach for step-wise mutation detection for haemophilia A patients in a diagnostic setting.

4.5 *FVIII* ANALYSIS IN OTHER CENTRES

Details of various laboratories from around the world, who offer mutation analysis for haemophilia A, were obtained from the Laboratory Directory on the GeneTests website (http://www.genetests.org). Five laboratories, who offer diagnostic testing for haemophilia A, were contacted via e-mail to enquire about their mutation detection methodologies. All five laboratories replied. A comparison of the techniques utilised for diagnostic *FVIII* analysis found that mRNA analysis was not routinely used in a diagnostic setting. mRNA is therefore a fairly underutilised tool for mutation detection for haemophilia A and possibly other conditions. It may be because of the technical difficulties entailed with handling mRNA (as already discussed in section 4.3.2) or the limited timescale and budget to undertake routine diagnostic testing (Bell, Bodmer, Sistermans, *et al.*, 2007). These

laboratories may have bigger budgets and therefore the cost of sequencing DNA would not be a problem.

The University Medical Centre, Utrecht, Netherlands performs diagnostic analysis on the genomic DNA using a PCR test for the inversion intron 22 and sequence analysis of the exons (including exonic splice regions) of the *FVIII* gene (Dr M.E. van Gijn, 2008, personal e-mail).

Leiden University Medical Centre in Netherlands use DNA, not mRNA for haemophilia A analysis. They analyse the complete coding sequence by sequence analysis and use PCR analysis for detection of the intron 1 and intron 22 inversions. They also can perform MLPA analysis for haemophilia A (Dr E.M.J. Boon, 2008, personal e-mail).

The Department of Genetics at the University of Pennsylvania, Philadelphia use a step-wise mutation detection method. They start by testing for the intron 22 inversion mutation using Southern Blot analysis, followed by full sequencing of all coding exons of the *FVIII* gene. If no mutations are found using these methods, intron 1 inversion testing, quantitative PCR and mRNA analysis to search for intronic mutations follow (Dr Ganguly, 2008, personal e-mail)

The Sheffield Molecular Genetics Service in the United Kingdom uses long-range-PCR for the intron 22 and intron 1 inversion mutations and direct DNA sequencing of all *FVIII* exons (Dr A. Goodeve, 2008, personal e-mail).

The most comprehensive mutation analysis strategy appears to be performed at the City of Hope Medical Centre in California. Initially they perform a PCR assay for the intron 22 inversion, followed by comprehensive mutation analysis where all coding regions and splice junctions of the *FVIII* gene are analysed using direct DNA sequencing. This is followed by testing for the intron 1 inversion mutation also using a PCR assay. Patients who test negative for these testing strategies are further tested for von Willebrand factor Normandy mutations (sequence analysis of exon 18, 19 and 20 of the von Willebrand gene) as mutations here may impact on FVIII stability. And lastly, in patients where no mutations are found, they perform a dosage analysis for known large gene deletions (their assay tests for large

deletions in exon 1, 6, 7, 8, 10, 12-14, 16, 20-23, 25-26 of the *FVIII* gene) (Weiss, 2008).

Therefore the use of mRNA analysis for *FVIII* in this study is a novel approach and its benefits have been proved in this project as it has the potential to reduce costs and detect a broad range of mutations, including large deletions and rearrangements which may be missed by DNA sequencing.

4.6 SUMMARY OF PROJECT FINDINGS

mRNA is a useful method for mutation analysis, especially for haemophilia A, and it is recommended that the mutation detection strategy outlined in section 4.4 be implemented into the routine haemophilia A diagnostic service available at the Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg.

The advantages of mRNA include the reduced cost of mutation detection compared to DNA analysis (section 4.3.1.1) and detection of additional mutations that may have been missed using DNA analysis (section 4.3.1.2). The limitations of mRNA analysis include the stability of mRNA (section 4.3.2.1) and the failure rate of mRNA analysis (section 4.3.2.2).

5 CONCLUSION

Knowledge of the haemophilia A disease-causing mutation may have clinical significance in determining the severity of the disease and in prediction of which patients are prone to developing inhibitors to FVIII replacement treatment or are potential candidates for gene therapy trials. Direct mutation detection is also useful in accurately determining the carrier status of at risk family members and in prenatal diagnosis.

In this study, the intron 22 inversion mutation was found in 40.9% of South African patients (32.6% of white patients, 50% of black patients and 42.9% of the Indian and mixed ancestry patients). The inversion mutation in intron 1 of the *FVIII* gene accounts for 3.4% of all patients (0% of white patients, 5.3% of black patients and 12.5% of patients from the Indian and mixed ancestry population). Analysis of the *FVIII* transcript identified an additional 55.7% of mutations (67.4% of white patients, 44.7% of black patients and 42.8% of Indian and mixed ancestry patients). The mutation profile of non-inversion patients was as follows: 34.7% missense, 18.4% nonsense, 4.1% splice-site mutations, 18.4% deletions, 14.3% insertions, while the remaining 10.2% patients had a complex mutation.

In addition to the intron 22 and intron 1 inversion mutations, mRNA proves to be highly effective in identifying disease-causing mutations. Using this testing strategy, a mutation was identified in 73.3% of South African Haemophilia A patients (76.8% of white patients, 67.9% black patients and 87.5% Indian and mixed ancestry patients). A higher mutation detection rate is expected if the technical difficulties related to mRNA are minimised.

Analysis of mRNA identified the c.3637insA mutation at a frequency of 14% in the white haemophilia A population, thereby providing further evidence for the presence of the previously suspected founder mutation in these patients.

mRNA analysis is a cost effective method of mutation analysis (almost half the cost of DNA based analysis) and has a high mutation detection rate (able to detect a wide variety of mutations, including rearrangements, deletions, insertions,

mutations affecting splicing and point mutations). mRNA based analysis is also able to determine the effects of certain mutations simultaneously.

Unfortunately, mRNA is less stable then DNA which may impact on the transportation of samples from remote locations and long term storage of diagnostic samples to be used as a positive control sample in future.

The Diagnostic Molecular Genetics Laboratory, Division of Human Genetics, NHLS and School of Pathology, the University of the Witwatersrand, Johannesburg is the only laboratory in South Africa to offer mutation analysis for haemophilia A, thereby providing a diagnostic service to the whole country and the whole Southern African region. From the findings of this study, it is recommended that this laboratory implement a step-wise based approach to mutation detection for haemophilia A. Initially, testing for the intron 22 inversion mutations (by Southern blotting or ideally by long range PCR) is recommended, followed by the detection of the intron 1 inversion mutation and the c.3637insA mutation (in white patients). Thereafter, mutation analysis of exon 14 (either by DNA or mRNA analysis) would be recommended before full mutation analysis using an mRNA based approach is carried out.

In summary, testing for a mutation using intron 1 and 22 direct detection and mRNA analysis identified a mutation in 76.8% (43/56) white patients, 67.9% (38/56) black patients and 87.5% (7/8) Indian and mixed ancestry patients. Thus a new mutation detection strategy has been developed that significantly improves the diagnostic service available to haemophilia A patients and their families in South Africa.

6 APPENDICES

APPENDIX A

ETHICS

A1 ETHICS CLEARANCE CERTIFICATE

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Krause/Hetem

CLEARANCE CERTIFICATE	**PROTOCOL NUMBER M040801**
PROJECT	Mutation Detection for Haemophilia A in South Africa
INVESTIGATORS	A/CH Prof/Ms Krause/Hetem
DEPARTMENT	Human Genetics
DATE CONSIDERED	04.08.27
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 04.10.18 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable

cc: Supervisor : Prof A Krasue

DECLARATION OF INVESTIGATOR(S)

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10005, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

A2 PATIENT CONSENT FORM

The National Health Laboratory Service

University of the Witwatersrand, School Of Pathology
Division Human Genetics

Hospital Street, Johannesburg 2001 PO Box 1038, Johannesburg 2000
Telephone: +27-11-489-9224/9223/9211 Telefax: +27-11-489-9226 or +27-11-489-9209
Prof A Christianson 489-9211/9239 Prof A Krause 489-9219 Prof M Ramsay 489-9214
Prof H Soodyall 489-9208 Dr T Lane 489-9221

HAEMOPHILIA A: MOLECULAR STUDY

PATIENT INFORMATION SHEET AND CONSENT FORM

Hello. My name is Claire Mitchell and I work in the Department of Human Genetics at the National Health Laboratory service, and try to identify what causes some diseases. I am doing a project on Haemophilia A and want to find what causes your Haemophilia.

Haemophilia A is a disorder that occurs when a person has a fault in their genetic make-up. Affected people, usually males, suffer from uncontrolled bleeding episodes, even after minor injury. Haemophilia is a result of a lower amount of the blood-clotting factor, factor VIII. Patients with the severe form of the disease have less than 1% of the normal level of factor VIII in their blood. Haemophilia is an X-linked genetic condition, which means that the fault (or mutation) in the gene is passed from the mother to her children, but only her sons may be affected.

There are many different mutations in the factor VIII gene causing Haemophilia. There is currently one common mutation, but only 40% of patients have this mutation. In this study, we want to find other mutations that are common in the South African population. To do this we need to study the blood of Haemophilia patients to look for the mutations causing the disease in these patients. If we find a common mutation, a more accurate diagnostic test can be offered to all Haemophilia patients. But, if we find any disease-causing mutation in you or your family, we would be able to test your female relatives to see if they carry the faulty gene. If they have this mutation, they are at risk of passing this gene on to their sons who will then be affected with Haemophilia.

We would like to ask you to participate in the study by allowing us to take **one-two teaspoons of blood** (5-10ml) from you / your child.

1. The research has been explained to me and I understand the test involves taking a sample of blood from my child / myself (which may cause slight discomfort).
2. I understand that the study may have direct benefits for my child or me. If not, it will help researchers to understand the disorder better so that they can develop treatments and ways to prevent it from happening.
3. I understand that I do not have to take part in this project and that if I choose not to take part this will not affect the way in which my child or I will be treated at the hospital / clinic. Similarly, if I chose to withdraw from the study at any stage, this will not prejudice any future treatment my child or I may require.

NAME: ..... DATE: .....

SIGNATURE:.....

A3 PATIENT ASSENT FORM

The National Health Laboratory Service

University of the Witwatersrand, School Of Pathology
Division Human Genetics

Hospital Street, Johannesburg 2001 PO Box 1038, Johannesburg 2000
Telephone: +27-11-489-9224/9223/9211 Telefax: +27-11-489-9226 or +27-11-489-9209
Prof A Christianson 489-9211/9239 Prof A Krause 489-9219 Prof M Ramsay 489-9214
Prof H Soodyall 489-9208 Dr T Lane 489-9221

HAEMOPHILIA A: MOLECULAR STUDY

ASSENT FORM

Hello,

My name is Claire Mitchell. I work in the Department of Human Genetics, and try to identify what causes some diseases. I am doing a project on Haemophilia and want to find what causes your Haemophilia.

To do this, we need to look at your hospital files, talk with your parents and take a small sample of your blood (about 10-15ml, which is the same as 3 teaspoons) from your arm. We can try to do this at the same time we take your other blood tests, if you need them. If not, we will just take blood for this project. As you know, when we take blood, it is a quick prick, which is uncomfortable as the needle goes through the skin, but does not cause any harm.

All the information in your files is private (no one else will be allowed to see them).

This study will may not help you directly – it will not make you better. Your treatment will continue just as before. This project will help us to understand what causes Haemophilia, and may help your family and other families in the future.

Please let us know if you agree to take part in this study.

APPENDIX B

TECHNIQUE PROTOCOLS

B1 DNA EXTRACTION PROTOCOL

Salting-out Procedure

Modified from Miller, *et al.* (1988).

1. Collect blood into purple-top EDTA or yellow-top ACD tubes
2. Decant no more than 10ml whole blood into a 50ml polypropanol (NUNC) tubes. If only 3-5ml blood is available, halve the volumes of solutions in the protocol
3. Freeze the blood at -20°C for storage. Thaw the whole blood for about 1 hour before extraction
4. The composition of all solutions used can be found in Appendix D

Day 1:

1. Fill each NUNC tube (up to 45-50ml mark) with cold sucrose-Triton-X lysing buffer (keep this solution cold during the procedure), and invert the NUNC tube several times to mix
2. Centrifuge the NUNC tube for 10 minutes at 1000x*g* (~2400r.p.m.) at 4°C in the Beckman GS-6R centrifuge
3. A reddish-white pellet should be visible. Pour off the supernatant fluid containing the lysed red blood cells, ensuring the pellet does not dislodge
4. Wash the pellet with 20-25ml cold sucrose-Triton-X lysing buffer and place the NUNC tube on ice for 5 minutes
5. Centrifuge the NUNC tube for 5 minutes at 1000x*g* (~2400r.p.m.) at 4°C in the Beckman GS-6R centrifuge
6. Pour off the supernatant. These washes may be carried out several times until a white pellet is visible
7. Add 3ml T20E5, 0.2ml (200μl) 10% SDS and 0.5ml (500μl) Proteinase-K mix. Mix the NUNC tubes well by inversion
8. Incubate at 42°C to 50°C overnight (no need to agitate)

Day 2:

9. Add 1ml saturated NaCl to the lysate the following day. Agitate vigorously for 15 seconds by inversion
10. Chill the NUNC tube with the salt at -20°C or on ice for 5-10 minutes
11. Centrifuge the NUNC tube for 30 minutes at 1000x*g* (~2400r.p.m.) at 4°C in the Beckman GS-6R centrifuge
12. A white pellet should be visible. If no pellet is visible, spin again for 10-20 minutes
13. Transfer the supernatant containing the DNA into a new NUNC tube
14. Add two volumes absolute ethanol (Merck) kept at room temperature
15. Agitate gently and spool the DNA. If no DNA is visible, precipitate DNA (see below)
16. Wash DNA in 70% ice-cold ethanol. Transfer the DNA into a labelled Eppendorf tube

17. Air-dry the DNA and resuspend in an appropriate amount of 1xTris-EDTA (TE) buffer (usually between 200μl and 1000μl) overnight

To Precipitate DNA

18. If DNA is unable to be spooled, place NUNC tube containing the absolute ethanol at -20°C overnight or at -70°C for 30 minutes
19. Centrifuge for 30 minutes (maximum) at 1000x*g* (~2400r.p.m.) at 4°C in the Beckman GS-6R centrifuge
20. Pour off the supernatant. Add about 15ml 70% ethanol. Centrifuge at 1000x*g* (~2400r.p.m.) at 4°C for 15 minutes
21. Pour off supernatant. DNA is the precipitate at the bottom of the NUNC tube. Allow the pellet to air-dry. Resuspend in appropriate amount of 1xTE buffer (between 50 and 250μl) overnight
22. Transfer the DNA into a labelled Eppendorf tube. Label the tube "ppt" to indicate that it is a precipitate

B2 RNA EXTRACTION PROTOCOL

QIAamp RNA Blood Mini Kit (QIAGEN)

Notes before starting

The RPE, RLT, EL and RW1 buffers, RNase free ddH_2O and the QIAshredder and QIAamp spin columns are supplied with the kit

Add 4 volumes absolute ethanol (Merck) to buffer RPE

Add 10µl β-mercaptoethanol (Sigma) to 1ml buffer RLT. This solution is stable at room temperature for 1 month

Cell lysates in buffer RLT (from step 6) can be stored at -70°C if necessary. Thaw products at 37°C for 10 minutes before continuing with protocol (step 7)

Steps 3-5 use the Beckman GS-6R centrifuge at 400x*g* (~1657r.p.m.) at 4°C. Steps 7-15 use the Eppendorf centrifuge 5415C, however the centrifugation speeds differ between the different steps. The vortex steps use the Vortex-Genie 2 (Scientific Industries)

Protocol

1. Mix 1 volume (0.5-1ml) fresh blood with 5 volumes (2.5-5ml) buffer EL. Incubate for 10-20 minutes on ice. Mix by vortexing twice. The cloudy suspension becomes clear
2. Centrifuge in the Beckman centrifuge at 400x*g* (~1657r.p.m.) for 10 minutes at 4°C. Discard supernatant
3. Add 2 volumes (2ml) buffer EL to the cell pellet. Resuspend pellet by vortexing. Centrifuge in the Beckman centrifuge at 400x*g* (~1657r.p.m.) for 10 minutes at 4°C. Discard supernatant
4. Add 600µl buffer RLT. Vortex or pipette to mix and remove all clumps

Cell lysate products can be stored at -70°c if needed. Thaw and incubate at 37°C for 10 minutes before continuing with step 5

5. Pipette lysate onto QIAshredder Spin Column (lilac column) in a 2ml collection column. Centrifuge in the Eppendorf centrifuge at maximum speed for 2 minutes to homogenise. Discard the QIAshredder and keep homogenised lysate
6. Add 1 volume (600µl) 70% ethanol to homogenised lysate. Mix by pipetting. Pipette sample onto QiAamp spin column (clear column) in collection tube. Centrifuge at ≥8000x*g* (≥10000r.p.m.) for 15 seconds in the Eppendorf centrifuge
7. Transfer QiAamp spin column into a new collection tube. Add 700µl buffer RW1 to the QiAamp spin column. Centrifuge at ≥8000x*g* (≥10000r.p.m.) for 15 seconds in the Eppendorf centrifuge
8. Transfer QiAamp spin column into a new collection tube. Check ethanol was added to buffer RPE. Apply 500µl buffer RPE to the QiAamp spin column. Centrifuge at ≥8000x*g* (≥10000r.p.m.) for 15 seconds in the Eppendorf centrifuge
9. Add 500µl buffer RPE to the QiAamp spin column. Centrifuge at ≥20000x*g* (≥14000r.p.m.) for 3 minutes in the Eppendorf centrifuge. Discard the flow through
10. Centrifuge again at ≥20000x*g* (≥14000r.p.m.) for 1 minute in the Eppendorf centrifuge
11. Transfer QiAamp spin column to a clean microfuge tube. Pipette 30-50µl RNase free ddH_2O onto the QiAamp column membrane. Centrifuge at ≥8000x*g*

(≥10000r.p.m.) for 1 minute in the Eppendorf centrifuge to elute the RNA. Repeat if >0.5ml blood was used

12. Store the RNA at -20°C or -70°C. Samples are stable for a year at these temperatures

TRIzol RNA Stabilisation:

The protocol for stabilisation of RNA in blood using TRIzol was received from Dr Marius Coetzee (Department of Haematology and Cell Biology, University of the Free State, Bloemfontein, 2009). This was to determine if TRIzol stabilised samples could be used for analysis in this project and considering future diagnostic use.

1. Decant 20ml fresh whole blood into a 50ml centrifuge tube
2. Add 30ml working lysis buffer and stand or mix for 10 minutes
3. Centrifuge for 10 minutes at 3000x*g* to pellet the white cells
4. Discard lysed red blood cells carefully without disturbing the white cell pellet
5. Add 20ml working lysis buffer to the tube and resuspend pellet
6. Stand or mix for 5 minutes and centrifuge for 10 minutes at 3000x*g*
7. Discard supernatant carefully, without disturbing the white cell pellet
8. Add 3.2ml TRIzol solution and mix by pipette action and ensure white cells are completely dissolved
9. Store the TRIzol homogenate at -70°C

TRIzol RNA extraction:

1. Thaw frozen sample. Transfer the sample to a 50ml NUNC tube
2. Add 0.2ml of chloroform:isoamylalcohol and shake vigorously for 15 seconds. Incubate at room temp for 3 minutes
3. Centrifuge at 12000x*g* (~14000r.p.m.) in the Eppendorf centrifuge 5415C for 15 minutes at 4°C
4. The aqueous phase (clear layer) was removed and placed in a new tube
5. Add 0.5 ml of isopropanol (propan-2-ol) (Sigma) to the aqueous phase and incubate at room temp for 10 minutes
6. Centrifuge at 12000x*g* (~14000r.p.m.) in the Eppendorf centrifuge 5415C for 10 minutes at 4°C. The RNA forms a pellet that may not be visible
7. Add 1ml 100% ethanol (Merck) and centrifuge at 7500x*g* (~10000r.p.m.) in the Eppendorf centrifuge 5415C for 5 minutes at room temperature
8. Remove the supernatant and air-dry the sample for 15 minutes
9. Add 50µl of RNase-free ddH_2O to resuspend the sample
10. Store the RNA at -20°C or -70°C

B3 RADIOACTIVE LABELLING OF P482.6 PROBE PROTOCOL

MegaPrime DNA labelling Kit (Amersham)

Protocol

1. Mix together: 10µl p482.6 probe
   5µl primer (supplied with the kit)
   11µl dH_2O
2. Incubate in a Hagar heating block at 95°C for 5 minutes, then cool on ice
3. Add: 4µl each dATP, dGTP, dTTP (supplied with the kit)
   5µl buffer (supplied with the kit)
   5µl 10mCi/µl ^{32}P dCTP
   2µl enzyme (supplied with the kit)
4. Incubate in a Hagar heating block at 37°C for 30 minutes
5. During the incubation step, prepare the spin columns as follows:
   a. Plug the bottom of a 1ml syringe with glass wool, up to the 0.1ml mark. Fill the syringe with 5% Sephadex (Appendix D) saturated with 1x TE buffer (Appendix D)
   b. Place the syringe in a 15ml falcon tube and spin in an IEC Clinical Centrifuge (International Equipment Company) at maximum speed for 5 minutes to compact the Sephadex beads. Pour the liquid off at the bottom of the falcon tube
   c. Top up the syringe with more Sephadex and spin again for 5 minutes. Pour off the liquid at the bottom of the falcon tube
   d. Repeat the top-up and spin until the column is packed to approximately 0.9ml
   e. Place a clean Eppendorf at the bottom of the falcon tube to make the collection of the probe easier
6. Add 5µl of 0.2M EDTA (Appendix D) to stop the labelling reaction
7. Add 50µl 1x TE buffer (pre-warmed to 37°C)
8. Add 100µl of the labelled probe mix to the column (see preparation below)
9. Spin at max speed for 5 minutes and collect eluted products into clean Eppendorf tube
10. Add 100µl 1x TE buffer and spin at max speed for 5 minutes. Collect eluted products into same tube as first spin
11. Dilute these products to 1ml with 1x TE buffer
12. Diluted probe can be used immediately or stored at -20°C in lead peg until required
13. To use labelled probe, denature in a Hagar heating block at 95°C for 5 minutes, then ice for 5 minutes

B4 REVERSE TRANSCRIPTION PROTOCOL

ImPromII Reverse Transcription System (Promega)

Notes before starting

The oligo$(dT)_{15}$ primer, nuclease-free ddH_2O, ImPromII buffer, $MgCl_2$, dNTP mix, RNaseIN and ImPromII Reverse Transcriptase enzyme are supplied with the kit.

Thaw all reagents on ice

Perform all steps on ice

Incubation steps are performed in a heating block or a PCR machine without a heated lid

Protocol

RNA/primer mix

Mix 4µl RNA with 1µl of 0.5µg/µl oligo$(dT)_{15}$ primer, and spin samples briefly

Incubate samples at 70°C for 5 minutes

Place samples on ice for 5 minutes

Spin for 10 seconds

Reverse transcriptase mix

Combine:	3.7µ	Nuclease-free ddH_2O
	4µl	5x ImPromII buffer
	4.8µl	$MgCl_2$
	1µl	dNTP mix
	0.5µl	RNaseIN (40U/µl) (optional)
	1 µl	ImPromII Reverse Transcriptase enzyme

Add 15µl reverse transcriptase mix to 5µl RNA/primer mix

Anneal at 25°C for 5 minutes

Extend at 42°C for 1 ½ hours

Inactivate the reverse transcriptase at 70°C for 15 minutes

Store cDNA products at 4°C until required

B5 PCR PRODUCT CLEAN-UP PROTOCOL

NucleoSpin Extract II Kit (Macherey-Nagel)

Notes before starting

Add absolute ethanol (Merck) to buffer NT3

All centrifugation steps performed at 11000x*g* in either the Eppendorf centrifuge 5415C (~11000r.p.m.) or the Beckman-Coulter Microfuge®18 centrifuge (~12500r.p.m.) for the specified time

All incubation steps performed in the Bloer heating block

Protocol

1. For extraction from agarose gel:
    a. Excise the product of interest from the agarose using a clean scalpel and placed in a clean Eppendorf tube
    b. Add two volumes buffer NT to one volume agarose (approximately equal to the weight of the gel, in mg)
    c. Incubate at 50°C for 5-10 minutes until the agarose has completely dissolved

OR For extraction directly from PCR products

    a. Make the volume of PCR products up to 50µl
    b. Add 100µl buffer NT

2. Label the NucleoSpin Extract II column and place into a collection tube
3. Add the sample to the NucleoSpin Extract II column. Centrifuge for 1 minute. Discard flow through
4. Add 600µl buffer NT3 to the NucleoSpin Extract II column. Centrifuge for 1 minute. Discard flow through
5. Centrifuge for 2 minutes. Discard collection column
6. Incubate the spin column at 70°C for 2-5 minutes to remove all residual ethanol from the sample
7. Place the NucleoSpin Extract II column into a clean, labelled 1.5ml Eppendorf tube
8. Add 15-50µl ddH_2O to the spin column membrane. Incubate at room temperature for 1 minute. Centrifuge for 1 minute to elute he cleaned sample

B6 CLEAN-UP OF CYCLE SEQUENCING PRODUCTS PROTOCOL

DyeEx™ 2.0 Spin Kit (QIAGEN)

Notes before starting

All centrifugation steps performed at 750x*g* using either the Eppendorf centrifuge 5415C (~3000r.p.m.) or the Beckman-Coulter Microfuge®18 centrifuge (~2500r.p.m.).

Protocol

1. Vortex sephadex column to remove all air bubbles
2. Loosen screw cap and break off the bottom of the tube
3. Place sephadex column into a collection tube
4. Centrifuge the column for 3 minutes
5. Discard the collection tube
6. Add 20µl of cycle sequencing product onto the centre of the sephadex
7. Place sephadex column into a clean, labelled 1.5ml Eppendorf tube
8. Centrifuge for 3 minutes
9. Discard sephadex column
10. Collect the flow-through and dry in the vacuum drier to concentrate the products

BigDye®X-Terminator™ Purification Kit (Applied Biosystems)

Notes:

This clean up method is used for cleaning 20µl cycle sequencing products in a 96-well plate

Halve the volumes of SAM™ Solution and BigDye® X-Terminator™ Solution if cleaning 10µl cycle sequencing products

SAM™ Solution and BigDye® X-Terminator™ Solution can be premixed in a 4.5:1 ratio up to 24 hours in advance

1. Add 90µl of SAM™ Solution and 20µl of BigDye® X-Terminator™ Solution to each well containing cycle sequencing product
2. Seal the plate with heat seal film or AB adhesive cover
3. Thoroughly mix the contents of the plate on a Vortex-Genie 2 (scientific Industries) for 30 minutes
4. Centrifuge the plate at 1000*xg* for 2 minutes
5. Run the plate on a 3130*xl* Genetic Analyser using the BigDye® X-Terminator™ run module (refer to section 2.3.8)

B7 PROTEIN SYNTHESIS PROTOCOL

EasyXpress™ Protein Synthesis Mini Kit (Promega)

and

EasyXpress™ Linear Template Kit (Promega)

Notes before starting

Once thawed, the *E. coli* extract should be stored on ice and used within 4 hours

The EasyXpress Protein Synthesis kits are very sensitive to multiple freeze-thaw cycles. Do not refreeze and thaw more than two times

Except for the actual transcription-translation incubation, all handling steps should be carried out on ice

Protocol

1. Thaw EasyXpress Reaction Buffer (blue screw-cap) and XE-solution (green screw-cap) on ice. Thaw RNase-free water (colourless screw-cap) at room temperature
2. For each sample (including a positive- and no-template control reactions), prepare a premix containing 20µl EasyXpress Reaction buffer and 2.5µl RNase-free water in an RNase- and DNase-free reaction tube
3. Thaw one tube *E. coli* extract (colourless snap-cap) on ice
4. Add 22.5µl premix from step 2 to each tube of *E. coli* extract
5. For each reaction add 2µl XE-solution to 0.7µg (~7µl) from the second PCR in a separate tube and mix by pipetting up and down
6. Add the XE-solution-PCR product mix from step 5 to each reaction
7. Bring the final reaction volume to 50µl with RNase-free water
8. Mix and centrifuge briefly to collect reactions at the bottom of the tube
9. Incubate at 37°C for 1 hour
10. Stop the reaction by placing on ice and proceed with sample analysis

APPENDIX C

PRIMER SEQUENCES

Table C1: The sequences of the primers used for amplifying the intron 1 inversion mutation of the *FVIII* gene.		
Primer name	**Flanks**	**Sequence (5'-3')**
9F	Int1h-1	GTTGTTGGGAATGGTTACGG
9cR	Int1h-1	CTAGCTTGAGCTCCCTGTGG
Int1h-2F	Int1h-2	GGCAGGGATCTTGTTGGTAAA

Table C2: The sequences of the primers used for amplifying the overlapping cDNA fragments of the *FVIII* gene

Fragment	Primer used for	Primer name	Sequence (5'-3')
1a	PCR and sequencing	HPA-RNA-1a-F	GGGAGCTAAAGATATTTTAGAGAAG
		HPA-seq1-R	TTTGTTTCTGAGTGCCAACTTTT
1b	1st round PCR	HPA-RNA-1a-F	GGGAGCTAAAGATATTTTAGAGAAG
		HPA-RNA-1b-R	CAACAGTGTGTCTCCAACTTCCCCAT
	2nd round PCR and sequencing	HPA-seq1-F	GGCCTCATTGGAGCCCTACT
		HPA-PTT-1d-R	ttttttttttttttttttttttttttttttatcCCTACCAATCCGCTGAGGGCCATTG
2	1st round PCR	HPA-RNA-2a-F	GAAGAAGCGGAAGACTATGATGATG
		HPA-RNA-2b-R	GCCTAGTGCTAGGGTGTCTTGAATTC
	2nd round PCR and sequencing	HPA-PTT-2c-Fnew	ctaatacgactcactataggaacagaccaccatgGATTCTGAAATGGATGTGGTCAGG
		HPA-RNA-2d-R	GGGAGAAGCTTCTTGGTTCAATGGC
	Internal sequencing	HPA-PTT2-seq-F	GATCCTCGGTGCCTGACCCGC
		HPA-PTT2-seq-R	CTCTTGTCTGACATTATCTG
3	PCR and sequencing	HPA-RNA-3c-F	ATGACCGCCTTACTGAAGGT
		HPA-PTT-3f-R-new	ttttttttttttttttttttttttttttttcaTAGGTTTCTGCTGCTTGGAA
	Internal sequencing	HPA-PTT3-seq-F	AAGTCATCTCCCCTTACTGA
		HPA-PTT3-seq-R	CCCCATGAACTTTCTTGGCT

Table C2: Continued.

Fragment	Primer used for	Primer name	Sequence (5'-3')
4a	PCR and sequencing	HPA-RNA-4c-F	TTTCCAAGCAGCAGAAACCTAT
		HPA-PTT4-seq-R	AATGACTGCTTTCTTGGACC
4b	PCR and sequencing	HPA-PTT4-seq-F	AGGGTCCTATTCCAAGACAACT
		HPA-RNA-4f-R-new	TAGAACATGTGGGGAGCTA
5	1st round PCR	HPA-RNA-5a-F	CTTCAGTCAGATCAAGAGGAAATTGAC
		HPA-RNA-5b-R	GAAGTCTGGCCAGCTTTGGGGCCCACTGT
	2nd round PCR	HPA-RNA-5c-F	TATGATGATACCATATCAGTTGAAATG
		HPA-RNA-5d-R	CTCTAATGTGTCCAGAAGCCATTCCC
	Internal sequencing	HPA-PTT5-seq-F	TCAAGCCTAATGAAACCAAA
		HPA-PTT5-seq-R	CCTGAGTGCACATCTTTTTC
6a	PCR and sequencing	HPA-PTT-6-F	ctaatacgactcactataggaacagaccaccatgAAAAAAGAGGAGTATAAAAT
		HPA-PTT6-seq-R	ATTCCCAATGGCATGCTGCAA
6b	1st round PCR	HPA-RNA-6a-F	TTCATTTCAGTGGACATGTGTTCAC
		HPA-RNA-6b-R	CAGGAGGCTTCAAGGCAGTGTCTGC
	2nd round PCR	HPA-PTT6-seq-F	TAAAACACAATATTTTTAACC
		HPA-PTT-6d-R	ttttttttttttttttttttttttttttttttGGAGGGACACTGCCCTGGAGCTGAG

Table C3: The relative positions of the *FVIII* mRNA sequence and exons covered by each cDNA fragment.

Fragment	***FVIII* mRNA sequence covered**		**Exons covered**	
	From	**To**	**From**	**To**
1a	102bp	866bp	1	6
1b	739bp	1491bp	4	9
2	1284bp	2458bp	8	14
3	2332bp	3750bp	14	14
4a	3730bp	4528bp	14	14
4b	4444bp	5382bp	14	14
5	5191bp	6328bp	14	20
6a	6127bp	6767bp	18	24
6b	6632bp	7295bp	23	26

Table C4: The sequences of the primers used for amplifying the DNA exons of the *FVIII* gene.

Exon	Primer Name	Sequence (5'-3')	Annealing temperature	PCR product size	Number of cycles
1	HPA-exon1F	GGGAGCTAAAGATATTTTAGAGAA	61°C	325bp	30x
	HPA-exon1R	GGGGCCCAGGTAGCATCACAACCATCC			
2	HPA-exon2F	GTATGGAAGCATTACTTCCAGCTGC	66°C	297bp	30x
	HPA-exon2R	TGGCAGCTGCACTTTTTAACTGCAACC			
3	HPA-exon3F	GCATGCTTCTCCACTGTGACCTTG	61°C	261bp	25x
	HPA-exon3R	CATAGAATGACAGGACAATAGGAGGG			
4	HPA-exon4F	GTACAGTGGATATAGAAAGGAC	58°C	319bp	30x
	HPA-exon4R	GATTCAGTTGTTTGTACTTCTCTGC			
5	HPA-exon5F	TTACTGTCAAGTAACTGATG	49°C	325bp	35x
	HPA-exon5R	TCTCCTTCATTCCTGAACAG			
6	HPA-exon6F	TGAATCCTGATGCCTCAAGC	52°C	397bp	30x
	HPA-exon6R	GTACAGAACTCTGGTGCTG			
7	HPA-exon7F	CAGATTCTCTACTTCATAGCC	50°C	361bp	30x
	HPA-exon7R	GTAGGACTGGATATTTATAATATTC			
8	HPA-exon8F	CCTAGTCTCTGGTATAGAACAGCC	61°C	319bp	30x
	HPA-exon8R	TGCCATTTGATTCCATACCTG			
9	HPA-exon9F	CTGATTCTGAAATGGATGTGGTCAGG	61°C	325bp	25x
	HPA-exon9R	GATCATGTCCATTGGAGACAAGGC			
10	HPA-exon10F	AAGAGACTTGAGCATCACAG	61°C	397bp	30x
	HPA-exon10R	AGACTGGAGCTTGAGGTCCG			

Table C4: Continued.

Exon	Primer Name	Sequence (5'-3')	Annealing temperature	PCR product size	Number of cycles
11	HPA-exon11F	TCTAATTGAGCTATTTATGG	50°C	261bp	35x
	HPA-exon11R-new	ACATACACTGAGAATGAAAC			
12	HPA-exon12F	CATCATAGACTGCTAGCTCCTACC	66°C	319bp	30x
	HPA-exon12R	CACCACCCACTGGACTTAAGTGCTGC			
13	HPA-exon13F	CCTAATTGTATCATGACAATCACAATCC	61°C	325bp	25x
	HPA-exon13R	GACAGCATGTGAGCTAGTGGGC			
14a1	HPA-exon14a-F	GGTCACAAACAGGCATAGTACAACAG	61°C	1029bp	35x
	HPA-PTT3-seqR	CCCCATGAACTTTCTTGGCT			
14a2	HPA-PTT3-seq-F	AAGTCATCTCCCCTTACTGA	61°C	868bp	35x
	HPA-exon14a-R	TCTGAGGCAAAACTACATTCTCTTG			
14b1	HPA-exon14b-F	CAAAGGACGTAGGACTCAAAGAGATGG	61°C	828bp	35x
	HPA-PTT4-seqR	AATGACTGCTTTCTTGGACC			
14b2	HPA-PTT4-seqF	AGGGTCCTATTCCAAGACAACT	61°C	1070bp	35x
	HPA-exon14b-R	CACCAGAGTAAGAGTTTCAAGACACC			
15	HPA-exon15F	GAGGATGTGAGGCATTTCTACCCAC	61°C	319bp	25x
	HPA-exon15R	GTGGGAATACATTATAGTCAGCAAG			
16	HPA-exon16F	CCCTAAGGACCTTAAGATCCTAGAAG	60°C	325bp	30x
	HPA-exon16R	AAAGCTTCTTATTGCACGTAGG			
17	HPA-exon17F	CCACTCTGGTTCATAGGTGAGAGAGC	61°C	297bp	25x
	HPA-exon17R	TCTCATTTGTCAAAGTGCAATCTGC			

Table C4: Continued.

Exon	Primer Name	Sequence (5'-3')	Annealing temperature	PCR product size	Number of cycles
18	HPA-exon18F	GTGGGAGTGGAATCCTCATAGATGT	61°C	261bp	25x
	HPA-exon18R	GATTGTGTTCCCAGTGCCTAGACC			
19	HPA-exon19F	AACTGTAAGGGTCACGTAGC	61°C	319bp	30x
	HPA-exon19R	GAACTCTGATGATTTCTACCAGC			
20	HPA-exon20F	GTGCACTTCTAGTTACTGTGTTCCAC	61°C	325bp	25x
	HPA-exon20R	GGGAGGATATTTAAGGAGGCAGTGTC			
21	HPA-exon21F	CTCAAGTGTCTAGGACTAACCCAGC	61°C	297bp	25x
	HPA-exon21F	GATGAATGTGATACATTTCCCATC			
22	HPA-exon22F	CCTGTAGCAATGTAGATTCTTCC	60°C	261bp	30x
	HPA-exon22F	AAATGACTAATTACATACCATTAAGG			
23	HPA-exon23F	AGATGTTGGATGCTGTTGAG	60°C	316bp	30x
	HPA-exon23F	GTCACCCTACCCATGGTTGAGGG			
24	HPA-exon24F	GCTGCTCAGTATAACTGAGGCTGAAGC	60°C	252bp	30x
	HPA-exon24F	CTCTGAGTCAGTTAAACAGTAAATC			
25	HPA-exon25F	GAGGGATTTGGGAATTTCTGGGAG	60°C	353bp	30x
	HPA-exon25R	TTGCTCTGAAAATTTGGTCATA			
26	HPA-exon26F	GCAGTGACCATTGTCCTGTCAGAC	66°C	331bp	30x
	HPA-exon26R	GGCTTCAAGGCAGTGTCTGCTAGGATTTAGCAC			
Special intron 13	HPA20-int13-F	AAAATCTTTCCCAGACCAAT	65°C	382bp	30x
	HPA20-int13-R	AGCTGGAGGCATGACACTAC			

Table C5: The sequences of the primers used for amplifying the three microsatellite markers of the *FVIII* gene.

Microsatellite name	**Reference**	**Primer name**	**Sequence (5'-3')**
Intron 13 (intragenic)	Lalloz, *et al.* (1991)	Primer 1A	FAM-TGCATGATTGTACATATGTATCTT
		Primer 2A	CCAAATTACATATGAATAAGCC
Intron 22 (intragenic)	Lalloz, *et al.* (1994)	Primer 1B	HEX-TTCTTAGAATGTAGTGTGTG
		Primer 2B	TAA TGCCCACATTATAGA
p39 (extragenic)	Wehnert, *et al.* (1993)	p39F	HEX-AGCACATGGTATAATGAACCTCCACG
		p39R	CAGTGTGAGTAGCATGCTAGCATTTG

APPENDIX D

REAGENT PREPARATION

1.5% agarose

1.5g agarose (Hispanger)

100ml 1xTBE

Mix together, and boil until agarose is completely dissolved

Allow to cool to ~60°C

Add 3µl 10mg/ml Ethidium Bromide (Sigma) to the solution

Pour into gel casting tray to set

The gel can be stored at 4°C until needed, if covered in gladwrap and foil

Other percentage gels can be used. An x% gel uses xg agarose

3% agarose

3g agarose (Hispanger)

100ml 1xTBE

Mix together, and boil until agarose is completely dissolved

Allow to cool to ~60°C

Add 3µl 10mg/ml Ethidium Bromide (Sigma) to the solution

Pour into gel casting tray to set

The gel can be stored at 4°C until needed, if covered in gladwrap and foil

Depurinating solution (0.025M HCl solution)

20ml Concentrated HCl (Saarchem)

780ml ddH_2O

Denaturing solution

87.66g NaCl (final concentration of 1.5M) (Saarchem)

20g NaOH (final concentration of 0.5M) (Analysed Analytical Reagent®)

Made up to 1L with ddH_2O

12.5mM dNTP mix (for PCR)

12.5µl 100mM dATP (Promega)

12.5µl 100mM dCTP (Promega)

12.5µl 100mM dGTP (Promega)

12.5µl 100mM dTTP (Promega)

950µl ddH_2O

Aliquot into smaller volumes and store at -20°C

70% Ethanol

70ml absolute Ethanol (Merck)

30ml ddH_2O

0.5M Ethylenediamine Tetra-Acetic Acid (EDTA)

93.06g EDTA disodium salt crystals (MW=372.24) (Saarchem)

Adjust to pH8.0 with 5M NaOH to allow crystals to dissolve completely

Make up to 500ml with ddH_2O

Autoclave and store at room temperature

Ficoll-Bromophenol Blue Loading Dye

50g 50% sucrose crystals (USB)

0.1ml 0.5M EDTA (pH7.0)

0.1g 0.1% bromophenol blue dye (Sigma)

10g 10% Ficoll®-400 powder (Sigma-Aldrich)

Make up to 100ml with ddH_2O

Aliquot into 1.5ml Eppendorf tubes

Store at 4°C

1M $MgCl_2$

20.33g $MgCl_2$ (MW = 203.31g/l) (Saarchem)

Make up to 100ml with ddH_2O

Autoclave and store at room temperature

1kb or 1kb+ Molecular Weight DNA Marker (1μg/μl)

250μl ladder (0.25μg/μl stock) (GIBCO BRL)

125μl Ficoll-Bromophenol Blue Loading Dye

2.1ml 1xTE

Store in smaller aliquots at 4°C

Neutralising solution

87.6g NaCl (Saarchem) (final concentration of 1.5M)

60.6g Tris base (Roche) (final concentration of 0.5M)

0.292g EDTA (Saarchem) (final concentration of 0.01M)

42ml concentrated HCl (Saarchem)

Make up to 1L with ddH_2O

pH should be close to 7.2. Adjust with HCl if necessary

0.144M NH_4Cl

77g NH_4Cl

1L ddH_2O

Store at room temperature

0.01M NH_4HCO_3

7.95g NH_4HCO_3

1L ddH_2O

Work in fume hood to avoid inhaling NH_4HCO_3 fumes during preparation

Store at room temperature

Prehybridisation buffer (modified Church and Gilbert Solution)

12.78g Na_2HPO_4 (AnalaR)

9.75g $Na_2H_2PO_4.2H_2O$ (PAL Chemicals)

17.5g SDS (Sigma)

5ml 0.5M EDTA (Saarchem)

Make up to 250ml with ddH_2O

pH should be close to 7.2 (use 5M NaOH or NaOH pellets)

Filter sterilise and store at room temperature

Heat to 65°C to dissolve crystals before use

10mg/ml Proteinase-K

10mg Proteinase-K powder (Roche)

1ml ddH_2O

Dissolve and store at -20°C until needed

Proteinase-K mix

850µl 10% SDS (Sigma)

34µl 0.5M EDTA (Saarchem)

1700µl 10mg/ml Proteinase-K

Make up to 8000µl with ddH_2O

Use 500µl per sample for DNA extraction

ROX500-HiDi Formamide mix

7µl GeneScan Rox500 marker (Applied Biosystems)

146µl Hi-Di™ Formamide (Highly deionised formamide) (Applied Biosystems)

This mix is sufficient for 1 run of 16 samples with an aliquot of 9µl per sample

Saturated NaCl

40g NaCl (Saarchem)

100ml autoclaved ddH_2O

Add NaCl slowly until ddH_2O is absolutely saturated

Sephadex solution

5g Sephadex® G-50 fine beads (Fluka)

100ml 1x TE buffer (pH 8.0)

Swell the Sephadex beads in the 1x TE buffer

10% Sodium dodecyl sulphate (SDS)

100g SDS powder (Sigma)

1L autoclaved ddH_2O

Weigh the SDS powder under a fume hood

Southern blot washes

Final concentration	Stock solution	For 100ml
2x SSC	20x SSC	10ml
1x SSC	20x SSC	5ml
0.5x SSC	20x SSC	2.5ml
0.1x SSC	20x SSC	0.5ml
0.1% SDS	10% SDS	1ml

Special Neutralising Solution

60.6g Tris(hydroxymethyl)aminomethane (Tris) (Roche)

175.4g NaCl (Saarchem) (final concentration of 3M)

Made up to 1L with ddH_2O after adjusting the pH to 7.0

Approximately 50ml of concentrated HCl (Saarchem) adjusts pH to 7.0

20x SSC solution

175.35g NaCl (Saarchem) (final concentration of 3M)

88.25g Tri-sodium citrate (Associated Chemical Enterprises) (0.3M final concentration)

Made up to 1L with ddH_2O

Sucrose-Triton-X Lysing Buffer

10ml 1M Tris-HCl (pH8.0)

5ml 1M $MgCl_2$

10ml Triton-X 100 (Merck)

Make up to 1L with ddH_2O, autoclave and keep at room temperature until needed

Add 109.5g sucrose (USB) just before use. Stir until dissolved and keep the solution chilled at 4°C

1.5M Tris (either pH8.8 or 6.8)

Dissolve 121.14g Tris(hydroxymethyl)aminomethane (Tris) (Roche)

in 800ml ddH_2O.

Adjust to required pH using concentrated HCl (Saarchem), approximately 4.2ml for pH8.0

Adjust the volume to 1L with ddH_2O

Autoclave and store at room temperature

10x Tris Borate EDTA buffer (TBE)

109.02g Tris base (Roche)

55.64g Boric acid (Promega)

7.44g NaEDTA (Saarchem)

Dissolve all reagents in 1L ddH_2O. Adjust the pH to 8.3 with HCl (Saarchem)

Autoclave and store at room temperature

Dilute 10-fold before use

$T_{20}E_5$

20ml 1M Tris-HCl

10ml 0.5M EDTA

Make up to 1L with ddH_2O with a pH of 8.0

Store at room temperature

1x Tris-EDTA (TE) Buffer

10ml 1M Tris-HCl (pH8.0)

2ml 0.5M EDTA

Make up to 1L with ddH_2O

Autoclave and store at room temperature

Working Lysis Buffer

100ml 0.144M NH_4Cl

100ml 0.01M NH_4HCO_3

Make up to 1L with ddH_2O

Autoclave and store at room temperature

APPENDIX E

PROTOCOL SHEETS FOR PCRS

E1 PCR PROTOCOL FOR DETECTION OF THE INTRON 1 INVERSION

PCR components

Component	Stock concentration	Final concentration	1x mix
DNA			1µl
Amplitaq buffer	10x	1x	2.5µl
$MgCl_2$	25mM	2.5mM	2.5µl
dNTP mix	10x	1x	2.5µl
DMSO	100%	5%	1.25µl
Primer: int1h-2F	20pmol/µl	20pmol	1µl
Primer: 9F	20pmol/µl	20pmol	1µl
Primer: 9cR	20pmol/µl	20pmol	1µl
Amplitaq polymerase	5U/µl	2.5U	0.5µl
ddH_2O			11.75µl
TOTAL			25µl

PCR conditions

95°C – 5 minutes	1 cycle
95°C – 30 seconds, 63°C – seconds; 72°C – 2 minutes	30 cycles
72°C – 5 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 15µl PCR products on a 1.5% agarose gel.
PCR product size = 1.5kb (normal band) or 1kb (mutant band)

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E2 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 1A

PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-RNA-1a-F	10pmol/µl	20pmol	2µl
Primer: HPA-1-seq-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes	1 cycle
95°C – 1:30 minutes, 54°C – 1:30 minutes; 72°C – 2:30 minutes	35 cycles
72°C – 20 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 765bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E3 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 1B

First round PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			1µl
High fidelity buffer 2	10x	1x	1.25µl
dNTP mix	10x	1x	1.25µl
Betaine			2.5µl
Primer: HPA-RNA-1a-F	10pmol/µl	20pmol	1µl
Primer: HPA-RNA-1b-R	10pmol/µl	20pmol	1µl
High fidelity taq polymerase	3.5U/µl	0.35U	0.1µl
ddH_2O			6.9µl
TOTAL			12.5µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 48°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____

Second round PCR components

Component	Stock concentration	Final concentration	1x mix
First round PCR product			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-PTT1-seq-F	10pmol/µl	20pmol	2µl
Primer: HPA-PTT-1d-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 48°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl second round PCR products on a 1.5% agarose gel.
Expected PCR product size = 737bp.

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			

E4 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 2

First round PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			1µl
High fidelity buffer 2	10x	1x	1.25µl
dNTP mix	10x	1x	1.25µl
Primer: HPA-RNA-2a-F	10pmol/µl	20pmol	1µl
Primer: HPA-RNA-2b-R	10pmol/µl	20pmol	1µl
High fidelity taq polymerase	3.5U/µl	0.35U	0.1µl
ddH_2O			6.9µl
TOTAL			12.5µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 60°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____

Second round PCR components

Component	Stock concentration	Final concentration	1x mix
First round PCR products			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Spermidine	1/40 dilution of 0.1M		2.5µl
Primer: HPA-PTT-2c-F-new	10pmol/µl	20pmol	2µl
Primer: HPA-RNA-2d-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			11.3µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 64°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl second round PCR products on a 1.5% agarose gel.
Expected PCR product size = 1243bp.

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			

E5 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 3

PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Primer: HPA-RNA-3c-F	10pmol/µl	20pmol	2µl
Primer: HPA-PTT-3f-R-new	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			13.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes	1 cycle
95°C – 1:30 minutes, 60°C – 1:30 minutes; 72°C – 2:30 minutes	35 cycles
72°C – 20 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 1483bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E6 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 4A

PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-RNA-4c-F	10pmol/µl	20pmol	2µl
Primer: HPA-PTT4-seq-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes	1 cycle
95°C – 1:30 minutes, 50°C – 1:30 minutes; 72°C – 2:30 minutes	35 cycles
72°C – 20 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 799bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E7 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 4B

PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-PTT4-seq-F	10pmol/µl	20pmol	2µl
Primer: HPA-RNA-4f-R-new	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes	1 cycle
95°C – 1:30 minutes, 50°C – 1:30 minutes; 72°C – 2:30 minutes	35 cycles
72°C – 20 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 983bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E8 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 5

First round PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			1µl
High fidelity buffer 2	10x	1x	1.25µl
dNTP mix	10x	1x	1.25µl
Primer: HPA-RNA-5a-F	10pmol/µl	20pmol	1µl
Primer: HPA-RNA-5b-R	10pmol/µl	20pmol	1µl
High fidelity taq polymerase	3.5U/µl	0.35U	0.1µl
ddH_2O			6.9µl
TOTAL			12.5µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 60°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____

Second round PCR components

Component	Stock concentration	Final concentration	1x mix
First round PCR products			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-RNA-5c-F	10pmol/µl	20pmol	2µl
Primer: HPA-RNA-5d-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 56°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 1275bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			

E9 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 6A

PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-PTT-6-F	10pmol/µl	20pmol	2µl
Primer: HPA-PTT6-seq-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes	1 cycle
95°C – 1:30 minutes, 54°C – 1:30 minutes; 72°C – 2:30 minutes	35 cycles
72°C – 20 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 641bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E10 PCR PROTOCOL FOR AMPLIFICATION OF cDNA FRAGMENT 6B

First round PCR components

Component	Stock concentration	Final concentration	1x mix
cDNA			1µl
High fidelity buffer 2	10x	1x	1.25µl
dNTP mix	10x	1x	1.25µl
Betaine			2.5µl
Primer: HPA-RNA-6a-F	10pmol/µl	20pmol	1µl
Primer: HPA-RNA-6b-R	10pmol/µl	20pmol	1µl
High fidelity taq polymerase	3.5U/µl	0.35U	0.1µl
ddH_2O			4.4µl
TOTAL			12.5µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 48°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____

Second round PCR components

Component	Stock concentration	Final concentration	1x mix
First round PCR products			2µl
High fidelity buffer 2	10x	1x	2.5µl
dNTP mix	10x	1x	2.5µl
Betaine			5µl
Primer: HPA-PTT6-seq-F	10pmol/µl	20pmol	2µl
Primer: HPA-PTT-6d-R	10pmol/µl	20pmol	2µl
High fidelity taq polymerase	3.5U/µl	0.7U	0.2µl
ddH_2O			8.8µl
TOTAL			25µl

PCR conditions

95°C – 10 minutes 1 cycle
95°C – 1:30 minutes, 52°C – 1:30 minutes; 72°C – 2:30 minutes 35 cycles
72°C – 20 minutes 1 cycle
15°C – hold
PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 1.5% agarose gel.
Expected PCR product size = 664bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			

E11 PCR PROTOCOL FOR AMPLIFICATION OF DNA EXONS

The primers, annealing temperatures and product size specific for each exon of the *FVIII* gene can be found in Appendix C, Table C4.

PCR components

Component	Stock concentration	Final concentration	1x mix
DNA			2µl
Amplitaq buffer	10x	1x	2.5µl
$MgCl_2$	25mM	2.5mM	2.5µl
dNTP mix	10x	1x	2.5µl
Primer: _____-F	10pmol/µl	10pmol	1µl
Primer: _____-R	10pmol/µl	10pmol	1µl
Amplitaq polymerase	5U/µl	1U	0.2µl
ddH_2O			14.3µl
TOTAL			25µl

PCR conditions

95°C – 5 minutes	1 cycle
95°C – 1 minute, ____°C – 1 minute; 72°C – 1 minute	_____ cycles
72°C – 10 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 3% agarose gel.
Expected PCR product size = _____bp

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E12 PCR PROTOCOL FOR AMPLIFICATION OF INTRON 13 INTRAGENIC MARKER

PCR components

Component	Stock concentration	Final concentration	1x mix
DNA			2µl
Amplitaq buffer	10x	1x	2.5µl
$MgCl_2$	25mM	2.5mM	2.5µl
dNTP mix	10x	1x	2.5µl
Primer: 1A-F (FAM)	10pmol/µl	10pmol	1µl
Primer: 2A-R	10pmol/µl	10pmol	1µl
Amplitaq polymerase	5U/µl	1U	0.2µl
ddH_2O			14.3µl
TOTAL			25µl

PCR conditions

95°C – 5 minutes	1 cycle
95°C – 1 minute, 45°C – 1 minute; 72°C – 1 minute	30 cycles
72°C – 10 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 5µl PCR products on a 3% agarose gel.
Expected PCR product size = 137-155bp
Analyse products further using microsatellite analysis. Run 1µl undiluted PCR products with 9µl Rox500-HiDi mix (for 16 samples use 146µl HiDi formamide and 7µl Rox500) on the 3130*xl* genetic analyser.

Interpretation of results

Number of AC repeats = (PCR product size – 108bp) ÷ 2

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E13 PCR PROTOCOL FOR AMPLIFICATION OF INTRON 22 INTRAGENIC MARKER

PCR components

Component	Stock concentration	Final concentration	1x mix
DNA			2µl
Amplitaq buffer	10x	1x	2.5µl
$MgCl_2$	25mM	2.5mM	2.5µl
dNTP mix	10x	1x	2.5µl
DMSO	100%	10%	2.5µl
Primer: 1B-F (HEX)	10pmol/µl	10pmol	1µl
Primer: 2B-R	10pmol/µl	10pmol	1µl
Amplitaq polymerase	5U/µl	1U	0.2µl
ddH_2O			10.8µl
TOTAL			25µl

PCR conditions

95°C – 5 minutes	1 cycle
95°C – 1 minute, 45°C – 1 minute; 72°C – 1 minute	30 cycles
72°C – 10 minutes	1 cycle
15°C – hold	

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 25µl PCR products on a 3% agarose gel.
Expected PCR product size = 69-87bp
Analyse products further using microsatellite analysis. Run 1µl undiluted PCR products with 9µl Rox500-HiDi mix (for 16 samples use 146µl HiDi formamide and 7µl Rox500) on the 3130*xl* genetic analyser.

Interpretation of results

Number of AC repeats = (PCR product size – 32bp) ÷ 2

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

E14 PCR PROTOCOL FOR AMPLIFICATION OF P39 EXTRAGENIC MARKER

PCR components

Component	Stock concentration	Final concentration	1x mix
DNA			2µl
Amplitaq buffer	10x	1x	2.5µl
$MgCl_2$	25mM	2.5mM	2.5µl
dNTP mix	10x	1x	2.5µl
Primer: p39-F (HEX)	10pmol/µl	10pmol	1µl
Primer: p39-R	10pmol/µl	10pmol	1µl
Amplitaq polymerase	5U/µl	1U	0.2µl
ddH_2O			14.3µl
TOTAL			25µl

PCR conditions

94°C – 45 seconds, 65°C – 50 seconds; 72°C – 30 seconds 30 cycles
72°C – 10 minutes 1 cycle
15°C – hold

PCR machine: Perkin Elmer GeneAmp 2720: PCR machine #:____
Check 5µl PCR products on a 3% agarose gel.
Expected PCR product size = ~152-166bp
Analyse products further using microsatellite analysis. Run 1µl undiluted PCR products with 9µl Rox500-HiDi mix (for 16 samples use 146µl HiDi formamide and 7µl Rox500) on the 3130*xl* genetic analyser.

Interpretation of results

Allele size = PCR product size (results are not reported as repeat numbers as with intron 13 and intron 22 microsatellites)

Samples

Lane	Patient name	HPA code	Result
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

APPENDIX F

SUMMARY TABLE OF RESULTS

Table F1: Mutations in the *FVIII* gene identified in South African haemophilia A patients.

Mutation type	Code	Ethnicity	Severity	Inhibitors	Intron 22 inversion	DNA change	Observed cDNA change	Predicted protein change	*FVIII* position	Novel mutation	Polymorphisms identified
Inversion	HPA007	White	Moderate	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA041/204	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA127	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA169	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA174	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA182	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA199	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA270	White	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA031	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA159	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA189	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA191	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA193	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA213	Black	Severe	Yes	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA237	Black	Unknown	Unknown	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA271	Black	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA277	Black	Unknown	Unknown	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA167	Indian	Severe	No	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA227	Mixed ancestry	Severe	Yes	Positive – type I	-	-	-	Intron 22	No	
Inversion	HPA035	White	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA138	White	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA170	White	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA195	White	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA203	White	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA196	Black	Severe	Yes	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA206	Black	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA207	Black	Severe	Yes	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA211	Black	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA235	Black	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA239	Black	Severe	Yes	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA240	Black	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA275	Black	Unknown	Unknown	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA198	Mixed ancestry	Severe	No	Positive – type II	-	-	-	Intron 22	No	
Inversion	HPA107	White	Severe	No	Positive – atypical	-	-	-	Intron 22	No	

Table F1: Continued

Mutation type	Code	Ethnicity	Severity	Inhibitors	Intron 22 inversion	DNA change	Observed cDNA change	Predicted protein change	*FVIII* position	Novel mutation	Polymorphisms identified
Inversion	HPA177	Black	Severe	No	Positive – atypical	-	-	-	Intron 22	No	
Inversion	HPA265	Black	Unknown	No	Negative	-	-	-	Intron 1	No	
Inversion	HPA266	Black	Severe	No	Negative	-	-	-	Intron 1	No	D1241E
Inversion	HPA273	Indian	Severe	Yes	Negative	-	-	-	Intron 1	No	
Missense	HPA186	Black	Severe	No	Negative	29618A→T	431A→T	D125V	Exon 4	Yes	M2238V
Missense	HPA091	White	Moderate	?	Negative	29700G→A	514G→A	V162M	Exon 4	No	D1241E
Missense	HPA172	White	Moderate	No	Negative	56198C→T	1171C→T	R372C	Exon 8	No	
Missense	HPA205	Black	Severe	No	Negative	56267A→G	1241A→G	T395C	Exon 8	Yes	
Missense	HPA084	Black	Severe	No	Inferred negative	56716G→A	1406G→A	G450E	Exon 9	Yes	M2238V
Missense	HPA183	Black	Severe	No	Inferred negative	56716G→A	1406G→A	G450E	Exon 9	Yes	A544A, M2238V
Missense	HPA261	Indian	Mild	No	Inferred negative	61568A→C	1457A→C	N467T	Exon 10	Yes	
Missense	HPA112	White	Severe	No	Negative	68826T→C	1898T→C	M614T	Exon 12	Yes	
Missense	HPA037	White	Moderate	No	Negative	91100G→A	2167G→A	A704T	Exon 14	No	
Missense	HPA218	Black	Severe		Inferred negative	91927T→G	2994T→G	H979Q	Exon 14	Yes	D1241E
Missense	HPA241	Black	Severe	No	Inferred negative	93692G→A	4759G→A	D1568N	Exon 14	Yes	D1241E
Missense	HPA272	Black	Moderate	No	Negative	116283G→A	5353G→A	E1766K	Exon 14	Yes	D1241E
Missense	HPA201	White	Moderate	No	Inferred negative	117871G→A	5544G→A	E1829K	Exon 16	No ¥	
Missense	HPA247	Black	Severe	No	Inferred negative	118700T→A	5881T→A	W1942R	Exon 18	Yes	
Missense	HPA232	White	Mild	No	Inferred negative	160972C→G	6690C→G	R2209G	Exon 24	No	
Missense	HPA234	White	Moderate	No	Inferred negative	162135G→T	6744G→T	W2229C	Exon 25	No	
Missense	HPA248	White	Mild	No	Inferred negative	185027C→T	6956C→T	P2300L	Exon 26	No	
Nonsense	HPA033	White	Severe	No	Negative	56262-56263 GG→AC	1236-1237 GG→AC	WD393-394XH	Exon 8	W393X = No ¥	
Nonsense	HPA254	White	Moderate	No	Inferred negative	56646C→T	1336C→T	R427X	Exon 9	No	
Nonsense	HPA175	Black	Severe	No	Negative	91220C→T	2287C→T	Q744X	Exon 14	Yes	
Nonsense	HPA269	Black	Severe	Yes	Inferred negative	No result	2440C→T	R795X	Exon 14	No	
Nonsense	HPA225	White	Severe	No	Negative	92024A→T	3091A→T	K1012X	Exon 14	Yes	D1241E
Nonsense	HPA115	White	Severe	No	Negative	92676T→G	3743T→G	L1229X	Exon 14	Yes	
Nonsense	HPA164	White	Severe	No	Negative	93875C→T	4942C→T	Q1629X	Exon 14	Yes	D1241E
Nonsense	HPA060	Black	Severe	No	Negative	94077C→T	5144C→T	R1696X	Exon 14	No	
Nonsense	HPA184	Indian	Severe	No	Negative	162270A→G	6879A→G	W2271X	Exon 25	No	
Deletion	HPA202	Black	Severe	No	Negative	?	1445-1538del	Ex10-11 deletion	Exon10&11	●	
Deletion	HPA258	Black	Severe	No	Inferred negative	74927delTTC	2014delTTC	ΔF652	Exon 13	No	
Deletion	HPA274	White	Severe	No	Inferred negative	92482delAA	3548delAA	K1164fsX13	Exon 14	No	
Deletion	HPA194	Black	Moderate	No	Negative	92571delA	3637delA	I1194fsX5	Exon 14	No	D1241E

Table F1: Continued.

Mutation type	Code	Ethnicity	Severity	Inhibitors	Intron 22 inversion	DNA change	Observed cDNA change	Predicated protein change	FVIII position	Novel mutation	Polymorphisms identified
Deletion	HPA160	Black	Severe	No	Negative	92679delCTGA	3745delCTGA	L1230fsX6	Exon 14	Yes	P1481P
Deletion	HPA256	White	Severe	No	Negative	93052delA	4118delA	Q1354fsX1	Exon 14	Yes	
Deletion	HPA068/215	White	Severe	No	Negative	93255delAGAA	4321delAGAA	K1424fsX20	Exon 14	No	
Deletion	HPA013	White	Carrier (dad severe)	Unknown	Negative	93313delA	4379delA	N1441fsX5	Exon 14	No	
Deletion	HPA128	White	Severe	No	Negative	93854delA	4920delA	E1621fsX1	Exon 14	Yes	S1269S
Insertion	HPA011/129	White	Severe	No	Negative	92571insA	3637insA	I1192fsX27	Exon 14	No	
Insertion	HPA015	White	Severe	No	Negative	92571insA	3637insA	I1194fsX27	Exon 14	No	
Insertion	HPA050	White	Severe	No	Negative	92571insA	3637insA	I1194fsX27	Exon 14	No	
Insertion	HPA118	White	Severe	No	Negative	92571insA	3637insA	I1194fsX27	Exon 14	No	
Insertion	HPA230	White	Severe	No	Inferred negative	92571insA	3637insA	I1194fsX27	Exon 14	No	
Insertion	HPA249	White	Severe	No	Negative	92571insA	3637insA	I1194fsX27	Exon 14	No	
Insertion	HPA029	White	Severe	No	Negative	93135insC	4201insC	P1381fsX5	Exon 14	Yes	S1269S
Splice site	HPA236	White	Moderate	No	Inferred negative	75023C→T	1904-2113del (ex13 del)	P687L = ex13 deletion	Exon 13	Yes	
Splice site	HPA176	Black	Severe	No	Inferred negative	91048G→C = IVS13-1G→C	2114-2143del (partial ex14 del)	686-696del (partial ex14 deletion)	Intron 13	Yes	
Complex	HPA219 *	Mixed ancestry	Severe	Yes	Inferred negative	?	1445-1538del (ex10 del)	463-494del (ex10 deletion)	Exon 10	●	M2238V
Complex	HPA020 *	White	Moderate	No	Inferred negative	?	113bp ins@2120	Insertion of 37.3amino acids	Intron 13	●	
Complex	HPA223	White	Severe	No	Inferred negative	?	5373-5419del	Partial ex16 del (del1773-1788)	Exon 16	●	IVS16+37delT ∂
Complex	HPA017 *	White	Severe	No	Negative	?	6005-6121del	Exon 19 deletion (1981-2020del)	Exon 19	●	
Complex	HPA210 *	Black	Severe	No	Negative	?	6005-6121del	Exon 19 deletion (1981-2020del)	Exon 19	●	IVS19-103T→C∂

Nomenclature of the *FVIII* mutations in this study is as follows (as per standard nomenclature used by the HAMSTeRS mutation database):

The DNA change is determined where the "A" of the start codon is the first base;
The cDNA sequence starts 171bp into the *FVIII* mRNA and
The amino acid sequence of the mature protein starts 19 amino acids into the native protein sequence.

* In these patients, a mutation was identified on one amplified product of cDNA. However, these mutations could not be confirmed on repeat cDNA sequencing, due to unsuccessful amplification or a mutation that was outside the region sequenced on DNA.

● The exact DNA change could not be elucidated. We were therefore unable to determine if these mutations were previously reported.

¥ These mutations were not previously reported on the HAMSTeRS mutation database, but were reported by Green, *et al.* (2008).

∂ These changes are assumed to be polymorphic as no evidence could be found for pathogenicity in the respective patients. They could be useful in linked marker analysis for at-risk relatives.

7 REFERENCES

Ahmed, R., Ivaskevicius, V. & Kannan, M., *et al.* 2005. Identification of 32 novel mutations in the factor VIII gene in Indian patients with haemophilia A. *Haematologica,* vol. 90, pp. 283-284.

Ananyeva, N.M., Kouiavskaia, D.V. & Shima, M., *et al.* 2001. Catabolism of the coagulation factor VIII: can we prolong lifetime of fVIII in circulation? *Trends in Cardiovascular Medicine,* vol. 11, pp. 25107.

Antonarakis, S.E., Rossiter, J.P. & Young, M., *et al.* 1995. Factor VIII gene inversions in severe haemophilia A: results of an international consortium study. *Blood,* vol. 86, pp. 2206-2212.

Arun, B. & Kessler, C.M. 2001. Clinical manifestations and therapy of the haemophilias. In: *Haemostasis and Thrombosis: basic Principles and Clinical practice*, 4th ed. Philadelphia: Lippncott-Raven, pp. 815-824.

Astermark, J., Berntorp, E. & White, G.C., *et al.* 2001. The Malmö International Brother Study (MIBS): further support for genetic predisposition to inhibitor development in haemophilia patients. *Haemophilia,* vol. 7, pp. 267-272.

Astermark, J., Oldenburg, J. & Escobar, M., *et al.* 2005. The Malmö International Brother Study (MIBS). Genetic defects and inhibitor development in siblings with severe hemophilia A. *Haematologica*, vol. 90, pp. 924-31.

Astermark, J. 2006. Basic aspects of inhibitors to factors VIII and IX and the influence of non-genetic risk factors. *Haemophilia,* vol. 12, pp. 52-60.

Bagnall, R.D., Waseem, N. & Green, P.M., *et al.* 2002. Recurrent inversion breaking intron 1 of the factor VIII gene is a frequent cause of severe haemophilia A. *Blood,* vol. 99, pp. 168-174.

Bagnall, R.D., Giannelli, F. & Green, P.D. 2005. Polymorphism and haemophilia A causing inversions in distal Xq28: a complex picture. *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 2598-2599.

Bagnall, R.D., Giannelli, F. & Green, P.D. 2006. *Int22h*-related inversions causing haemophilia A: a novel insight into their origin and a new more discriminant PCR test for their detection. *Journal of Thrombosis and Haemostasis*, vol, 4, pp. 591-598.

Barrow, R.T., Healey, J.F. & Gailani, D., *et al.* 2000. Reduction of the antigenicity of factor VIII toward complex inhibitory antibody plasmas using multiply-substituted hybrid human/porcine factor VIII molecules. *Blood,* vol. 95, pp. 564-568.

Baty, B.J., Drayna, D. & Leonard, C.O., *et al.* 1986. Prenatal diagnosis of factor VIII deficiency to help with the management of pregnancy and delivery. *Lancet,* vol. 1, pp. 207.

Becker, J., Schwaab, R. & Möller-Taube, A., *et al.* 1996. Characterisation of the factor VIII defect in 147 patients with sporadic haemophilia A: Families studies indicate a mutation type-dependent sex ratio of mutation frequencies. *American Journal of Human Genetics,* vol. 58, pp. 657-670.

Bhopale, G.M. & Nanda, R.K. 2003. Blood coagulation factor VIII: An overview. *Journal of Biosciences,* vol. 28, pp. 783-789.

Boekhorst, J., Verbruggen, B. & Lavergne, J.M., *et al.* 2005. Thirteen novel mutations in the *factor VIII* gene in the Nijmegen haemophilia A patient population. *British Journal of Haematology,* vol. 131, pp. 109-117.

Bowen, D.J. 2002. Haemophilia A and Haemophilia B: molecular insights. *Journal of Clinical Pathology: Molecular Pathology,* vol. 55, pp. 127-144.

Brandstetter, H., Bauer, M. & Huber, R., *et al.* 1995. X-ray structure of clotting factor IXa: active site and module structure related to Xase activity and haemophilia B. *Proceedings of the National Academy of Science of the United States of America,* vol. 92, pp. 9796-9800.

Brettler, D.B. 2005. Gene therapy for haemophilia? *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1317-1319.

Clarke, G.M. 1997. Probability, the binomial and the geometric distributions. In *Statistics & Experimental Design, An introduction for Biologists & Biochemists,* 3rd ed. London: Arnold, pp 35-40.

Coleman, R., Genet, S.A. & Harper, J.I., *et al.* 1993. Interaction of incontinentia pigmenti and factor VIII mutations in a female with biased X inactivation, resulting in haemophilia. *Journal of Medical Genetics,* vol. 30, pp. 497-500.

Cutler, J.A., Mitchell, M.J. & Smith, M.P., *et al.* 2002. The identification and classification of 41 novel mutations in the factor VIII gene (F8C). *Human Mutation,* vol. 19, pp. 274-278.

Dahlbäck, B. 2000. Blood coagulation. *Lancet,* vol. 355, pp. 1627-1632.

Dangerfield, B.T.F., Manga, P. & Field, S.P., *et al.* 1997. Feasibility of prenatal diagnosis and carrier detection in South African haemophilia A patients. *British Journal of Haematology,* vol. 97, pp. 558-560.

Darby, S.C., Keeling, D.M. & Spooner, R.J., *et al.* 2004. The incidence of factor VIII and factor IX inhibitors in the haemophilia population of the UK and their effect on subsequent mortality, 1977-99. *Journal of Thrombosis and Haemostasis,* vol. 2, pp. 1047-1055.

Dean, G. 1971. *The Porphyrias.* London: Pitman medical. pp. 103-112.

Den Dunnen, J.T. & van Ommem, G.J. 1999. The protein truncation test: a revew. *Human Mutations,* vol. 14, pp. 95-102.

Den Dunnen, J.T. & Antonarakis, S.E. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Human Mutations,* vol. 15, pp. 7-12.

DiMichele, D. 2005. Gene therapy for hemophilia? The debate reframed. *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1315-1316.

DiMichele, D., Chuansumrit, A. & London, A.J., *et al.* 2006. Ethical issues in haemophilia. *Haemophilia,* vol. 12, pp. 30-35.

Eaton, D., Rodriguez, H. & Vehar, G.A. 1986. Proteolytic processing of human factor VIII. Correlation of specific cleavages by thrombin, factor Xa, and activated protein C with activation and inactivation of factor VIII coagulant activity. *Biochemistry,* vol. 25, pp. 505-512.

El-Maarri, O., Singer, H. & Klein, C., *et al.* (2006). Lack of *F8* mRNA: a novel mechanism leading to hemophilia A. *Blood*, vol. 107, pp2759-2765.

Evatt, B.L. & Robillard, L. 2000. Establishing haemophilia care in developing countries: using data to overcome the barrier of pessimism. *Haemophilia, vol.* 6, pp. 131-134.

Fang, H., Wang, L. & Wang, H. 2007. The protein structure and effect of factor VIII. *Thrombosis research,* vol. 119, pp. 1-13.

Freije, D. & Schlessinger, D. 1992. A 1.6-Mb contig of yeast artificial chromosomes around the human factor VIII gene reveals three regions homologous to probes for the DXS115 locus and two for the DXYS64 locus. *American Journal of Human Genetics,* vol. 51, pp. 66-80.

Fulcher, C.A., Roberts, J.R. & Zimmerman, T.S. 1983. Thrombin proteolysis of purified factor viii procoagulant protein: correlation of activation with generation of a specific polypeptide. *Blood,* vol. 61, pp. 807-811.

Gabrovsky, V. & Calos, M.P. 2008. Factoring gene therapy into a cure for haemophilia A. *Current opinions in Molecular Therapeutics*, vol. 10, pp464-470.

Ganguly, A., Rock, M.J. & Prockop, D.J. 1993. Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: Evidence for solvent-induced bends in DNA heteroduplexes. *Proceedings of the National Academy of Science of the United States of America*, vol. 90, pp.10325-10329.

Giangrande, P.L. 2004. Gene therapy for hemophilia? No. *Journal of Thrombosis and Haemostasis,* vol. 2, pp. 1238-1237.

Gill, J.C. 1999. The role of genetics in inhibitor formation. *Thrombosis and Haemostasis,* vol. 82, pp. 500-504.

Gitschier, J., Wood, W.I. & Goralka, T.M., *et al.* 1984. Characterization of the human factor VIII gene. *Nature,* vol. 312, pp. 326-330.

Goodeve, A.C. 1998. Laboratory methods for the genetic diagnosis of bleeding disorders. *Clinical Laboratory Haematology,* vol. 20, pp. 3-13.

Graw, J., Brackmann, H.H. & Oldenburg, J., *et al.* 2005. Haemophilia A: from mutation analysis to new therapies. *Nature Reviews Genetics,* vol. 6, pp. 488-501.

Green, P.M., Bagnall, R.D. & Waseem, N.H., *et al.* 2008. Haemophilia A mutations in the UK: results of screening one-third of the population. *British Journal of Haematology,* vol. 143, pp. 115-128.

Hacein-Bey-Abina, S., Von Kalle, C. & Schmidt, M., *et al.* 2003. LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science*, vol. 302, pp. 415-419.

Hay, C.R.M. 1998. Factor VIII inhibitors in mild and moderate-severity haemophilia A. *Haemophilia,* vol. 4, pp 558-563.

Hazewinkel, M.H., Hoogerwerf, J.J. & Hesseling, P.B., *et al.* 2003. Haemophilia patents aged 0-18 years in the Western Cape. *South African Medical Journal,* vol. 93, pp. 793-796.

Hemker, H.C., Hermans, W.T. & Muller, A.D., *et al.* 1980. Oral treatment of haemophilia A by gastrointestinal absorption of factor VIII entrapped in liposomes. *The Lancet,* vol. 1, pp. 70-71.

Higuchi, M., Antonarakis, S.E. & Kasch, L., *et al.* 1991. Molecular characterization of mild-to-moderate hemophilia A: detection of the mutation in 25 of 29 patients by denaturing gradient gel electrophoresis. *Proceedings of the National Academy of Science of the United States of America,* vol. 88, pp. 8307-8311.

Hoots, W.K. 2005. Gene therapy for haemophilia? *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1316-1317.

Hoyer, L.W. 1994. Haemophilia A. *The New England Journal of Medicine,* vol. 330, pp. 38-47.

Jenkins, T., Nicholls, E. & Gordon, E., *et al.* 1980. Familial hypercholesterolaemia--a common genetic disorder in the Afrikaans population. *South African Medical Journal,* vol. 57, pp. 943-947.

Kaufman, R.J. 1992. Biological regulation of factor VIII activity. *Annual Review of Medicine,* vol. 43, pp. 325-339.

Kazazian, H.H., Wong, C. & Youssoufian, H., *et al.* 1988. Haemophilia A resulting from de novo insertion of L1 sequences represents a novel mechanism for mutation in man. *Nature,* vol. 332, pp. 164-166.

Keeney, S., Mitchell, M. & Goodeve, A. 2005. The molecular analysis of haemophilia A: a guideline from the UK haemophilia centre doctors' organisation haemophilia genetics laboratory network. *Haemophilia*, vol. 11, pp387-397.

Kulkarni, R. & Lusher, J. 2001. Perinatal management of newborns with haemophilia. *British Journal of Haematology*, vol. 112, pp. 264-274.

Lakich, D., Kazazian, H.H. & Antonarakis, S.E., *et al.* 1993. Inversions disrupting the factor VIII gene are a common cause of severe haemophilia A. *Nature Genetics,* vol. 5, pp. 236-241.

Lalloz, M.R.A., McVey, J.H. & Pattinson, J.K., *et al.* 1991. Haemophilia A diagnosis by analysis of a hypervariable dinucleotide repeat within the factor VIII gene. *The Lancet,* vol. 338, pp. 207-211.

Lalloz, M.R.A., Schwaab, R. & McVey, J.H., *et al.* 1994. Haemophilia A diagnosis by simultaneous analysis of two variable dinucleotide tandem repeats within the factor VIII gene. *British Journal of Haematology,* vol. 86, pp. 804-809.

Laprise, S.L., Mak, E.K. & Killoran, K.A., *et al.* 1998. Use of denaturing gradient gel blots to screen for point mutations in the factor VIII gene. *Human Mutation,* vol. 12, pp. 393-402.

Laurian, Y. 2005. Future perspective for the treatment of hemophilia A. *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1320-1321.

Lenting, P.J., van Mourik, J.A. & Mertens, K. 1998. The life cycle of coagulation factor VIII in view of its structure and function. *Blood*, vol. 92, pp. 3983-96.

Le Saux, O., Beck, K. & Sachsinger, C., *et al.* (2002). Evidence for a founder effect for pseudoxanthoma elasticum in the Afrikaner population of South Africa. *Human Genetics*, vol. 111, pp.331-338.

Leuer, M., Oldenburg, J. & Lavergne, J.M., *et al.* 2001. Somatic mosaicism in hemophilia A: a fairly common event. *American Journal of Human Genetics,* vol. 69, pp. 75-87.

Levinson, B., Kenwrick, S. & Lakich, D., *et al.* 1990. A transcribed gene in an intron of the human factor VIII gene. *Genomics,* vol. 7, pp. 1-11.

Levinson, B., Kenwrick, S. & Gamel, P., *et al.* 1992. Evidence for a third transcript from the human factor VIII gene. *Genomics,* vol. 14, pp. 585-589.

Liang, Y., Yan, M. & Xiao, B, *et al.* 2008. Analysis of intron 1 factor VIII gene inversion for direct diagnosis in patients with severe haemophilia A in China. *Prenatal Diagnosis,* vol. 28, pp. 160-161.

Lillicrap, D. 2005. Gene therapy: molecular engineering of factor VIII and factor IX. In: *Textbook of Haemophilia*. Oxford: Blackwell Publishing Ltd, pp.229-234.

Ludlam, C.A., Pasi, K.J. & Bolton-Maggs, P, *et al.* 2005. A framework for genetic service provision for haemophilia and other inherited bleeding disorders. *Haemophilia,* vol. 11, pp. 145-163.

Lusher, J.M., Arkin, S. & Abildgaard, C.F., *et al.* 1993. Recombinant factor VIII for the treatment of previously untreated patients with hemophilia A. Safety, efficacy, and development of inhibitors. Kogenate Previously Untreated Patient Study Group. *New England Journal of Medicine,* vol. 328, pp. 453-459.

Mahlangu, J.N. & Gilham, A. 2007. Guideline for the treatment of haemophilia in South Africa. *South African Medical Journal,* vol. 97, pp. 1296-1310.

Mannucci, P.M. & Tuddenham, E.G.D. 2001. The haemophilias--from royal genes to gene therapy. *New England Journal of Medicine,* vol. 344, pp. 1773-1779.

Mannucci, P.M. 2008. Back to the future: a recent history of haemophilia treatment. *Haemophilia*, vol. 14, pp10-18.

Maugard, C., Tuffery, S. & Aguilar-Martinez, P., *et al.* 1998. Protein Truncation Test: Detection of severe Haemophilia A mutation and analysis of factor VIII transcripts. *Human Mutation,* vol. 11, pp. 18-22.

Maxam, A.M. & Gilbert, W. 1977. A new method for sequencing DNA. *Proceedings of the National Academy of Science of the United States of America,* vol. 74, pp. 560-564.

McKibbin, J.C. 2002. Assesment of a polymorphic microsatellite marker to increase informativity in Afrikaans haemophilia A families [dissertation]. Johannesburg: University of the Witwatersrand.

McKusick, V.A. 1965. The Royal Haemophilia. *Scientific America,* vol. 213, pp. 88-95.

Migeon, B.R., Axelman, J. & Jan de Beur, S., *et al.* 1989. Selection against lethal alleles in females heterozygous for incontinentia pigmenti. *American Journal of Human Genetics,* vol. 44, pp. 100-106.

Miller, S.A., Dyk, D.D. & Pelesky, H.F. 1988. A Simple Salting-Out Procedure for Extracting DNA from Human Nucleated Cells. *Nucleic Acid Research,* vol. 16, pp. 1215.

Mueller, R.F. & Young, I.D. 1998. Glossary. In: *Emery's Elements of Medical Genetics*, tenth edition. United States of America: Churchill Livingstone, pp. 23, 345.

Nathwani, A.C. & Tuddenham, E.G.D. 1992. Epidemiology of coagulation disorders. *Baillière's Clinical Haematology,* vol. 5, pp. 383-439.

Naylor, J.A., Green, P.M. & Rizza, C.R., *et al.* 1992. Analysis of factor VIII mRNA reveals defects in everyone of 28 haemophilia A patients. *Human Molecular Genetics,* vol. 2, pp. 11-17.

Naylor, J., Brinke, A. & Hassock, S., *et al.* 1993. Characteristic mRNA abnormality found in half the patients with severe haemophilia A is due to large DNA inversions. *Human Molecular Genetics,* vol. 2, pp. 1773-1778.

Naylor, J.A., Buck, D. & Green, P., *et al.* 1995. Investigation of the factor VIII intron 22 repeated region (int22h) and the associated inversion junctions. *Human Molecular Genetics,* vol. 4, pp. 1217-1224.

Naylor, J.A., Nicholson, P. & Goodeve, A., *et al.* 1996. A novel DNA inversion causing severe haemophilia A. *Blood,* vol. 87, pp. 3255-61.

Negrier, C. 2004. Gene therapy for hemophilia? Yes. *Journal of Thrombosis and Haemostasis,* vol. 2, pp. 1234-5.

O'Brien, L.M., Mastri, M. & Fay, P.J. 2000. Regulation of factor VIIIa by human activated protein C and protein S: inactivation of cofactor in the intrinsic factor Xase. *Blood,* vol. 95, pp. 1714-1720.

Oefner, P.J. & Underhill, P.A. 1995. Comparative DNA sequencing by denaturing high performance liquid chromatography (dHPLC). *American Journal of Human Genetics*, vol. 57, pp. A266.

Oldenburg, J., Rost, S. & El-Maarri, O., *et al.* 2000. De novo factor VIII gene intron 22 inversion in a female carrier presents as a somatic mosaicism. *Blood* vol. 96, pp. 2905-2906.

Olednburg, J., Schroder, J. & Brackman, H.H., *et al.* 2004a. Environmental and genetic factors influencing inhibitor development. *Seminars in Haematology*, vol. 41, pp.82-88.

Oldenburg, J., Ananyeva, N.M. & Saenko, E.L. 2004b. Molecular basis of haemophilia A. *Haemophilia,* vol. 10, pp. 133-139.

Pandey, G.S. & Mittal, B. 2001. Molecular diagnosis of Haemophilia A. *Journal of Postgraduate Medicine*, vol. 47, pp. 274-280.

PCR clean-up Gel extraction, User manual, NucleoSpin Extract II, February 2006/Rev. 03.

Peters, M.F. & Ross, C.A. 2001. Isolation of a 40-kDa Huntingtin-associated protein. *Journal of Biological Chemistry,* vol. 276, pp. 3188-3194.

Peyvandi, F., Jayandharan, G. & Chandy, M., *et al.* 2006. Genetic diagnosis of haemophilia and other inherited bleeding disorders. *Haemophilia,* vol. 12, pp. 82-89.

Pipe, S.W. & Kaufman, R.J. 1997. Characterization of a genetically engineered inactivation-resistant coagulation factor VIIIa. *Proceedings of the National Academy of Science of the United States of America*, vol. 28, pp 11851-6.

Pittman, D.D. & Kaufman, R.J. 1988. Proteolytic requirements for thrombin activation of anti-haemophilic factor (factor VIII). *Proceedings of the National Academy of Science of the United States of America,* vol. 85, pp. 2429-2433.

Pola, V. & Svojitka, J. 1957. [Classic hemophilia in women]. *Folia Haematological,* vol. 75, pp. 43-51.

Poustka, A., Dietrich, A. & Langenstein, G., *et al.* 1991. Physical map of human Xq27-qter: localizing the region of the fragile X mutation. *Proceedings of the National Academy of Science of the United States of America,* vol. 88, pp. 8302-8306.

Pros, E., Gómes, C. & Martín, T., *et al.* 2008. Nature and mRNA effect of 282 different NF1 point mutations: focus on splicing alterations. *Human Mutations*, Vol. 29, pp. E173-E193.

QIAGEN EasyXpress™ Protein Synthesis Handbook, September 2003.

Reeves, M.D., Yawitch, T.M. & van der Merwe, N.C., *et al.* (2004). BRCA1 mutations in Aouth African breast and/or ovarian cancer families: evidence of a novel founder mutation in Afrikaner families. *International Journal of Cancer*, vol. 10, pp. 677-682.

Rhei, E., Bogomolniy, F. & Federici, M.G., *et al.* 1998, Molecular genetic characterization of BRCA1- and BRCA2-linked hereditary ovarian cancers. *Cancer Research*, vol. 58, pp. 3193-3196.

Rimoin, D.L., Connor, J.M. & Pyeritz, R.E., *et al.* 2007. Approaches to specific disorders. In: *Emery's and Rimoin's Principles and Practice of Medical Genetics*, 5th edition. New York: Elsevier, pp. 1708

Rizza, C.R., Spooner, R.D.J. & Giangrande, P.L.F. 2001. Treatment of haemophilia in the United Kingdom 1981-1996. *Haemophilia,* vol. 7, pp. 349-359.

Roest, P.A., Bout, M. & van der Tuijn, A.C., *et al.* 1996. Splicing mutations in DMD/BMD detected by RT-PCR/PTT: detection of a 19aa insertion in the cysteine rich domain of dystrophin compatible with BMD. *Journal of Medical Genetics*, vol. 33, pp. 935-939.

Ross, M.T., Grafham, D.V. & Coffey, A.J., *et al.* 2005. The DNA sequence of the human X chromosome. *Nature,* vol. 434, pp. 325-337.

Rossiter, J.P., Young, M. & Kimberland, M.L., *et al.* 1994. Factor VIII gene inversions causing severe haemophilia A originate almost exclusively in male germ cells. *Human Molecular Genetics,* vol. 3, pp. 1035-1039.

Sambrook, J. & Russell, D.W. 2001. In: *Molecular Cloning, a laboratory manual*, third edition. New York: Cold Spring Harbour Press, pp 6.33-6.39; 8.3-8.6; 12.3-12.6.

Sanger, F., Nicklen, S. & Coulson, A.R. 1977. DNA sequencing with chain-termination inhibitors. *Proceedings of the National Academy of Science of the United States of America*, vol. 74, pp. 5463-5467.

Santacroce, R., Santoro, R. & Sessa, F., *et al.* (2008). Screening of mutations of hemophilia A in 40 Italian patients: a novel G-to-A mutation in intron 10 of the F8 gene as a putative cause of mild hemophilia A in southern Italy. *Blood Coagulation Fibrinolysis*, vol. 19, pp. 197-202.

Scallan, C.D., Lillicrap, D. & Jiang, H., *et al.* 2003. Sustained phenotypic correction of canine haemophilia A using an adeno-associated viral vector. *Blood,* vol. 102, pp. 2031-2037.

Scharrer, I., Bray, G.L. & Neutzling, O. 1999. Incidence of inhibitors in haemophilia A patients--a review of recent studies of recombinant and plasma-derived factor VIII concentrates. *Haemophilia,* vol. 5, pp. 145-154.

Snyder, R.O., Miao, C. & Meuse, L, *et al.* 1999. Correction of hemophilia B in canine and murine models using recombinant adeno-associated viral vectors. *Nature Medicine,* vol. 5, pp. 64-70.

Sorenson, J.R., Jennings-Grant, T. & Newman, J. 2003. Communication about carrier testing within hemophilia A families. *American Journal of Human Genetics,* vol. 119, pp. 3-10.

Spearman, A.D., Sweet, K. & Zhou, X.P., *et al.* 2008. Clinically applicable models to characterise BRCA1 and BRCA2 variants of uncertain significance. *Journal of Clinical Oncology*, vol. 26, pp. 5393-5400.

Spiegel, P.C., Murphy, P. & Stoddard, B.L. 2004. Surface-exposed hemophilic mutations across the factor VIII C2 domain have variable effects on stability and binding activities. *Journal of Biological Chemistry,* vol. 279, pp. 53691-53698.

Srivastava, A. 2005. Gene therapy for hemophilia? *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1313.

Stevens, R.F. 1999. The history of haemophilia in the royal families of Europe. *British Journal of Haematology*, vol. 105, pp. 25-32.

Strauss, H.S. 1967. The perpetuation of haemophilia by mutations. *Pediatrics,* vol. 39, pp. 186-193.

Tipping, A.J., Pearson, T. & Morgan, N.V., *et al.* (2001). Molecular and genealogical evidence for a founder effect in Fanconi anemia families of the Afrikaner population of South Africa. *Proceedings of the National Academy of Science of the United States of America*, vol. 98, pp. 5734-5739.

Toole, J.J., Knopf, J.L. & Wozney, J.M., *et al.* 1984. Molecular cloning of a cDNA encoding human antihaemophilic factor. *Nature,* vol. 312, pp. 342.

Tuddenham, E.G.D., Cooper, D.N. & Gitschier, J., *et al.* 1991. Haemophilia A: database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene. *Nucleic Acid Research* vol. 19, pp. 4821-4833.

Tuddenham, E.G.D, Schwaab, R. & Seehafer, J., *et al.* 1994. Haemophilia A: database of nucleotide substitutions, deletions, insertions and rearrangements of the factor VIII gene, second edition. *Nucleic Acid Research*, vol. 22, pp. 4851-4868.

Tuddenham, E.D. 2005. Gene therapy for haemophilia? Gene therapy for haemophilia is both desirable and achievable in the near future. *Journal of Thrombosis and Haemostasis,* vol. 3, pp. 1314.

Vehar, G.A., Keyt, B. & Eaton, D., *et al.* 1984. Structure of human factor VIII. *Nature,* vol. 31, pp. 337-342.

Viel, K.R., Machiah, D.K. & Warren, D.M., *et al.* 2007. A sequence variation scan of the coagulation factor VIII (FVIII) structural gene and associations with plasma FVIII activity levels. *Blood,* vol. 109, pp. 3713-3724.

Voet, D. & Voet, J.G. 1995. Molecular physiology. In: *Biochemistry*, second edition. United States of America: John Wiley & Sons, Inc, pp. 1196-1207.

Vogel, F. 1977. A probable sex difference in some mutation rates. *American Journal of Human Genetics,* vol. 29, pp. 312-319.

Walker, F.J., Chavin, S.I. & Fay, P.J. 1987. Inactivation of factor VIII by activated protein C and protein S. *Archives of Biochemistry and Biophysics,* vol. 252, pp. 322-328.

Wehnert, M., Reiner, O. & Caskey, C.T. 1993. Four STR polymorphisms map to a 500 kb region between DXS15 and DXS134. *Human Molecular Genetics,* vol. 2, pp. 1503.

White, G.C. 2001. Gene therapy in hemophilia: clinical trials update. *Journal of Thrombosis and Haemostasis,* vol. 86, pp. 172-177.

White, G.C. & Samulski, R.J. 2005. Gene therapy introduction and overview. In: *Textbook of Haemophilia.* Oxford: Blackwell Publishing Ltd, pp. 214-219.

Wilson, K. & Walker, J.M. 1997. *Principles and techniques of practical biochemistry*, fourth edition Great Britain: Cambridge University Press, pp124-125, 142-146 & 154-156.

Windsor, S., Lyng, A. & Taylor, S.A.M., *et al.* 1995. Severe haemophilia A in a female resulting from two de novo factor VIII mutations. *British Journal of Haematology,* vol. 90, pp. 906-909.

Wion, K.L.., Kelly, D. & Summerfield, J.A., *et al.* 1985. Distribution of factor VIII mRNA and antigen in human liver and other tissues. *Nature,* vol. 317, pp. 726-729.

Wood, W.I., Capon, D.J. & Simonsen, C.C., *et al.* 1984. Expression of active human factor VIII from recombinant DNA clones. *Nature,* vol. 312, pp. 330.

Yamazaki, E., Mohri. H. & Inaba, H., *et al.* 1997. Variant of intron 22 inversions in the factor VIII gene in severe hemophilia A. *Blood Coagulation Fibrinolysis,* vol. 8, pp. 445-449.

8 ELECTRONIC REFERENCES

Bell, J., Bodmer, D. & Sistermans, E., *et al.* 2007. “Practice Guidelines for the Interpretation and Reporting of Unclassified Variants (UVs) in Clinical Molecular Genetics.” <http://cmgsweb.shared.hosting.zen.co.uk/BPGs/pdfs%20current%20bpgs/UV%20GUIDELINES%20ratified.pdf>

Boon, E.M.J. <e.m.j.boon@lumc.nl> “Haemophilia A genetic testing.” 4 December 2008. Personal e-mail.

Ganguly, A. <ganguly@mail.med.upenn.edu> “Haemophilia A genetic testing.” 4 December 2008. Personal e-mail.

Goodeve, A. <anne.goodeve@sch.nhs.uk> “Haemophilia A genetic testing.” 4 December 2008. Personal e-mail.

Green, P. <peter.green@genetics.kcl.ac.uk> “Haemophilia A information.” 7 October 2004. Personal e-mail.

HAMSTeRS (Haemophilia A Mutation, Structure, Test and Resource Site) database. 2003. “Review” <http://europium.csc.mrc.ac.uk/WebPages/Database/Review/Review_txt.html>

HAMSTeRS (Haemophilia A Mutation, Structure, Test and Resource Site) database. 2004. <http://europium.csc.mrc.ac.uk/WebPages/Main/main.htm>

HAMSTeRS (Haemophilia A Mutation, Structure, Test and Resource Site) database. 2007. “Mutation Database” <http://europium.csc.mrc.ac.uk/WebPages/Database/Quiz/intrquiz.htm>

Johnson, M.J. & Thompson, A.R. 2005. “GeneReviews” <http://www.genetests.com/>.

Keeney, S., Mitchell, M. & Goodeve, A. 2003. “Practice guidelines for the molecular diagnosis of haemophilia A”. <http://cmgsweb.shared.hosting.zen.co.uk/BPGs/pdfs%20current%20bpgs/Haemophilia%20A%20new.pdf>

van Gijn, M.E. <m.e.vangijn@umcutrect.nl> “Haemophilia A genetic testing.” 4 December 2008. Personal e-mail.

Weiss, L.M. 2008. “Assay Summary, Factor VIII Gene Mutation Analysis, Hemophilia A”. <http://mdl.cityof hope.org/pdfs/HpA_AssaySum.pdf>

World Federation of Haemophilia. 2005. “Guidelines for the Management of Haemophilia. Montreal, World Federation of Haemophilia”. <http://wfh.org/2/docs/Publications/Diagnosis_and_Treatment/Guidleines_Mng_Haemophilia.pdf>