THE MOLECULAR GENETICS OF FABRY DISEASE

This thesis is submitted in fulfilment of
conditions governing candidates for the degree of

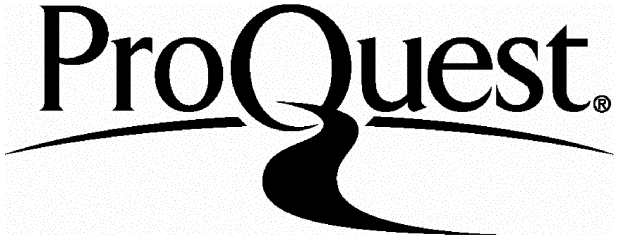
Doctor of Philosophy
in the
Faculty of Life Sciences
of the
University of London

by

Joanna Pauline Davies

Department of Biochemistry and Genetics
Institute of Child Health
30 Guilford Street
London
WC1N 1EH

April 1995



ABSTRACT

Anderson-Fabry disease is the lysosomal storage disorder resulting from a deficiency of α–galactosidase A (E.C. 3.2.1.22.), which cleaves α-linked galactose residues from the non-reducing ends of oligosaccharide chains. Loss of activity leads to the progressive accumulation of partially catabolised glycosphingolipids in the lysosomes of all cells but particularly those of the vascular endothelium. The α–galactosidase A gene, which is located at Xq22, spans 12kb and has 7 exons. The disease affects an estimated 1/40 000 and there is no cure. Although hemizygotes can be diagnosed reliably by demonstrating a deficiency of α-galactosidase A, heterozygote detection is unreliable due to mosaicism in the pattern of α-galactosidase A expression, caused by random inactivation of the X-chromosome.

The aims of this study were to develop methods for mutation detection, which would allow conclusive identification of heterozygotes and a study of the relationship between the genotype and the phenotype. Mutation analysis by a combination of single strand conformation polymorphism (SSCP) analysis and sequencing has detected 6 polymorphisms in the non-coding regions of the α-galactosidase A gene and 26 different, putative disease-causing mutations in 28 out of 30 unrelated Fabry patients. The polymorphisms were present in 7/61 unaffected males and in 4/30 Fabry patients. The mutations consisted of 16 missense, 5 nonsense, one splice site, 3 small deletion and 1 insertion mutation, 25 of which had not been identified previously. Two of these mutations were not detected by SSCP analysis but were found by sequencing. The mutations in 2 families were not identified by these methods.

The effects of the mutations on the α-galactosidase A activity were investigated by studying the residual activity in cultured cells from the patients and by the transient expression of mutated cDNA in COS-1 cells. Only 1 out of the 9 cell cultures had detectable activity, from a patient with the N215S mutation and a presumed, mild phenotype. Six missense mutations found in this study and a further 3, found by others, were transiently expressed from cDNA in COS-1 cells. No α-galactosidase A activity was produced by 5 mutant constructs, R49L, R112H, V269A, V316E and Q327K, while 4, G35R, Q279E, R301Q and G361R, produced low activity, compared with the wild-type cDNA. Mutations that produced detectable intracellular α-galactosidase A activity, G35R, Q269E and R301Q, were found in patients with variant phenotypes.

The identification of 26 different mutations in 28 families and of 6 polymorphisms within the α–galactosidase A gene has allowed accurate carrier detection and a study of the relationship between genotype and phenotype. This will improve patient management and indicate those most likely to benefit from enzyme therapy in the future.

PREFACE

Parts of the work described in this thesis are described in the following publications.

1. Davies JP, Winchester BG, Malcolm S (1993). Mutation analysis in patients with the typical form of Anderson-Fabry disease. *Hum. Mol. Genet.* 2:1051-1053.

2. Davies JP, Winchester BG, Malcolm S (1993). Sequence variations in the first exon of α-galactosidase A. *J. Med. Genet.* 30:658-663.

3. Davies JP, Winchester BG, Malcolm S (1993). Abstract for the CGS meeting, March 1993, London.

4. Davies J, Malcolm S, Winchester B (1993) Mutations for Fabry disease and polymorphisms in the α-galactosidase A gene. Abstract for the Second International Duodecim Symposium on Molecular Biology of lysosomal diseases, Finland: p105.

5. Davies J, Malcolm S, Winchester B (1993) Mutations for Fabry disease and polymorphisms in the α–galactosidase A gene. Abstract for the Ninth European Study Group on Lysosomal Diseases (ESGLD) Workshop, Delphi: p113.

6. Davies J, Christomanou H, Winchester B, Malcolm S (1994). Detection of 8 new mutations in the α-galactosidase A gene in Fabry disease. *Hum. Mol. Genet.* 3:667-669.

7. Davies JP, Eng, CM, MacDermot, K., Malcolm S, Winchester BW, Desnick, RJ (1995). Eleven mutations in the alpha-galactosidase A gene causing Fabry disease, manuscript in preparation.

ACKNOWLEDGEMENTS

Many thanks to my supervisors, Dr Bryan Winchester and Dr Sue Malcolm, to whom I am indebted for their help and encouragement throughout this study. Thanks also to the many people in the Biochemistry and Genetics labs. whose expertise and entertainment has helped me during the last three years. In particular, to Sandra Strautnieks, for many tours of the West End, in search of food, drink and solace. Also, to Paul Rutland, for his speed in synthesising the majority of primers used in this study and for solving many computer and technical problems, without complaint! I am grateful to Mrs Elisabeth Young, in the Enzyme Laboratory, for helpful discussions and for providing all specific α-galactosidase activity data from leukocyte and plasma measurements in Fabry families and normals.

I am also indebted to Dr R.J. Desnick, whose invitation to visit his laboratories at Mount Sinai Medical Center in New York, for 8 months, proved a most stimulating and enjoyable experience. My thanks go to everybody in these labs., in particular, to Dr Christine Eng, for welcoming me to New York, for encouragement and for introducing me to such a nice group of people in her lab.; to Dr Yiannis Iounnou, for putting up with me in his lab. for 4 months and whose enthusiasm was inspirational; and to Dana Niehaus, Dr Ken Astrin, Dr Lois Resnick-Silverman and Ken Zeidner for communicating their technical expertise. Finally, to Teresa Licholai, Robyn and Phil Cotter, Shai and Tamar Erlich and numerous others, for their friendship during my stay in New York.

I am very grateful to the following clinicians who provided patient samples and clinical information during this study. To Dr K. MacDermot, Royal Free Hospital, London; Prof. P.S. Friedmann, University of Liverpool; Dr T. Lesley, Middlesex Hospital; Prof. J.M. Connor, Duncan Guthrie Institute, Glasgow; Dr N.R. Dennis, Princess Ann Hospital, Southampton; Dr D. Robinson, Salisbury District Hospital; Prof. R. Harris, St. Mary's Hospital, Manchester; Prof. T. Cox, Addenbrooke's Hospital, Cambridge; Dr Delaney, Tewkesbury Hospital, Gloscestershire; Dr R. Staughton, Queen Mary's Hospital, Roehampton; Prof. J. Leonard and Dr A. Vellody at Gt. Ormond St., London; Dr G. Gray, Birmingham Childrens' Hospital; Prof. Anita Harding, Inst. Neurology, London; Dr M.M. Black, St Thomas's Hospital, London; Dr E. Legius, Leuven, Belgium; Dr D.J.J. Halley and Prof. Lindhout,

Rotterdam, Netherlands; Dr W. Lissens, Brussels, Belgium; Dr M. Beck and Dr K. Harzer, Mainz, Germany; Prof. H. Christomanou, Athens, Greece; and Dr H. Mikelakakis, Athens.

Finally, I would like to thank the Wellcome Trust for funding this research.

CONTENTS

CHAPTER 1: INTRODUCTION

CHAPTER 2: MATERIALS AND METHODS

CHAPTER 3: RESULTS

CHAPTER 4: DISCUSSION

LIST OF FIGURES

Chapter 1: Introduction

Chapter 2: Materials and methods

Chapter 3: Results

Chapter 4: Discussion

Appendix

LIST OF TABLES

Chapter 1: Introduction

Chapter 2: Materials and methods

Chapter 3: Results

Chapter 4: Discussion

ABBREVIATIONS

a	Adenine
A	Adenine
A	Ampere
α-gal	α-Galactosidase
APS	Ammonium persulphate
β-gal	β-Galactosidase
bp	Base pair
BMT	Bone marrow transplant
BSA	Bovine serum albumin
CDI	Carbodiimide
c	Cytosine
C	Cytosine
°C	Degrees centigrade
$CaCl_2$	Calcium chloride
cDNA	Complementary deoxyribonucleic acid
CHO cells	Chinese hamster ovary cells
Ci	Curie
cm	Centimetre
CO_2	Carbon dioxide
COS-1 cells	African green monkey CV-1 cells transformed with origin-defective mutant SV40 virus
CpG	Cytosine-guanine dinucleotide
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytidine 5'-triphosphate
ddATP	2'3'-Dideoxyadenosine 5'-triphosphate
ddCTP	2'3'-Dideoxycytidine 5'-triphosphate
ddGTP	2'3'-Dideoxyguanosine 5'-triphosphate
ddH_2O	Double distilled water
ddNTP	2'3'-Dideoxynucleoside 5'-triphosphate
ddTTP	2'3'-Dideoxythymidine 5'-triphosphate
DGGE	Denaturing gradient gel electrophoresis
dGTP	2'-Deoxyadenosine 5'-triphosphate
dH_2O	Distilled water
DMEM	Dulbecco's modified Eagle's medium

DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
dNTP	2'-Deoxynucleoside 5'-triphosphate
DSCA	Double-stranded conformation analysis
DTT	Dithiothreitol
dTTP	2'-Deoxythymidine 5'-triphosphate
EDTA	Ethylenediamine tetra-acetic acid
ER	Endoplasmic reticulum
fmol	Femtomole
Fuc	Fucose
g	Gram
g	Acceleration due to gravity
g	Guanine
G	Guanine
Gal	Galactose
GalNAc	N-acetylgalactosamine
Glc	Glucose
GlcNAc	N-acetylglucosamine
HCl	Hydrochloric acid
hr	Hour
IMS	Industrial methylated spirits
IPTG	Isopropylthiogalactoside
kb	Kilobase
KCl	Potassium Chloride
kDa	Kilodaltons
KH_2PO_4	Potassium dihydrogen phosphate
K_m	Michaelis Menten constant
l	Litre
LB	Luria-Bertani medium
LSD	Lysosomal storage disease
M	Molar (moles/litre)
Man-6-P	Mannose-6-phosphate
mg	Millligram
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulphate
ml	Millilitres
mm	Millimetre
mM	Millimolar

mmol	Millimole
mol	Mole
MOPS	3-N-morpholino-propane sulphonic acid
MPR	Mannose 6-phosphate receptor
mRNA	Messenger ribonucleic acid
4-MU	4-methylumbelliferyl
4-MU-α-gal	4-methylumbelliferyl-α-D-galactopyranoside
NaCl	Sodium chloride
NaH_2PO_4	Sodium dihydrogen phosphate
Na_2HPO_4	Disodium hydrogen phosphate
NaOH	Sodium hydroxide
ng	Nanogram
$(NH_4)_2SO_4$	Diammonium sulphate
nm	Nanometer
nmol	Nanomole
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PMSF	Phenyl methyl sulphonyl fluoride
pmol	Picomole
PEG	Poly ethylene-glycol
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
RPMI 1640	Roswell Park Memorial Institute formula 1640
SDS	Sodium dodecyl sulphate
SSCP	Single strand conformation polymorphism
t	Thymine
T	Thymine
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TE	Tris-EDTA buffer
TEMED	(N, N, N',N')-tetramethylethylenediamine
Tm	Melting temperature
Tris base	Tris (hydroxymethyl) aminomethane
U	Uracil
UDP	Uridine diphosphate

UV	Ultraviolet
V	Volt
V_m	Maximum enzyme velocity
v/v	Volume for volume
W	Watt
w/v	Weight for volume
X-Gal	5-bromo-4-chloro-3-indolyl β-D-galactoside
µM	Micromolar
µg	Microgram
µl	Microlitre

Addendum

CRIM	Cross-reacting immunologic material

1. INTRODUCTION

1.1 THE LYSOSOMAL SYSTEM

The anabolic and catabolic processes within the cell are in a dynamic state of flux and involve a vast number of enzymes and accessory proteins that act in a coordinated manner. All of these proteins are encoded by the genome and a defect in a single protein can cause severe pathological effects if there is not a compensatory mechanism to maintain the normal metabolic processes. Since the primary defect is in the gene sequence it can be inherited.

Enzyme-catalysed processes occur both extracellularly and intracellularly. Within cells enzymes are distributed in specific compartments or organelles. This concentrates the enzymes that are involved in a particular function into a small, membrane-bound space and serves several purposes. It allows the amount and types of substrate which are acted upon by these proteins to be controlled and it increases the reaction efficiencies by bringing the substrate and enzymes into closer proximity. The organelle membranes allow ion gradients to be established to facilitate metabolic reactions. One such organelle is the lysosome, whose primary function is the catabolism of macromolecules. It contains a characteristic set of enzymes and accessory proteins which collectively have the capacity to catalyse the breakdown of most naturally occurring macromolecules to their constituent monomeric components. These can then pass through the lysosomal membrane and be reutilised by the cell. A defect in a particular lysosomal protein will create a metabolic block and the resultant partially degraded substrates will progressively accumulate in the lysosomes because they cannot pass through the lysosomal membrane. This results in the pathological condition known as a lysosomal storage disease.

Lysosomes are found in all cell types in humans, except erythrocytes. They were first identified by ultracentrifugation, as vesicles containing latent acid phosphatase activity (de Duve *et al.*, 1955). They were subsequently shown to be identical to particles that stained positive for acid phosphatase, which had been detected by light and electron microscopy (Novikoff *et al.*, 1956). Since their discovery, extensive research has revealed much of the cell biological and biochemical processes involved in the formation and function of the lysosome (Holtzmann, 1989; Kornfeld, S *et al.*, 1989; Winchester, 1992). From these studies

it has become clear that the lysosomal fraction is heterogeneous, reflecting the different cellular functions of the lysosomal system. This heterogeneity has complicated the definition of the lysosome. Terms such as multivesicular bodies, autophagosomes, primary lysosomes and secondary lysosomes have been used to describe the morphologically different forms of the lysosome. However, lysosomes can be defined more precisely by their biochemical description as organelles with an intraluminal pH of 4.5-5.5 that contain lysosomal hydrolases and characteristic membrane proteins. In addition, unlike the acidic, pre-lysosomal compartments, they do not contain mannose 6-phosphate receptors. Lysosomes have also been defined functionally as the organelles that are the final destination for soluble, endocytosed proteins. For many macromolecules, the mature lysosome is the metabolic endpoint from which their constituents can be reutilised in further anabolic processes or excreted. However, for some compounds, such as cholesterol (Goldstein *et al.*, 1977) and vitamin B_{12} (Idriss *et al.*, 1991), the lysosome plays an essential role in their processing for metabolic use. Knowledge of the lysosomal system, its catalytic activities and transport systems, is essential in understanding the pathology of lysosomal storage diseases.

1.1.1 Transport of substrates to the lysosome by autophagy and heterophagy

Materials for degradation reach the lysosome by either autophagy, for the digestion of materials from within the cell, or heterophagy, for the digestion of exogenous materials (Holtzmann, 1989; Winchester, 1992).

Autophagy occurs continuously, as observed from histological studies and three non-selective mechanisms have been described (Seglen *et al.*, 1992; Dice, 1992). In microphagy the lysosomal membrane invaginates to internalise cytoplasmic constituents directly, while in macroautophagy whole organelles and large portions of the surrounding cytoplasm are engulfed to form membrane-bound vacuoles called autophagosomes. Subsequently, their contents are degraded in the lysosome. In crinophagy, mature secretory granules fuse with lysosomes or pre-lysosomal structures for degradation. A fourth process for degradation of cytoplasmic proteins is not included in the invagination processes of autophagy and is specific for certain proteins that possess a specific peptide sequence (Lys-Phe-Glu-Arg-Gln, or KFERQ). The mechanism involves direct transport though the lipid bilayer of the lysosome, aided by a chaperone, a member of the Hsp70 family of heat shock proteins (Chiang *et al.*, 1989).

Heterophagy is more selective than autophagy and multiple processes are involved in transport across the plasma membrane and into the lysosomes. Materials are internalised into the cell by endocytosis, an invagination of the plasma membrane. Two endocytic mechanisms operate. Phagocytosis is involved in the uptake of large solid materials such as cell debris and bacteria, leading to the formation of phagosomes, vesicles of 1-2μm in diameter, while pinocytosis refers to the uptake of fluids, forming smaller vesicles of ~0.1μm in diameter. Phagocytosis is probably not constitutive and may involve specific receptors that on binding ligand induce the internalisation of membrane to form the phagosome. The phagosome fuses or is engulfed by the lysosome for destruction and this mechanism of heterophagy is commonly used by macrophages for digesting foreign materials. Pinocytosis does occur constitutively and can be a non-specific uptake mechanism, forming fluid phase pinocytosis vesicles which ultimately become part of the lysosome. Pinocytosis can also be selective when uptake is mediated by receptors clustered in clathrin-coated pits. Coated vesicles are formed and the receptors are often recycled whilst their ligand is transported, via endosomes, to the lysosome. Receptors have been found that bind specific carbohydrate moieties including mannose 6-phosphate, low density lipoprotein (LDL), α_2-macroglobulin, EGF, insulin, the constant region (F_c) of IgG antibodies and human chorionic gonadotrophin (HCG) (Winchester, 1992).

Therefore, there is a continuous flow of extracellular and cytoplasmic material into the lysosomes, by both specific and non-specific mechanisms. The details of these mechanisms are not fully understood.

1.1.2 The lysosomal membrane

The lysosomal membrane is a typical phospholipid bilayer membrane, although its protein content is different from that of other membranes (Burnside *et al.*, 1982). Its major function is to compartmentalise the catabolic processes, thereby limiting degradation of macromolecules to those that are transported into the lysosomes and protecting the cell against destruction. Specific systems for transport of small molecules across the membrane are known, including those for certain monosaccharides, vitamin B_{12} (cobalamin), calcium, folate derivatives, phosphate, sulphate, chloride, nucleosides and amino acids (Pisoni *et al.*, 1991; Chou *et al.*,

1992). A Mg^{2+}-ATPase proton pump resides in the lysosomal membrane and this is responsible for maintaining the acidic interior of the lysosome (Arai *et al.*, 1993).

A set of proteins, known as lysosome-associated membrane proteins (LAMPS), lysosome glycoproteins (lgps) or lysosome integral membrane proteins (LIMPS) are found predominantly in the membranes of lysosomes. At least four of these, lgpA, lgpB, LIMP1 and LIMP II are known to possess specific amino acid sequences that are required for their transport to the lysosome (Williams *et al.*, 1990; Harter *et al.*, 1992; Ogata *et al.*, 1994) and they are either transported via the plasma membrane or directly from the Golgi to the lysosome (Honing *et al.*, 1995). They are not known to perform catalytic activity and their proposed function is to protect the membrane from autodigestion (Fukada, M, 1991; Peters *et al.*, 1994).

1.1.3 Lysosomal Hydrolases

The interior of the lysosome holds about 50 different glycosylated hydrolytic enzymes whose pH optima range from 3.5-5.5 (Barrett *et al.*, 1977; Winchester, 1992). These enzymes can be grouped into several classes: proteases, nucleases, lipases, glycosidases, sulphatases and phosphatases and, collectively, they can degrade most macromolecules. Most of the glycosidases are exohydrolases and so degradation of a particular carbohydrate chain requires multiple enzymes to act in a specific order for its breakdown into monosaccharide units. Each enzyme may operate in several pathways for the degradation of different oligosaccharide chains, causing cleavage of a particular glycosidic bond.

Deficiencies of most of the lysosomal enzymes have been shown to cause disease, indicating the importance of the lysosomal system. It also suggests that they do not have overlapping specificities, which could compensate for the insufficiency of individual enzymes. However, deficiencies of proteases and nucleases have not been found and this may reflect either overlapping specificities in these groups of lysosomal enzymes or it may indicate that the resulting phenotype is lethal in utero.

In addition to the hydrolases, a protective protein (Galjart *et al.*, 1988) and five sphingolipid activator proteins (Furst *et al.*, 1992; Kishimoto *et al.*, 1992) exist to aid in catabolism (table 1).

Accessory protein	Function	Enzymes activated or protected
Saposin A	activator	β-glucosylceramidase β-galactosylceramidase
Saposin B	activator	α-galactosidase A β-galactosidase β-glucosylceramidase arylsulphatase A sphingomyelinase
Saposin C	activator	β-glucosylceramidase β-galactosylceramidase sphingomyelinase
Saposin D	activator	sphingomyelinase
GM2-activator	activator	β-hexosaminidase A (a heterodimeric protein composed of an α and β chain)
Protective protein (carboxypeptidase L)	protector	β-galactosidase and sialidase form a complex

Table 1: Accessory proteins that enhance or are essential for lysosomal enzyme function

Four of the activator proteins are derived from a common precursor protein, prosaposin, which is encoded by a single genetic locus. Saposins A, B (SAP-1), C (SAP-2) and D are released from prosaposin by proteolysis in the lysosome. They are highly homologous and thought to be evolutionarily related. Defects in saposins B and C and prosaposin have been shown to cause lysosomal storage diseases.

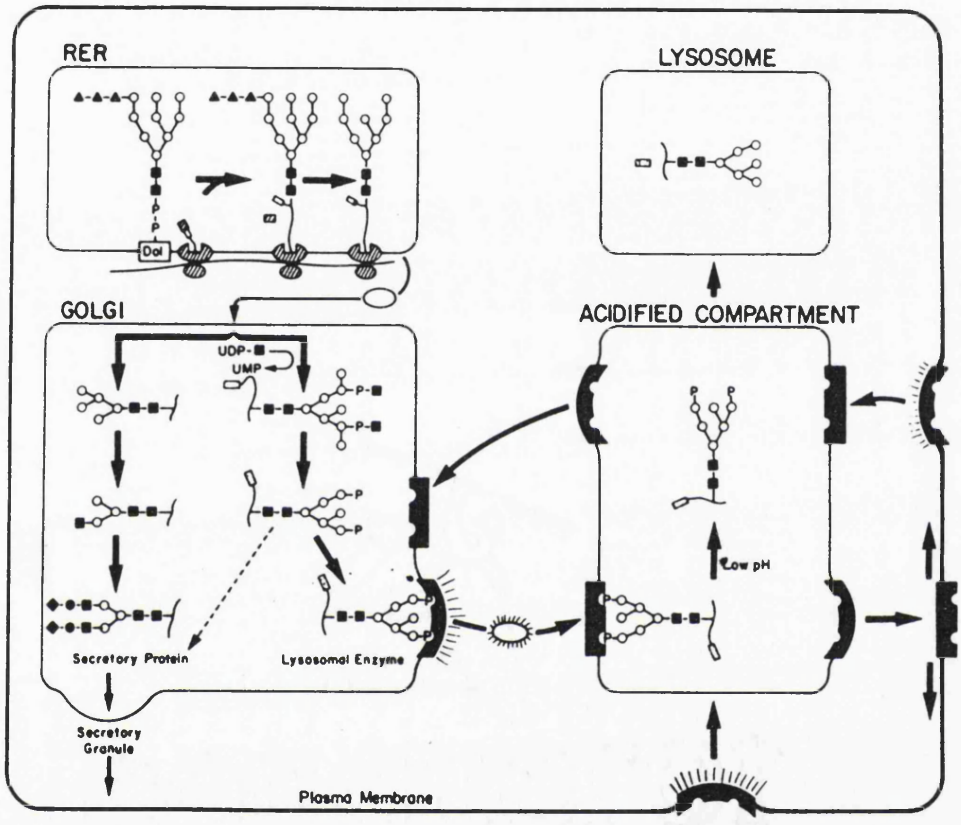
1.1.3.1 Biosynthesis and intracellular transport of lysosome proteins

1.1.3.1.1 Processing of lysosomal proteins in the ER and Golgi system

Lysosomal proteins are synthesised on ribosomes that are attached to the rough endoplasmic reticulum (ER) and their transport into the ER and Golgi has been studied by cell biological methods and microscopy, combined with immunodetection. The steps involved are indicated in figure 1 (Kornfeld, R *et al.*, 1985). As with all proteins processed in the ER, an N-terminal, hydrophobic signal peptide of about 20-30 amino acids is cotranslationally translocated across the ER membrane (Walter *et al.*, 1984; Simon *et al.*, 1991). The signal peptide is then cleaved in the ER and N-linked glycosylation is initiated. A pre-assembled oligosaccharide is transferred to some asparagine residues which are part of the recognition sequence Asn-X-Thr or Ser sites. 'X' is any amino acid, although sites containing proline or aspartic acid are rarely glycosylated (Aubert *et al.*, 1976). In the ER, disulphide bond formation, protein folding and oligomerisation occur and 3 glucose units and a mannose residue are removed from the oligosaccharide chains by specific ER glycosidases.

From the ER, the lysosomal enzyme precursors are transported to the Golgi compartments in vesicles for further processing. Proteins destined for lysosomes are phosphorylated at certain mannose residues and this acts as unique marker for transport to the lysosome. Phosphorylation occurs as a 2 step enzymic process in which GlcNAc-1-P is transferred from UDP-GLcNAc to the 6 position of mannose to form Man 6-P-GlcNAc and then GlcNAc is removed to form Man 6-P. The first step is thought to occur in a post ER/pre-Golgi compartment and the second in the cis Golgi (Lazzarino *et al.*, 1988, 1989; Kornfeld, S *et al.*, 1989).

Figure 1: Post-translational processing and transport of lysosomal enzymes



Filled triangles, squares and diamonds are glucose, N-acetylglucosamine and sialic acid residues, respectively. Open circles are mannose residues. 'P' are phosphate moieties and 'Dol' is the carbohydrate donor, dolichol (Kornfeld R, *et al.*, 1985).

The mechanism for recognising lysosomal enzymes for mannose 6-phosphorylation is unknown. No conserved, linear protein sequences have been implicated but mutation studies support the hypothesis that the location of amino acids in the tertiary structure, i.e. the local conformation, is important (Kornfeld, S, 1989; Kornfeld, S *et al.*, 1989; Baranski *et al.*, 1990). Modification of surface lysine residues on lysosomal proteins, by binding to sulpho-N-hydroxysuccinamide, inhibits phosphorylation of numerous lysosomal enzymes, indicating the importance of lysine in the recognition process (Cuozzo *et al.*, 1994).

N-linked glycans that do not acquire the Man-6-P marker can be further modified in the Golgi apparatus by glycosidases and glycosyltransferases that sequentially remove and add monosaccharide units. In addition, both the protein, at tyrosine residues (Huttner *et al.*, 1988) and the sugar components can be sulphated (Braulke *et al.*, 1987). The resultant oligosaccharides may be complex, high mannose or hybrid in structure (von Figura *et al.*, 1986) and depending on the cell type, the post-translational processing may modify a single protein heterogeneously, as seen for many lysosomal enzymes, for example, β-hexosaminidase (Sonderfeld-Fresco *et al.*, 1989).

In the terminal trans compartment of the Golgi, lysosomal proteins with the man-6-P marker bind to receptors specific for Man-6-P and are segregated from other glycoproteins which travel to the plasma membrane for incorporation or secretion (Kornfeld, R *et al.*, 1985; Kornfeld, S *et al.*, 1989).

1.1.3.1.2 Transport of lysosomal proteins with mannose 6-phosphate

The presence of a specific recognition marker for some lysosomal proteins was first discovered in 1977 (Kaplan *et al.*, 1977) based on the observation that in I-cell disease (mucolipidosis type II, Hasilik *et al.*, 1981; Ben-Yoseph *et al.*, 1986) phosphorylation at the mannose residues is defective and many lysosomal enzymes are mislocalised. Two specific mannose 6-phosphate receptors (MPR) have now been identified and cloned (Dahms *et al.*, 1989). One is a 46kDa Ca^{2+}-dependent receptor (CD-MPR46) and the other is a cation-independent receptor (CI-MPR300) of 250-300kDa, which also binds insulin-like growth factor 2 (IGF2), Kornfeld, S, 1992. The two receptors are found together in the pre-lysosomal compartment, the plasma membrane and the Golgi apparatus of most cell types. The reason for

having two proteins with apparently the same function is unknown, as is the significance of the bifunctional IGF2/MPR300 protein.

The MPR-lysosomal protein complex is transported from the Golgi, probably by a vesicular transport mechanism, to form part of a pre-lysosomal or late endosomal compartment. The pH of this compartment is lower than that in the Golgi and causes dissociation of the complex. The receptor is recycled to the Golgi or to the plasma membrane, while the enzyme is channelled to the lysosomes.

Although the major fraction of lysosomal enzymes is transported directly to the lysosome by the MPR pathway, a proportion is secreted constitutively, in low amounts (Rosenfeld *et al.*, 1982; Zuhlsdorf *et al.*, 1983). If this protein is phosphorylated at mannose residues then it can be re-captured by binding to the MPRs in the plasma membrane of the same or an adjacent cell, followed by endocytosis and delivery to the lysosome.

1.1.3.1.3 Transport of lysosomal proteins independently of the MPRs

Evidence for the transport of lysosomal enzymes by mechanisms independent of the MPR transport system has come from the study of different cell types from I-cell patients, in which MPR transport cannot operate. In I-cell fibroblasts, multiple enzymes are deficient in the lysosomes due to their lack of transport, while in hepatocytes, leukocytes and Kupffer cells, the same enzymes reach nearly normal levels in the lysosomes, indicating the presence of alternative, cell-specific transport systems (Owada *et al.*, 1982; Kornfeld, S, 1986). In addition, in normal cells a few lysosomal enzymes, including acetyl Co A: α-glucosaminide N-acetyl-transferase, some proteinases (e.g. cathepsin M), β-glucocerebrosidase (Willemsen *et al.*, 1987; Aerts *et al.*, 1988) and lysosomal acid phosphatase (Gottschalk *et al.*, 1989) are not phosphorylated and transported by the MPRs. These proteins are all membrane-bound or membrane-associated and are transported by a different, unknown mechanism. In the case of lysosomal acid phosphatase, its membrane-bound precursor has been shown, by mutagenesis studies, to possess a peptide sequence that is required for targeting to the lysosome (Krentler *et al.*, 1986; Williams *et al.*, 1990; Harter *et al.*, 1992). Once it reaches the lysosome, its precursor is subjected to two steps of proteolytic cleavage to release the active, soluble phosphatase into the lysosomal lumen (Gottschalk *et al.*, 1989).

Secreted lysosomal enzymes may also be transported to the lysosomes by endocytosis (Mellman, 1987), perhaps mediated by receptors for galactose, mannose and N-acetylglucosamine (Hasilik, 1992), on glycoproteins. Other, uncharacterised endocytic mechanisms may also exist (Koster *et al.*, 1994).

1.1.3.2 Post-Golgi processing of lysosomal enzymes

Once the lysosomal proteins reach the pre-lysosomal vesicles and the lysosome they are often processed further (Hasilik, 1992). The oligosaccharides may be trimmed, as with the α-chain of β-hexosaminidase, which is modified both in the Golgi apparatus and during lysosome biosynthesis, to produce heterogeneously glycosylated enzyme (Little *et al.*, 1988). Some lysosomal enzymes are subjected to partial proteolysis after transport from the Golgi, for example, iduronate 2-sulphatase (Bielicki *et al.*, 1990; Wilson *et al.*, 1990), lysosomal acid GM2-activator and β-hexosaminidase (Hubbes *et al.*, 1989; Mahuran, 1990). This generates increased heterogeneity in the number of forms that exist for a single enzyme but for most of the lysosomal enzymes, it is not necessary for generation of enzymic activity. The activity of a few proteins is dependent on cleavage. The proteases, cathepsins D, L, B and H are believed to autocatalyse their own peptide fragmentation (Hasilik, 1992). The prosaposin polypeptide also requires proteolytic cleavage to form saposins A, B, C and D (Furst *et al.*, 1992), section 1.1.3.

1.2 LYSOSOMAL STORAGE DISEASES

Lysosomal storage diseases (LSDs) were first recognised in 1965 (Hers, 1965). The biochemical defect in Pompe's disease was identified as a deficiency of a single enzyme, of α-glucosidase. This caused a metabolic block, resulting in the progressive accumulation of undegraded glycogen within the lysosomes and an increase in lysosome size and number. Subsequently, many other clinically described diseases were recognised as LSDs, in which either partially or fully catabolised substances accumulated in the lysosome. Most were caused by a primary defect in a specific lysosomal hydrolase, table 2 (Neufeld, 1991). However, in a few disorders, the deficient catabolic activity was secondary to a defect in another protein. For example, in I-cell disease the lack of mannose 6-phosphorylation caused mislocalisation of multiple enzymes in fibroblasts. Lack of an accessory protein cofactor, such as saposin B or the protective protein for β-galactosidase and sialidase also leads to enzymic deficiencies. Transport defects result in abnormal storage of fully catabolised in the lysosomes, as in cystinosis, in which cobalamin transport is deficient. In one LSD, sulphatidosis, in which there is a deficiency of several sulphatases, the underlying defect is unknown.

LSDs account for 15% of known inherited metabolic diseases. All are inherited as autosomal recessive disorders, except for Fabry and Hunter diseases which are X-linked. Historically, LSDs were classified by the predominant type of enzyme substrates that accumulated, i.e. sphingolipids, glycoproteins or mucopolysaccharides. The clinical pathology is dependent on the type, cellular location and rate of transport of the storage products to the lysosome. Now the primary protein defects for all but four LSDs are known. The observed enzyme deficiencies in sulphatidosis and mucolipidosis IV are secondary to an unknown defect, as is the defective cholesterol esterification observed in Niemann-Pick C. In a fourth disorder, Batten disease, subunit c of the mitochondrial ATP synthase complex accumulates in the lysosomes and is believed to occur due to defective degradation of subunit c (Ezaki *et al.*, 1995). Since the protein defects are known for most LSDs, they can be classified more precisely as hydrolase defects or defects in membrane transport, lysosome targeting, substrate activators or protective proteins, table 2 (Scriver *et al.*, 1989; Neufeld, 1991).

Table 2: The lysosomal storage disorders and their protein defects

Disease	Protein Deficiency	Gene size (No. Exons)	Reference, gene (or cDNA)
Sphingolipidoses			
Fabry	α-D-galactosidase A	12kb (7)	Kornreich, 1989
Farber	Ceramidase	-	-
Gaucher	β-glucocerebrosidase	7.6kb (11)	Horowitz, 1989
GM1 gangliosidosis	β-D-galactosidase	>62.5kb (16)	Morreau, 1991
Tay-Sachs (GM2 gangliosidosis)	β-hexosaminidase A, (α chain)	35-40kb (14)	Proia, 1987
Sandhoff (GM2 gangliosidosis)	β-hexosaminidase A,B (β chain)	45kb (14)	Neote, 1988
Krabbe	β-galactosylcerebrosidase	-	(Chen, 1993)
Metachromatic leukodystrophy	Arylsulphatase A	3.2kb (8)	Kreysing, 1990
Mucolipidosis IV	Ganglioside sialidase (secondary)	-	-
Niemann-Pick, types A and B	Sphingomyelinase	4.7kb (6)	Schuchman, 1992
Niemann-Pick, type C	(Cholesterol esterification)	-	-
Schindler	α-N-acetylgalactosaminidase	13.7kb (9)	Wang, 1991
Glycoproteinoses			
Aspartylglucosaminuria	N-aspartyl-β-glucosaminidase	13kb (9)	Park, 1991
Fucosidosis	α-L-fucosidase	23kb (8)	Kretz, 1992
α-Mannosidosis	α-D-mannosidase	-	(Nebes, 1994)
β-Mannosidosis	β-D-mannosidase	-	-
Sialidosis (Mucolipidosis I)	N-acetyl-α-neuraminidase (sialidase)	-	-
Mucopolysaccharidoses			
Hunter (MPS II)	Iduronate-2-sulphatase	24kb (9)	Flomen, 1993 Wilson, 1993
Hurler/Scheie (MPS I)	α-L-iduronidase	8kb (14)	Scott, 1992
Maroteaux-Lamy (MPS VI)	GalNAc-4-sulphatase (arylsulphatase B)	<24kb (8)	Modaressi, 1993
Morquio A (MPS IV, A)	N-GalNAc-6-sulphatase	40kb (14)	Morris, 1994
Morquio B (MPS IV, B)	β-D-galactosidase	<62.5kb (16)	Morreau, 1991
Sanfilippo, A (MPSIII, A)	Heparan N-sulphatase	-	-
Sanfilippo, B (MPSIII, B)	α-N-acetylglucosaminidase	-	-
Sanfilippo, C (MPSIII, C)	AcetylCoA: α-glucosaminide N-transferase	-	-
Sanfilippo, D (MPSIII, D)	GlcNAc 6-sulphatase	-	(Robertson, 1988)
Sly (MPS VII)	β-glucuronidase	21kb (12)	Miller, 1990; Shipley, 1991
Accessory protein disorders			
Galactosialidosis	Protective protein for β-galactosidase & sialidase	-	(Galjart, 1988)
GM2 gangliosidosis (AB variant)	GM2 activator (for β-hexosaminidase A)	15kb (3)	Klima, 1991
Sphingolipid activator disorders (multiple enzymes affected)	Deficiency of one or more saposins, A, B, C or D	20kb (13)	Rorman, 1992
Other storage disorders			
Pompe (glycogenosis II)	α-glucosidase	20kb (20)	Hoefsloot, 1990
Wolman disease	Acid lipase	36kb (10)	Anderson, 1994
Sulphatidosis	Multiple sulphatases	-	-
Batten (Ceroid lipofuscinosis)	Unknown	-	-
Membrane transport defects			
Cystinosis	Cystine transporter	-	-
Sialic acid storage (eg. Salla)	Sialic acid transporter	-	-
Methylmalonic aciduria variant	Vit B12 (cobalamin) transport	-	-
Lysosome targeting defect			
I cell disease (Mucolipidosis II) & Pseudo-Hurler (Mucolipidosis III)	N-acetylglucosaminyl-phosphotransferase	-	-

There is phenotypic heterogeneity in all the disorders. This can result from the existence of residual enzyme activity towards one or more types of substrate and therefore differences in the amount and sometimes the type of accumulated substrate. For example, a deficiency of β-galactosidase can cause two clinically distinct diseases. GM1 gangliosidosis occurs in those with negligible activity for gangliosides whist in Morquio B (MPS IV, B), the defect only seems to affect the breakdown of glycosaminoglycans.

The biochemical and clinical heterogeneity of the LSDs can, in part, be attributed to heterogeneity of the underlying genetic defects. Many of the genes and/ or cDNA nucleotide sequences for lysosomal proteins are known (table 2) and the genetic defects in many patients have been identified. The value of this information is apparent in studies of patients with Gaucher disease. Like other LSDs, Gaucher disease is clinically variable in severity and is sub-classified into 3 groups, the milder type 1 non-neuronopathic form and the more severe types 2 and 3, which are neuronopathic. Mutation analysis has identified four common mutations and genotype-phenotype correlation has shown that all patients homozygous for the N370S mutation had the mildest, type 1 form of Gaucher disease. Three other mutations, L444P, an insertion mutation 84GG and a splice site mutation IVS2(+1) were found predominantly in patients with the more severe types 2 and 3 (Grabowski, 1993). This information is important in family counselling and has allowed efficient screening programmes to be developed for carrier detection. Knowledge of the ethnic origin of patients is also important in some diseases. In the Ashkenazi-Jewish population, Gaucher disease affects about 1:450, 94-97% of whom have one of the four common mutations. Detection of these same four mutations in non-Jewish populations identifies fewer carriers. Another example is aspartylglucosaminuria, which is relatively rare, except in the Finnish population. Its incidence in Finland is an estimated 1:3600 and the C163S mutation is found in 98% of these patients (Mononen *et al.*, 1993). Therefore, knowledge of the ethnic origin, the presence of common mutations and the severity of symptoms associated with specific mutations are of value for carrier screening and counselling.

There is no cure for the majority of lysosomal storage diseases and treatment is palliative. The progression of disease has been successfully reduced in two disorders, cystinosis, which is treated using cysteamine and type 1 Gaucher disease, treated by enzyme replacement. Cysteamine reacts with the storage product, cystine, to form a mixed disulphide and cysteine, which can then be

transported out of the lysosomes, thus preventing accumulation (Markello *et al.*, 1993). Patients with non-neuronopathic type 1 Gaucher disease have been successfully treated with a modified form of the deficient enzyme, glucocerebrosidase (Barton *et al.*, 1991; Barranger *et al.*, 1993). The commercially available enzyme, ceredase, is purified from placenta and the oligosaccharide side-chains of the enzyme are degraded to expose terminal mannose residues. This form of the enzyme more efficiently endocytosed by macrophages, the major site for substrate degradation, than the native enzyme (Furbish *et al.*, 1981). Recently, a recombinant form of glucocerebrosidase, secreted from CHO cells, has proved useful in treatment (Grabowski *et al.*, 1995). Thus, knowledge of the primary defect, the properties of the enzyme or protein, the sites of storage product accumulation and methods for targeting exogenous, therapeutic substances, have enabled the development of effective therapy in these two LSDs. Characterisation of the residual activity and cross-reacting material in cells from patients with LSDs and expression of mutated cDNAs in mammalian cells have provided some insight into the effect of mutations on the activity, stability, processing and transport of some enzymes (e.g. Weitz *et al.*, 1992; Oshima *et al.*, 1994). Future therapeutic strategies may be influenced by information gained from these types of studies.

Therapy in LSDs other than Gaucher and cystinosis has been attempted by a variety of methods. Most recent studies have been aimed at replacing the deficient enzyme. Enzyme therapy may become available for several lysosomal diseases, in addition to Gaucher disease and will be facilitated by the development of methods for large scale production of recombinant enzymes. A major disadvantage of this strategy is that it relies upon repeat administration of enzyme. In some LSDs, such as Hurler (Hopwood *et al.*, 1993b), MPS VI (Krivit *et al.*, 1984) and others (Hobbs, 1990), bone marrow transplants (BMT) have provided a continuous source of enzyme and partial metabolic correction in human and animal models, with enzyme activity present in numerous tissues, including the neurons and glia of the central nervous system (Walkley *et al.*, 1994). The success of BMT is limited by the availability of histocompatible tissue and has a high mortality rate due to infection and organ rejection. This problem may be circumvented by gene therapy. Although gene therapy for LSDs has not yet been tested in humans, preliminary studies are in progress. Retroviral vectors have been used to introduce a corrective cDNA sequence into cultured cells with an enzyme deficiency, for example, in lymphoblastoid cells from a Hunter patient (Braun *et al.*, 1993) and fibroblasts from a patient with metachromatic leukodystrophy (Rommerskirch *et al.*, 1991), resulting in expression of the deficient enzyme. Recently, autologous bone marrow

from the murine model of Gaucher was transfected with a corrective retroviral vector and these cells were transplanted back into the mice. The modified cells successfully infiltrated the liver, lung, brain, spinal cord and macrophages, producing recombinant enzyme (Krall *et al.*, 1994). This approach may allow treatment of the neuronopathic forms of Gaucher disease, currently untreatable by enzyme therapy and has been approved for clinical trial. In another study, β-glucuronidase cDNA was transferred into MPS VII, murine fibroblasts, which were subsequently embedded into a matrix of rat collagen and polytetra-fluoroethylene fibres. This neo-organ matrix was transplanted back into the peritoneal cavity of the MPS VII mouse , producing complete metabolic correction in the liver and spleen (Moullier *et al.*, 1993). Treatment of numerous LSDs by these methods is imminent. The extent of success will rely upon the ability of the enzyme to reach its target, its catalytic activity and stability. The patient must also be immunologically tolerant to the enzyme (Hobbs, 1990) and since patients with a absence of protein, such as those with a complete deletion, may not be tolerant it is important to establish the nature and effects of mutations in patients requiring treatment.

At the start of this study, genetic information regarding the LSD, Fabry disease, was limited. Although the gene had been cloned, specific defects in patients had been determined in only 13 cases. No common mutations had been identified and genetic methods for carrier detection had not been developed. There was no information regarding the correlation of genotype with phenotype. Since there is no effective therapy for this disorder, carrier detection and counselling is important. Therefore a study of the molecular genetics of Fabry disease was undertaken to develop methods for the detection of mutations and carriers.

1.3 FABRY DISEASE

1.3.1 Introduction

Fabry disease, also known as Anderson-Fabry disease or angiokeratoma corporis diffusum (McKusick No. 30150) is an inherited lysosomal storage disease which is estimated to affect 1 in 40,000 panethnically. It is caused by a deficiency of α-galactosidase A (E.C. 3.2.1.22), a lysosomal hydrolase, which catalyses the removal of terminal α-galactosyl residues from the non-reducing ends of neutral glycosphingolipids. It results in the progressive accumulation of glycosphingolipids, terminated in α-linked galactose, in the lysosomes of many body tissues and fluids but predominantly in the vascular endothelium. Fabry disease is X-chromosome-linked, consistent with the localisation of the α-galactosidase A gene at Xq22 (Desnick *et al.*, 1989).

1.3.2 Clinical Aspects

1.3.2.1 Symptoms in hemizygous males

Anderson-Fabry disease was first described clinically in male patients in 1898 by the British surgeon, Anderson (1898), and independently in the same year by the German physician, Fabry (1898). Since their early clinical observations, the description has been extended to include the following major diagnostic symptoms, which are present to varying degrees of severity in different patients (Desnick *et al.*, 1989; Morgan *et al.*, 1990; Radcliffe *et al.*, 1990).

1. Angiokeratomas are skin lesions that are dark-red to blue in colour. They are often observed in clusters between the umbilicus and knees, exhibiting bilateral symmetry and are also found in the oral mucosa and conjunctiva.
2. Peripheral (autonomic) nerve system manifestations include burning pains in the hands and feet. The pain may be either acute, episodic crises, usually induced by exercise, stress and temperature changes, or constant discomfort, acroparesthesias.
3. Ocular features, such as corneal dystrophy, conjuctival and retinal vascular lesions are seen.
4. Hypohidrosis, deficient perspiration.

5. Cardiovascular deposition of glycosphingolipids can result in electrocardiographic (ECG) abnormalities, angina, myocardial ischaemia and heart failure.
6. Cerebrovascular involvement which can cause neurologic symptoms. Transient ischaemic attacks, hemiplegia, haemorrhagic lesions and thrombosis may occur. Personality and behavioural changes, confusion and disorientation can also occur and may be caused by an electrolyte imbalance as a result of either renal insufficiency or cerebral damage.
7. Kidney complications include proteinuria and progressively impaired renal function.

Other clinical features of the disease are found in some Fabry patients. These include gastrointestinal malfunction (Rowe *et al.*, 1974), anaemia, due to decreased erythrocyte survival (Krivit *et al.*, 1968), lymphoedema (Gemignani *et al.*, 1979), retarded growth and late onset of puberty (Desnick *et al.*, 1989).

In patients with the typical clinical picture, symptoms begin during childhood or teenage years, with angiokeratoma, acroparesthesia and ocular features. Cerebrovascular, cardiovascular and renal malfunction occur at a later stage and may cause death, which occurs at an average age of 41 years (Columbi *et al.*, 1967), if untreated. In this group of patients, the residual activity of α-galactosidase is negligible.

Eighteen atypical hemizygotes with mild or limited sets of symptoms have been described. Their residual α-galactosidase A activity ranged from 1-37% of the normal, where measured (table 3). Of these, one had no symptoms, one had retarded growth and skin dysplasia, another had acroparesthesias only, five had proteinuria only, nine had cardiac involvement only and one was milder but no details were available.

Table 3: Phenotypes, enzyme activities and mutations in atypical hemizygotes

Age (yr)	Phenotype in hemizygote	Mutation	α-galactosidase A activity*		Other comments (references)
			Material	% normal	
-	Dutch patient with retarded growth and ectodermal dysplasia only.	P146S	Leukocytes	7% of highest normal value	(Ploos van Amstel *et al.*, 1994)
56	Persian patient with mild proteinuria only.	R112H	No enzyme assays.		Sibling was severely affected and died. (Eng *et al.*, 1994)
42	Italian patient with reumatoid arthritis and proteinuria only (renal involvement).	N215S	Fibroblasts Platelets Liver Lymphocytes Granulocytes Plasma Urine	37% 17% 17% 11% 9% 1% 1%	(Bishop *et al.*, 1981; Eng *et al.*, 1993)
63	Czech patient died, at age 63, of cardiac malfunction. No renal involvement or other symptoms.	N215S	No enzyme assays. Glycosphingolipid deposition found in heart only.		(Elleder *et al.*, 1990; Eng *et al.*, 1993)
63	German patient had cardiac involvement only.	N215S	No enzyme assays.		(Eng *et al.*, 1993)
63	Japanese patient died of cardiac complications aged 64. No other symptoms.	Q279E	Lymphoblasts Plasma	3.8% 3%	(Nagao *et al.*, 1991; Ishii *et al.*, 1992)
54	Japanese patient with cardiomyopathy as the only symptom.	M296V	Lymphoblasts Plasma Fibroblasts Lymphocytes Granulocytes	25.1% 12.5% 9.2% 4% 1.4%	(von Scheidt *et al.*, 1991)
52	Japanese patient with cardiomyopathy as the only symptom.	R301Q	Lymphoblasts Plasma	3.6% 1.6%	(Sakaruba *et al.*, 1990; Nagao *et al.*, 1991, Ishii *et al.*, 1992)
-	Danish, milder symptoms	R301Q	No enzyme assays		(Madsen *et al.*, 1995)
50s	English, pulmonary & cardiac symptoms only.	1208d3	No enzyme assays		(Eng *et al.*, 1993)

Table 3, cont.: Phenotypes and enzyme activities and mutations in atypical hemizygotes

Age (yr)	Phenotype in hemizygote	Mutation	α-galactosidase A activity*		Other comments (references)
			Material	% normal	
-	Japanese patient with cardiomyopathy only.	-	-		(Yamamoto *et al.*, 1989)
58	Japanese patient with cardiac involvement only	-	No enzyme assay. Glycosphingolipid accumulation in heart only		(Ogawa *et al.*, 1990)
71	Japanese patient with cardiac involement only	-	No enzyme assay. Glycosphingolipid accumulation in heart only		(Ogawa *et al.*, 1990)
51	Arab male with no symptoms	-	Leukocytes Fibroblasts	12% (α-gal.A & B) 23%, natural substrate used	Daughter had renal problems (Bach, *et al.* 1982)
26	Japanese patient with acroparesthesias only.	-	Fibroblasts	No activity with artificial substrate. Residual activity is seen using natural substrate	(Kobayashi *et al.*, 1985)
39	Italian patient with proteinuria only.	-	Leukocytes Fibroblasts	30.5% (α-gal.A & B) 20% (α-gal.A)	(Clarke *et al.*, 1971) (Romeo *et al.*, 1975)
48	Canadian patient with proteinuria only.	-	Leukocytes	16.5% (α-gal.A & B)	(Clarke *et al.*, 1971)

* = α-galactosidase A activity only, measured using the artificial substrate 4-methylumbelliferyl-α-D-galactopyranoside, except where the natural substrate, ceramide trihexoside, was used as indicated. Where indicated, α-galactosidase A and B activities were measured simultaneously.

1.3.2.2 Symptoms in female heterozygotes

Heterozygotes are either asymptomatic or show attenuated classic clinical symptoms. This is presumed to be caused by variations in the X-inactivation pattern and therefore expression of the defective and normal copies of the gene. About 70% have corneal dystrophy, 30% angiokeratoma and less than 10% suffer infrequent acroparesthesias. Other more severe, late onset symptoms are seen rarely (Desnick *et al.*, 1989). Many heterozygotes with atypical phenotypes have been reported, including two sisters, one with proteinuria as the only symptom and the other with haematuria (Chen, HC *et al.*, 1990). In a pair of monozygotic twins one sister was asymptomatic and the other was affected (Levade *et al.*, 1991).

The enzyme activities in these females range from values within the normal range (Avila *et al.*, 1973) to undetectable (Desnick *et al.*, 1973, 1989), so making heterozygote detection by enzyme assay difficult.

1.3.2.3 Histological examination

Histological examination provides evidence for the pathology in Fabry patients. Accumulation of glycosphingolipids in Fabry patients has been observed in the lysosomes of almost all tissues examined, with the exception of the hepatocytes, liver sinus endothelial cells and alveolar macrophages (Desnick *et al.*, 1989). They are also excreted into the urine. The glycosphingolipid deposits are crystalline and birefringent. They can be observed by polarising microscopy and show a 'Maltese cross' configuration. Glycosphingolipids can also be stained histochemically (Farragina *et al.*, 1981) and ultrastructural studies have shown the presence of concentric, lamellar lipid inclusions and other atypical deposits in the lysosomes of Fabry patients (Hashimoto *et al.*, 1976; Elleder *et al.*, 1990b).

1.3.3 Biochemical Aspects

1.3.3.1 Accumulation in glycosphingolipids in Fabry patients

The biochemical nature of Fabry disease was not investigated until about 1950, when it was discovered that the disease was characterised by the accumulation of

lipids. In 1963 the undegraded storage products were identified as the neutral glycosphingolipids, whose structures are indicated in figure 2 (Sweeley *et al.*, 1963).

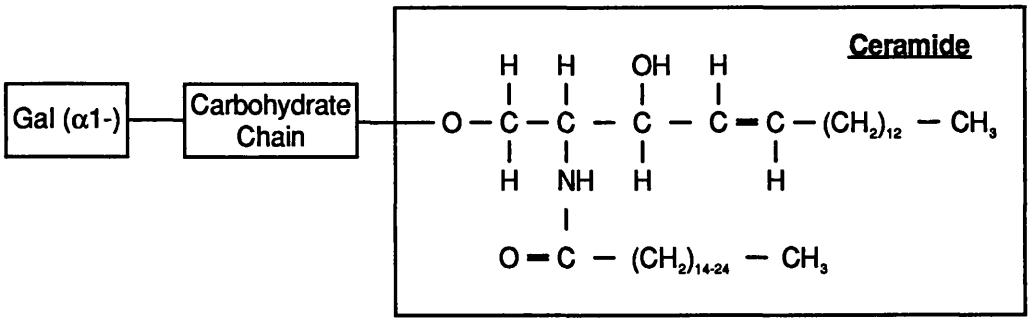
Two glycosphingolipid substrates for α-galactosidase A are found in all individuals. Globotriaosylceramide is the major substrate and in Fabry hemizygotes it accumulates to abnormally high proportions in most tissues and cells, except for the erythrocytes, which do not have lysosomes. In particular the vascular endothelium, muscle, kidney, lymph nodes, heart, prostate and autonomic ganglia are affected. Galabiosylceramide is less widely distributed but is elevated in the kidney, pancreas, heart, lung, urine and autonomic ganglia of Fabry hemizygotes (Desnick *et al.*, 1989). These observations fit with the clinical picture, in which involvement of the kidney, heart, nerve and vascular tissues are responsible for the major clinical manifestations.

Three additional glycosphingolipids are found in some individuals. Those with blood groups AB and B produce the B or B1 antigens, which contain terminal α-linked galactose and these have been found to accumulate in the epithelial tissues of the stomach, intestine and pancreas of Fabry hemizygotes (Wherret *et al.*, 1973). The third blood group substrate is the P_1 antigen but its accumulation in Fabry patients has not been investigated. The presence of these blood group substances in some individuals may increase the substrate load on the catabolic system and could explain why Fabry hemizygotes and heterozygotes with the AB or B blood groups are more severely affected than those without (Kint *et al.*, 1973). An additional variation in the substrate synthesis system is the existence of two alternative α-galactosyltransferases, the B^1 and B^2 enzymes, that produce the group B antigens (Yoshida, 1983). The B^2 enzyme is found in black and Oriental groups and in the heterozygous AB blood type the B^2 transferase has a sufficiently high activity to produce nearly as much B antigen as a blood group B homozygote with the B^1 enzyme. The significance of these factors in Fabry patients has not been investigated.

Analysis of the storage products by histological examination of tissue biopsies is often indicative of Fabry disease. Excretion of elevated amounts of storage product, globotriaosylceramide, in the urine is also diagnostically useful. Neutral glycosphingolipids can be easily detected by separation using thin layer chromatography and comparison with normal controls (Schlinder *et al.*, 1990).

Figure 2:The structure of the glycosphingolipids

(A) Basic Structure



(B) The structures of the specific neutral glycosphingolipids

Name	Chemical structure (C=ceramide)
Globotriaosylceramide	Gal (α1->4) Gal (β1->4) Glc (β1->1')-C
Galabiosylceramide	Gal (α1->4) Gal (β1->1')-C
Blood group B antigen	Gal (α1->3) Gal (2<-1αFuc) (β1->3) GlcNAc (β1->3) Gal (β1->4) Glc (β-1->1')-C
Blood group B1 antigen	Gal (α1->3) Gal (2<-1αFuc) (β1->4) GlcNAc (β1->3) Gal (β1->4) Glc (β-1->1')-C
Blood group P_1 antigen	Gal (α1->4) Gal (β1->4) GlcNAc (β1->3) Gal (β1->4) Glc (β-1->1')-C

They can be measured quantitatively by densitometric scanning of the plates and by gas-liquid chromatography (Hozumi *et al.*, 1990; Ogawa *et al.*, 1990).

In normal individuals, glycosphingolipids are found in the plasma membranes of all cells and are usually present in small quantities, about 5% of the total lipid, although higher concentrations are found in neuronal cells and the apical membranes of epithelial cells (Hoekstra *et al.*, 1992). The catabolic enzymic activities required to degrade the storage products found in Fabry disease are indicated in figure 3. The primary source of the neutral glycosphingolipids is *de novo* synthesis (Hoekstra *et al.*, 1992; Sandhoff *et al.*, 1992) in the liver (Brown *et al.*, 1981). A second major source is believed to be senescent erythrocytes (Dawson *et al.*, 1970; Desnick *et al.*, 1989), whose primary site of synthesis is in the bone marrow. The lack of significant pathology in the liver and bone marrow of Fabry patients indicates that they are not major catabolic sites.

1.3.3.2 The function and properties of α-galactosidase A

1.3.3.2.1 Alpha-galactosidase A deficiency in Fabry patients

The deficient enzyme in Fabry patients was found to be a ceramide trihexosidase (Brady *et al.*, 1967) thereby explaining the accumulation of ceramide trihexoside in the disease. In 1970 this enzyme was shown to catalyse the hydrolysis of α-galactosidic linkages in 4-methylumbelliferyl-α-D-galactosylpyranoside (4MU-α-gal) and the p-nitrophenyl analogue, permitting the use of these substrates for the diagnosis of Fabry disease. Leukocytes from Fabry patients were deficient in this activity (Kint, 1970). The hydrolase activity detected by these artificial substrates consisted of two component enzymes named α-galactosidase A and B. Purification of the B form showed that it was actually an α-N-acetylgalactosaminidase with activity towards synthetic substrates with α-galactosidic linkages (Elleder *et al.*, 1990b; Ogawa *et al.*, 1990). It has since been renamed α-N-acetylgalactosaminidase. It is not affected in Fabry disease. The A form is the true α-galactosidase and is the activity deficient in Fabry disease. Diagnosis of Fabry hemizygotes, in both classic cases and variants, is achieved by demonstrating a decrease in the rate of hydrolysis of the artificial substrate 4MU-α-gal, relative to the normal activity (Kint, 1970). A more accurate assay uses the same substrate but in the presence of N-acetylgalactosamine, an inhibitor of α-galactosidase B (Mayes *et al.*, 1980).

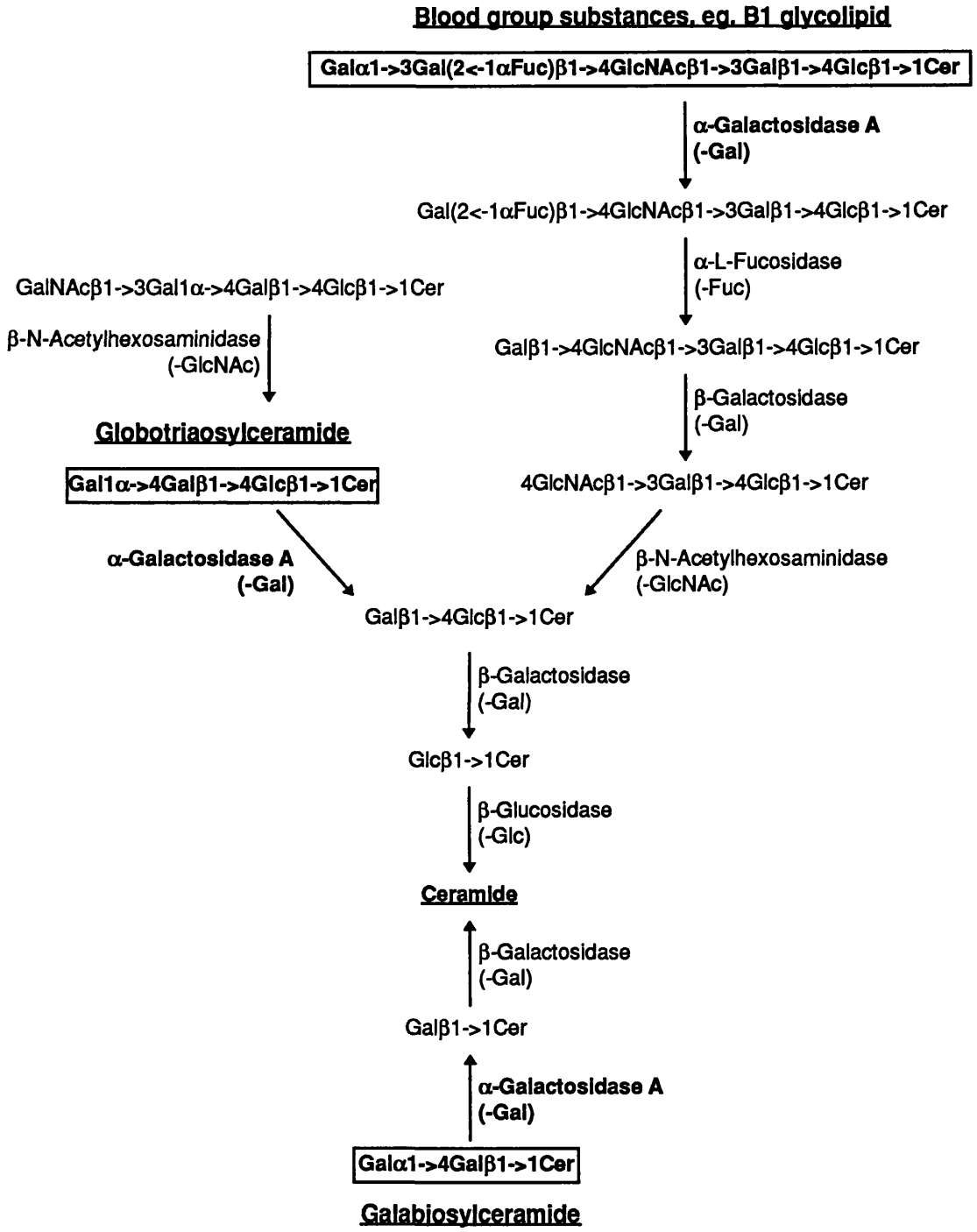


Figure 3: Degradation pathways for the catabolism of neutral glycosphingolipids

The catabolic enzyme activities required in the degradation of neutral glycosphingolipids, adapted from Desnick *et al.*, 1989. The boxes indicate substrates that accumulate in Fabry patients. Gal, Glc, Fuc, GalNAc, GlcNAc and Cer are galactose, glucose, fucose, N-acetylgalactosamine, N-acetylglucosamine and ceramide, respectively.

The enzyme activity may also be measured by using the native radiolabelled (^3H) substrate, ceramide trihexoside but preparation of the substrate is laborious (Rietra *et al.*, 1976). Deficient activity has been demonstrated in plasma, serum, leukocytes, tears, and cloned skin fibroblasts (Kint, 1970; Desnick *et al.*, 1973; Johnson *et al.*, 1975; Mayes *et al.*, 1980). Prenatal detection of Fabry hemizygotes has been performed on chorionic villus samples or amniocytes (Brady *et al.*, 1971).

The presence of cross-reacting immunologic material (CRIM) in Fabry patients has been investigated. Analysis of fibroblasts, leukocytes, renal tissue and urine has shown that there is no CRIM in some patients with no detectable enzyme activity (Beutler *et al.*, 1973; Rietra *et al.*, 1974; Hamers *et al.*, 1979) but it is present in others. In one study the processing of the mutant enzyme was investigated, showing that one patient did not produce an enzyme precursor, two produced precursor but no mature enzyme and one produced mature, inactive protein (Lemansky *et al.*, 1987). Thus, the disease is biochemically heterogeneous.

1.3.3.2.2 Amino acid sequence of human α-galactosidase A

Normal, human α-galactosidase A has been purified from human tissue. It consists of a 429 amino acid precursor protein that contains a cleavable, 31 amino acid signal peptide and a 398 amino acid mature peptide (Bishop *et al.*, 1991). The signal peptide is typical of other sequences required for transport into the endoplasmic reticulum (Watson, 1984; von Heijne, 1986) and has 5 charged amino acids at its N-terminus, a central hydrophobic core of 15 residues and a more polar C-terminus containing a proline or glycine at -4 to -5 from the cleavage site and the sequence alanine-X-alanine. The mature peptide contains four consensus sites for protein glycosylation at asparagines 139, 192, 215 and 408, although only three of these are utilised, based on studies using N-glycanase treatment of the oligosaccharides and estimates of the molecular weight (Bishop *et al.*, 1991). Site-directed mutagenesis studies have confirmed that full enzyme activity requires glycosylation at all sites except for asparagine 408 (Desnick *et al.*, 1994).

As with the other lysosomal enzymes (Neufeld, 1991), the amino acid sequence of α-galactosidase A showed little similarity to other human lysosomal hydrolases, with the notable exception of α-N-acetylgalactosaminidase (Wang, AM *et al.*, 1990, 1991). The amino acid sequences of the full length α-galactosidase A gene shared 46.9% overall identity with that of α-N-acetylgalactosaminidase. The carboxyl

terminus sequence encoded by exon 7 of α-galactosidase A showed only 15.8% homology, while the remaining amino terminal region had 56.4% identity. All of the N-glycosylation sites and 8 of the 10 cysteine residues in α-galactosidase A were conserved in α-N-acetylgalactosaminidase. Alpha-galactosidase A also had similarity with short regions of the amino acid sequences for α-galactosidases in *E. coli* and yeast, encoded by the Mel A and Mel 1 genes, respectively.

1.3.3.2.3 The physical and kinetic properties of α-galactosidase A

The properties of purified α-galactosidase A have been well studied (Bishop *et al.*, 1978, 1981a; Kusiak *et al.*, 1978; Dean *et al.*, 1979). The enzyme is a homodimer with a molecular weight of about 100kdal (two 50kdal units), containing 5-15% (w/w) asparagine-linked oligosaccharide chains of the complex or high mannose type. Multiple forms are found, with slightly different physical properties caused by the heterogeneity in their glycosylation. Their different isoelectric points may result from the presence of different numbers of sialic acid residues (Bishop *et al.*, 1981a).

The optimal activity is at pH 3.8-4.0 for the natural substrate, ceramide trihexoside and at pH 4.6 with the artificial fluorogenic substrate 4-MU-α-gal, with Michaelis-Menten constants (K_m) of 0.1-0.2mM and 2mM, respectively. *In vivo*, a sphingolipid activator protein, SAP-1, or saposin B, is required as a cofactor, which binds to the substrate and is essential for the enzyme to function (Furst *et al.*, 1992; Kishimoto *et al.*, 1992; Sandhoff *et al.*, 1994). Deficiency of saposin B results in a secondary deficiency of α-galactosidase A, β-galactosidase and cerebroside sulphatase activity, *in vivo* (Li *et al.*, 1985). The clinical symptoms are distinct from Fabry disease and are similar to those of metachromatic leukodystrophy (Stevens *et al.*, 1981).

The biosynthesis and processing of α-galactosidase A have been investigated. Pulse-labelling and immunoprecipitation studies have shown that α-galactosidase A is synthesised as a glycosylated 50.5kDa precursor in fibroblasts (Lemansky *et al.*, 1987). The lysosomal form was processed to 47-50kDa intermediates, over a period of 2 days and within 7 days a stable 46kDa protein was detected. A secreted form of the enzyme had a molecular weight of 52kDa. Like most lysosomal hydrolases the enzyme is phosphorylated, allowing targeting to the lysosome via the mannose 6-phosphate pathway. Secretion of only the 52kDa form of α-galactosidase A occurred in fibroblasts from a patient with I cell disease, in which

enzyme phosphorylation and therefore targeting to the lysosome by the mannose 6-phosphate receptor mechanism is defective. In Chang liver cells (Le Donne *et al.*, 1983) a 58kDa α-galactosidase A precursor is processed to a mature 49kDa much faster than in fibroblasts, within 3 hours of synthesis. The normal transport and processing of α-galactosidase A follows pathways shared by other lysosomal enzymes.

1.3.4 The α-galactosidase A gene and cDNA

Fabry disease was first described as an X-chromosome-linked disease in 1965 by pedigree analysis (Grace *et al.*, 1991). The α-galactosidase A gene (GLA) has been localised to the long arm of the X-chromosome at Xq22 by *in situ* hybridisation (Desnick *et al.*, 1989) and more recently it has been placed at Xq22.1, between the Bruton's tyrosine kinase gene (BTK) and the DXS178 locus, in a region of less than 300kb, on the physical map (Vetrie *et al.*, 1993, 1994; Sweatman *et al.*, 1994). The order of the DXS polymorphic markers in this region is indicated below:

centromere-DXS366-DXS442-BTK-**GLA**-DXS178-DXS265-DXS178CA-DXS101-telomere

The nearest polymorphic marker, DXS178, is estimated to be 100kb distal to the α-galactosidase A gene.

The full length cDNA (Bishop *et al.*, 1988, 1991) and the gene for α-galactosidase A (Kornreich *et al.*, 1989) have been isolated and sequenced. The GenBank accession numbers for the cDNA and gene are X14448 and D00039, respectively. The gene is 12kb long and has a promoter region that contains the consensus GC box and CCAAT transcription promoter sequences and the Ap-1, OCTA and c-fos enhancer binding sites (Bishop *et al.*, 1988). These are elements that are typical of constitutively expressed 'housekeeping genes, as is the presence of a TATA box, located about 25 bases upstream of the transcription start site. It contains 7 exons of 92-291bp in length, which are separated by introns of 200-3.7kb long. The gene also has 12 *Alu* repeats (30% of the gene) and is therefore comparatively rich in these sequences, since the genome contains a distribution density of 1 *Alu* repeat per 0.5-4kb (Makalowski *et al.*

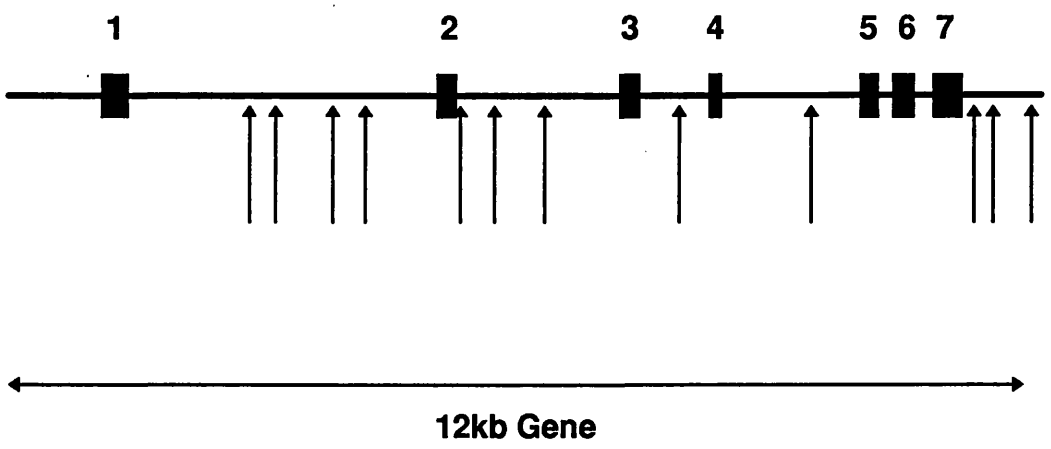


Figure 4: The α-galactosidase A gene structure and position of *Alu* repeats

Exon positions and sizes are indicated by the filled boxes and the exon numbers are indicated above. The arrows indicate the positions of the 5' ends of the 12 *Alu* repeats located in the introns and at the 3' end of the gene.

The messenger RNA is formed by splicing at sites with the normal invariant GT and AG consensus sites (Shapiro *et al.*, 1987) and it contains 60 bases of 5' untranslated sequence, 1290 nucleotides of protein coding sequence and a poly-A tail (Bishop *et al.*, 1988, 1991). Unusually, the 3' untranslated sequence of the mRNA is either absent or is very short, only 5-7 bases. It also has two polyadenylation signals, one of which is a rare sequence, and an snRNA binding site which binds U4 and may be involved in polyadenylation. No other human nuclear mRNAs lack a 3' untranslated region (Bishop *et al.*, 1991).

The gene and cDNA nucleotide sequences for α-galactosidase A have homology with the α-N-acetylgalactosaminidase gene. Introns 1-6 of the α-galactosidase A gene interrupt the protein-coding sequence at similar positions to introns 2-7 of the α-N-acetylgalactosaminidase gene, indicating that they may be evolutionarily related. The full length cDNA nucleotide sequences of these two gene have 55.8% identity. Only one other nucleotide sequence for human lysosomal enzymes had any similarity. At the 5' end of the cDNA for α-fucosidase (Fukushima *et al.*, 1985), bases 31-178 showed 52% identity with bases 34-168 of the α-galactosidase cDNA when four gaps were introduced (Desnick *et al.*, 1989).

1.3.5 Family counselling, methods of diagnosis and their problems

In Fabry families it is important to diagnose hemizygotes and heterozygotes for counselling because both can be clinically affected and can transmit the defect in a single chromosome to produce an affected child.

1.3.5.1 The problem of heterozygote detection by clinical and biochemical methods

Hemizygotes for Fabry disease can be identified conclusively by enzyme assay and their diagnosis is not a problem. Heterozygotes can be identified when the symptoms are clearly observed and the α-galactosidase A activities are below the normal range. However, the diagnosis of heterozygotes that are clinically asymptomatic and have enzyme activities in the normal range is a major problem in Fabry families. Phenotypic variability is caused by the process of random X-inactivation (Lyon, 1961; Migeon, 1994). In all somatic cells, only one X-chromosome is genetically active and the other is inactivated, for most genes. Random inactivation occurs early during embryonic development and before tissue differentiation. Once this occurs, the same chromosome, either maternal or paternal, is the only active chromosome in all daughter cells. This causes mosaicism in the pattern of enzyme expression and some cells will produce the normal protein while others express the defective gene. The extent of accumulation of the storage product and therefore the phenotype and disease progression in carrier females depends on the proportion and location of cells expressing the mutant. This phenomenon accounts for the variant phenotypes in female monozygotic twins, in which one twin was clinically normal and the other was affected (Levade *et al.*, 1991; Winchester *et al.*, 1992) and highlights the problem of predicting the clinical effect in heterozygotes and in diagnosis of asymptomatic females.

Another potential, complicating factor in the detection of Fabry heterozygotes is the possibility of age-related reactivation of the inactive X-chromosome. The ornithine transcarbamylase gene on the inactive X chromosome has been shown to be reactivated in older mice (Wareham *et al.*, 1987). This occurrence may explain the observation that older heterozygotes of Fabry disease tend to have an α-galactosidase activity nearer to the normal range than younger heterozygotes (Desnick *et al.*, 1987) and may also account for the increase in fibroblast α-galactosidase A activity with cell age (Hozumi *et al.*, 1990).

The problem of X-inactivation and diagnosis by enzyme assay has led to the development of modified biochemical assays. The heterozygote detection frequency was improved by measuring the activity of other lysosomal enzymes such as the β-galactosidase, in addition to α-galactosidase, as a control (Spence *et al.*, 1977; Sheth *et al.*, 1981). Others have measured enzyme activity in single cell fibroblast clones (Romeo *et al.*, 1970; Jongkind *et al.*, 1983) or in hair roots (Spence *et al.*, 1977) of heterozygotes to detect the presence of single cells with no enzyme activity. Hair follicles originate from a small cell population and so a significant proportion are likely to have cells with the same inactivated chromosome. Therefore, by testing several hair follicles for enzyme activity, diagnosis may be made if a significant number of enzyme-deficient follicles are found and a trimodal distribution of activities is observed (Ropers *et al.*, 1977). However, these methods are difficult, very labour intensive and may still yield inconclusive results. Therefore, no satisfactory biochemical methods for heterozygote detection have been developed.

1.3.5.2 Molecular genetic detection of heterozygotes

Genetic detection methods can identify heterozygotes (Desnick *et al.*, 1987). Analysis of closely linked polymorphisms can be used to track the inheritance of the defective gene. The polymorphic markers DXS17, DXS87 and DXS88 have been used in analysis of Fabry families, using Southern blot analysis of RFLPs (MacDermot *et al.*, 1987; Morgan *et al.*, 1987). These three regions can also be amplified by PCR for RFLP analysis and have been used for carrier detection in some families (Kornreich *et al.*, 1992). The closest known marker to the α-galactosidase A gene is DXS178, which is about 100kb telomeric to the gene (Sweatman *et al.*, 1994; Vetrie *et al.*, 1994). CA repeat markers have also been identified at the DXS178 locus (Allen *et al.*, 1992; de Weers *et al.*, 1992) and may prove useful for carrier detection in Fabry families.

In addition to polymorphic markers located outside of the α-galactosidase A gene, two other polymorphisms have been found within, or close to the GLA locus. A *Sac* I RFLP in intron 4 of the gene was observed at a frequency of 0.08 in normals and 0.02 in Fabry families. A rare *Nco* I RFLP, located about 10kb downstream from the last gene exon, has been found at a frequency of 0.13 in normals and 0.12 in Fabry families (Desnick *et al.*, 1987). The *Nco* I RFLP was used for diagnosis in a Fabry

family with 60 members. In this family, 7 obligate heterozygotes that were undetected by enzyme assay were identified (Kirkilionis *et al.*, 1991). No intragenic polymorphisms, except for the *Sac* I RFLP were known prior to this study.

Two problems exist for the identification of heterozygotes by polymorphism analysis. One is that the polymorphisms are not always informative. The second problem is that in the absence of a family history of Fabry disease, the presence of a new disease-causing mutation cannot be determined. Detection of a linked polymorphism in an affected child and subsequent detection of the polymorphism in the mother does not conclusively show that the mother has the mutation as it may have arisen *de novo* in the child. Therefore, direct detection of the specific defect for heterozygote identification is desirable.

The cloning of the gene and the lack of informative, intragenic polymorphic markers for heterozygote detection has prompted the study of the family-specific mutations for use in carrier detection. This allows direct analysis of inheritance of the genetic defect. One problem exists with this type of analysis. If the defect is a new mutation and the individual is a mosaic then it is possible that analysis of DNA from somatic cells will not show the presence of a mutation which is in the germ-line cells (Hall, 1988). In this case, although the individual appears to be genetically normal, the children are at risk from inheriting the defect and should be considered for genetic counselling.

1.3.6 Treatment

1.3.6.1 Palliative treatment

No cure for Fabry disease has yet been developed and so palliative treatment has been used to counteract the clinical manifestations of the disease. Analgesics such as phenoxybenzamine, carbamazepine and diphenylhydantoin have been administered to alleviate pain, although these are not entirely effective. Renal insufficiency can be treated by dialysis and kidney transplantation (Desnick *et al.*, 1989). Kidney transplantation as a method of supplying endogenous, normal α-galactosidase A has also been carried out and results in improvement of symptoms in some patients (Desnick *et al.*, 1972; Philippart *et al.*, 1972; Mosnier *et al.*, 1991) but not in others (Spence *et al.*, 1976). However, even when symptoms are alleviated, transplantation of a normal kidney does not completely prevent

substrate deposition in the allograft (Mosnier *et al.*, 1991) or death by cardiac involvement (Kramer *et al.*, 1984).

1.3.6.2 Therapy

Treatment of Fabry disease at a biochemical level has been attempted, although with little success so far (Desnick *et al.*, 1981, 1989; Desnick, 1983). It has been shown that α-galactosidase from fig, coffee bean and human sources is internalised by cultured fibroblasts from Fabry patients and can catabolise the intracellular glycosphingolipid deposits (Desnick *et al.*, 1989). Enzyme replacement has also been carried out *in vivo* by repeated injection of purified human enzyme into Fabry patients. This was shown to rapidly lower the levels of accumulated glycosphingolipids in the plasma but only temporarily, since the half-life of α-galactosidase A is short, about 10 minutes for the splenic isozyme and 70 minutes for the injected plasma enzyme , Desnick *et al.*, 1979. The availability of purified enzyme is a limiting factor in this type of treatment. Foetal liver transplantation has also been used to provide a source of normal enzyme but the long-term effect has not been evaluated (Touraine *et al.*, 1979; Grosshans, 1986; Desnick *et al.*, 1989). A more indirect approach is depletion of the substrate. Globotriaosylceramide has been removed from plasma by plasmapheresis to try to limit the clinical effects of the enzyme deficiency (Pyeritz *et al.*, 1980) but it cannot reverse the effects of substrate deposition once the substrate reaches the lysosomes. Removal of senescent erythrocytes, a source of substrate, has also been used but was not therapeutic (Beutler *et al.*, 1981).

Progress in enzyme therapy will depend on improving methods for producing and isolating large amounts of catalytically active protein. Recently, expression of the cDNA sequence in insect cells, using a baculovirus vector, has produced α-galactosidase enzyme that is glycosylated and catalytically active (Coppola *et al.*, 1994). Approximately 5.6mg of recombinant α-galactosidase A per litre of cell culture was obtained and purified. The enzyme was an estimated 46.6-46.9kDa, by mass spectral analysis and was glycosylated to contain fucose, galactose, mannose and N-acetylglucosamine but did not include xylose, N-acetylgalactosamine or sialic acid. Human α-galactosidase A is known to possess sialic acid residues and so the glycosylated, recombinant protein was not identical. Uptake of the recombinant baculovirus protein in cultured fibroblasts from a Fabry patient caused an increase in α-galactosidase A activity to 60% of the normal activity and

this was partially inhibited by mannose 6-phosphate, indicating the involvement of the MPR for transport, as with the human enzyme. After covalent modification with poly ethylene-glycol (PEG), a procedure designed to enhance stability during enzyme replacement therapy, the recombinant enzyme retained 24% of activity with the natural substrate, trihexosylceramide. Further investigation of this baculovirus enzyme and its stability is required. The lack of sialic acid residues may be a problem for enzyme therapy as it is proposed that the human plasma α-galactosidase A has a longer half-life in the circulation than splenic enzyme due to the presence of more sialic acid residues (Desnick *et al.*, 1989). A study of glycoproteins has shown that desialylated proteins are removed from the blood circulation more rapidly than those that are sialylated (Ashwell *et al.*, 1974). Alpha-galactosidase A has also been over-expressed in CHO cells and the resultant enzyme was shown to be glycosylated, catalytically active and could bind to the MPR (Iounnou *et al.*, 1992). Recombinant proteins such as these may be useful for enzyme replacement therapy but no *in vivo* studies have yet been described.

No animal models have been found for the study of Fabry disease and gene therapy for this disorder has not been investigated.

1.3.7 Mutation detection in the α-galactosidase A gene

At the start of this study, analysis of genomic DNA by Southern blotting had shown that most of the mutations in the α-galactosidase A gene were small rearrangements or point mutations. Gross gene rearrangements were identified in only 6 out of 130 families (5%) by this method (Bernstein *et al.*, 1989; Kornreich *et al.*, 1990). A multiplex PCR-based method was subsequently developed to detect this type of mutation in genomic DNA (Kornreich *et al.*, 1993). In addition, 5 point mutations (Bernstein *et al.*, 1989; Koide *et al.*, 1990; Sakuraba *et al.*, 1990; von Scheidt *et al.*, 1991), a 13bp deletion (Ishii *et al.*, 1991) and a splice site mutation (Yokoi *et al.*, 1991) had been reported. All of these were detected by reverse transcription of the RNA and sequencing of the resultant cDNA, except for one point mutation detected by Southern blot analysis (Bernstein *et al.*, 1989). No methods had been reported for PCR amplification of the α-galactosidase A gene from genomic DNA and no methods for mutation screening in Fabry patients, except for Southern blotting, had been described. Therefore there was a need to develop methods for detecting small mutations that were undetected by Southern

blotting techniques. This information would allow the detection of heterozygotes by genetic analysis and would also permit genotype comparisons with the phenotype.

During the course of this investigation, many new mutations were reported. In total, 93 different putative disease-causing mutations have been found in 106 families, providing conclusive evidence for the genetic heterogeneity of the Fabry disease. These are listed in table 4. Allele designations for the point mutations in the coding region were given according to the predicted amino acid substitutions and their position in the protein. For the small deletions and insertions the allele designations were based on their position in the cDNA and the size of the change, as follows: the cDNA number immediately 5' to the first nucleotide which showed a sequence difference was given, followed by del or ins for deletions and insertions respectively and then its size. Thus, 716ins1 indicates that when reading in a 5'-3' direction, the first base alteration is observed at cDNA nucleotide 717 and that it is a 1 base insertion. Splice-site mutations were named according to their position in the intron and the base substitution or deletion. For example, IVS3-1,t indicates that the mutation is in intron (IVS) 3, it is at the 3' end of the intron, -, at position 1bp upstream of exon 4. The conserved guanine of the 'ag' dinucleotide is replaced with a thymine. A +2 would indicate a mutation at the 5' end, 2bp downstream of the preceeding exon.

Table 4: Mutations in the α-galactosidase A gene

(A): Point mutations in the α-galactosidase A cDNA

Exon	Ref.	Nucleotide		Allele	Nucleotide Substitution	At CpG	R/E Site
		cDNA (1)	Gene (2)				
1	3	95	1274	L 32 P	CTG->CCG		
	3,14	101	1280	N 34 S	AAT->AGT		
	12	103	1282	*G 35 R	GGA->AGA		
	4	118	1297	P 40 S	CCT->TCT		
	13	124	1303	*M 42 V	ATG->GTG		-*Nco* I
	5	132	1311	W 44 X	TGG->TGA		
	13	145	1324	*R 49 S	CGC->AGC	Yes	-*Hha* I
	12	146	1325	*R 49 L	CGC->CTC	Yes	-*Hha* I
	16	155	1334	C 52 S	TGC->TCC		
	14	166	1345	C 56 G	TGC->GGC		+*Hae* III
	16	167	1346	C 56 F	TGC->TTC		
	13	167	1346	*C 56 Y	TGC->TAC		
	16	175	1354	E 59 K	GAG->AAG		
2	7	196	5095	E 66 Q	GAG->CAG		
	16	242	5141	W 81 X	TGG->TAG		
	3	253	5152	G 85 N	GGT->AGT		
	16	266	5165	L 89 R	CTC->CGC		
	13	274	5173	*D 92 H	GAT->CAT		
	16	295	5194	Q 99 X	CAA->TAA		
	16	299	5198	R 100 K	AGA->AAA		
	13	319	5218	*Q 107 X	CAG ->TAG		
	7	334	5233	R 112 C	CGC->TGC	Yes	
	16	335	5234	R 112 H	CGC->CAC	Yes	
3	16	392	7291	L 131 P	CTA->CCA		
	16	427	7226	A 143 P	GCA->CCA	Yes	
	16	431	7230	G 144 V	GGC->GTC		
	8	436	7335	P 146 S	CCT->TCT		-*Cel* II
	3	466	7365	A 156 T	GCC->ACC		
	16	469	7364	Q 157 X	CAG->TAG		
	14	484	7383	W 162 R	TGG->CGG		+*Msp* I
	12	494	7393	*D 165 V	GAT->GTT		-*Bgl* II
	16	515	7414	C 172 Y	TGT->TAT		
4	8	606	8379	C 202 W	TGT->TGG		+*Eco*R II

*** Mutations found in this study**

Table 4 (A), cont: Mutations in the α-galactosidase A gene

Exon	Ref.	Nucleotide		Allele	Nucleotide	At CpG	R/E Site
		cDNA (1)	Gene (2)				
5	9,14	644	10135	*N 215 S	AAT->AGT		-*Tsp*E I
							+*Bsr* I
	15	658	10149	R 220 X	CGA->TGA	Yes	
	12	677	10168	*W 226 X	TGG->TAG		+*Rma* I
							+*Mae* I
	9,14	679	10170	*R 227 X	CGA->TGA	Yes	+*Hph* I
	14	680	10171	*R 227 Q	CGA->CAA	Yes	
	13	708	10199	*W 236 C	TGG->TGC		
	16	730	10221	D 244 N	GAC->AAC		
	14	791	10282	D 264 V	GAC->GTC		
	14	797	10288	D 266 V	GAT->GTT		*Bsr* I
6	9	806	10514	*V 269 A	GTG->GCG		
	16	816	10524	N 272 K	AAC->AAA		
	7	835	10543	Q 279 E	CAG->GAG		
	13	859	10567	*W 287 G	TGG->GGG		
	9	861	10569	*W 287 X	TGG->TGA		-*Bgl* I
	16	863	10571	A 288 D	GCT->GAT		
	10	886	10594	M 296 V	ATG->GTG		
	14	890	10598	S 297 F	TCT->TTT		+*Mse* I
	16	901	10609	R 301 X	CGA->TGA	Yes	
	5,3	902	10610	R 301 Q	CGA->CAA	Yes	
	14	937	10645	D 313 Y	GAT->TAT		
	12	947	10655	*V 316 E	GTA->GAA		-*Mae* II
	9	979	10687	*Q 327 K	CAA->AAA		
	7	982	10690	G 328 R	GGG->AGG		
	14	983	10691	G 328 A	GGG->GCG		-*Kpn* I
7	13	1018	10996	*W 340 R	TGG->CGG		
	14	1020	10998	W 340 X	TGG->TGA		
	9	1024	11002	*R 342 X	CGA->TGA	Yes	
	8,13	1025	11003	*R 342 Q	CGA->CAA	Yes	
	11	1066	11044	R 356 W	CGG->TGG	Yes	
	9	1081	11059	*G 361 R	GGA->AGA		-*Sau*96 I
	14	1192	11170	E 398 X	GAA->TAA		+*Ase* I
							+*Mse* I

*** Mutations found in this study**

Table 4, cont: Mutations in the α-galactosidase A gene

(B): Splice site mutations in the α-galactosidase A gene

Splice Site	Ref.	Allele name	Gene Nucleotide (1)	Nucleotide Substitution	Resuit
5' Intron 2	5	IVS 2+2, g	5270	exon 2-gt->gg	
3' Intron 3	12	IVS 3-1, t	8321	*ag-exon 4 >at	
3' Intron 3	19	IVS 3-1, a	8321	ag-exon 4 -> aa	Skips exon 4
3' Intron 5	14	IVS 5-del 2,3	10507-8	tcag-exon 6-> tg-exon 6	
5' Intron 6	6	IVS 6+1, t	10708	exon 6-gt -> tt	Skips exon 6

(C): Large rearrangements in the α-galactosidase A gene

Exons/ introns rearranged	Ref.	Direct repeat sequences	Deletion (del)/ insertion (ins)/ inversion (inv)/ duplicated (dup) size	Rearranged gene [1] (& cDNA [2]) nucleotides
Intron 2 -> Intron 4	11	38bp repeats (Alu repeats)	3197 bp (del)	6587-9745 + 38 bp repeat
Exon 1-> Intron 2	11	CCA	4651 bp (del)	1198-5845 + 3 bp repeat
Intron 2 -> 3' gene	11	AAG	4519 bp (del)	6739-11258 + 3 bp repeat
Intron 2 -> Intron 3	11, 17	AGAACT	402 bp (del)	7086-7487 + 5 bp repeat
Intron 1 -> Exon 6	11	TAGACA	8112 bp (dup)	2594 (or 5) - 10705 (or 6)
Exon 5 -> 3' gene	11	See text	1710 bp (del) 151 bp (inv)	10237-11932 del. 12084-12097 del. 11933-12083 inv.

*** Mutations found in this study**

Table 4, cont: Mutations in the α-galactosidase A gene

(D): Small rearrangements in the α-galactosidase A gene

Allele (Exon)	Ref.	Sequence change & cDNA nucleotides Normal -> mutant	Rearranged gene [1] (& cDNA [2]) nucleotides
25del1 (1)	16	22-CTACATC-28 -> CTAC-TC	1204 (26)
124del13 (1)	18	123-CATGGGC——ACTGGGAG-144 -> CA-CTGGGAG	1304-1316 (125-137)
333del18 (2)	16	331-CAGCGC——ATTCGC-354 -> CAG-CGC	5233-5250 (334-351)
358del6 (2)	16	358-CTAGCTAA-365 -> C-A	5258-5263 (359-364)
***716ins1 (5)**	9	714-TATAAAG-720 -> TATTAAAG	10208 (717)
***717del2 (5)**	12, 13	714-TATAAAG-720 -> TATAG	10209-10210 (718-719)
773del2 (5)	14	772-GGACCAGG-779 -> GG-CAGG	10266-10267 (774-775)
954ins5 (6)	14	949-ATTGCCAT-956 -> ATTGCC-TTGCC-AT	10658-10662 (950-954)
***1010del19 (7)**	12	1008-CTTTGA——CCTCT-1031 -> CTT-CT	10989-11007 (1010-1029)
1016del11 (7)	14	1014-AGTGTGGGAACGACCTCT-1031 -> AGT-TCT	10996-11006 (1017-1027)
1020del1 (7)	16	1018-TGGGA-1022 -> TGG-A	10999 (1021)
1040ins1 (7)	8	1038-CTTA-1041 -> CTTTA	11018 (1040) repeated
1073del2 (7)	8	1070-AGGAGATT-1077 -> AGGA-TT	11052-11053 (1073-1074)
***1087del1 (7)**	13	1085-CTCGCTC-1091 -> CTC-CTC	11066 (1088)
1123del53 (7)	8, 14	1120-AAAGGAGT——AAAAGGAA-1178-> AAAG-AA	11101-11153 (1124-1176)
1208del3 (7)	14	1207-TTAAGAAGTC-1216 -> TTAAG-TC	11190-11192 (1208-1213)

*** Mutations found in this study**

Table 4, cont: Mutations in the α-galactosidase A gene

(E): Complex mutations in the α-galactosidase A gene

Exon	Ref.	Nucleotides		Allele	Nucleotide substitutions Normal -> mutant	R/E site
		cDNA (1)	Gene (2)			
2	16	359 & 361	5258 & 5260	L 120 P-A 121 T	356-AG CTA GCT -> AG CCA ACT	*Nhe* I
5	16	656 & 657	10147 & 10148	I 219 N	652-GAA ATC C -> GAA AAT C	-

REFERENCES

1. Bishop, D.F., Kornreich, R., Eng, C.M., Ioannou, Y.A., Fitzmaurice, T.F. and Desnick, R.J. (1991) In Salvayre, R., Douste-Blazy, L. and Gatt, S. (ed), *Lipid storage disorders - biological and medical aspects.* Plenum press, New York, pp. 809-822.
2. Kornreich, R., Desnick, R.J. and Bishop, D.F. (1989) *Nucleic Acids Res*, **17**, 3301-3302.
3. Madsen, K.M., Hasholt, L., Sorensen, S.A., Largerstrom-Fermer, M. and Dahl, N. (1995) *Hum Mutat*, **5**, 277-278.
4. Koide, T., Ishiura, M., Iwai, K., Inoue, M., Kaneda, Y., Okada, Y. and Uchida, T. (1990) *FEBS Lett*, **259**, 353-356.
5. Sakuraba, H., Oshima, A., Fukuhara, Y., Shimmoto, M., Nagao, Y., Bishop, D.F., Desnick, R.J. and Suzuki, Y. (1990) *Am J Hum Genet*, **47**, 784-789.
6. Sakaruba, H., Eng, C., Desnick, R.J. and Bishop, D.F. (1992) *Genomics*, **12**, 643-650.
7. Ishii, S., Sakaruba, H. and Suzuki, Y. (1992) *Hum Genet*, **89**, 29-32.
8. Ploos van Amstel, J.K., Jansen, R.P.M., de Jong, J.G.N., Hamel, B.C. and Wevers, R.A. (1994) *Hum Mol Genet*, **3**, 503-505.
9. Davies, J.P., Winchester, B.G., and Malcolm, S. (1993) *Hum Mol Genet*, **2**, 1051-1053.
10. von Scheidt, W., Eng, C.M., Fitzmaurice, T.F., Erdmann, E., Hubner, G., Olsen, G.J., Christomanou, H., Kandolf, R., Bishop, D.F. and Desnick, R.J. (1991) *N Eng J Med*, **324**, 395-399.
11. Bernstein, H.S., Bishop, D.F., Astrin, K.H., Kornreich, R., Eng, C.M., Sakuraba, H. and Desnick, R.J. (1989) *J Clin Invest*, **83**, 1390-1399.
12. Davies, J.P., Winchester, B.W., and Malcolm, S. (1994) *Hum Mol Genet*, **3**, 667-669.
13. Davies, J.P., MacDermot, K., Eng, C.M., Winchester, B.W., Malcolm, S., and Desnick, R.J., *unpublished data.*
14. Eng, C.M., Resnick-Silverman, L.A., Niehaus, D.J., Astrin, K.H. and Desnick, R.J. (1993) *Am J Hum Genet*, **53**, 1186-1197.
15. Meaney, C., Blanch, L.C. and Morris, C.P. (1994) *Hum Mol Genet*, **3**, 1019-1020.
16. Eng, C.M., Niehaus, D.J., Enriquez, A.L., Burgert, T.S., Ludman, M.D. and Desnick, R.J. (1994) *Hum Mol Genet*, **3**, 1795-1799.
17. Fukuhara, Y., Sakuraba, H., Oshima, A., Shimmoto, M., Nagao, Y., Nadaoka, Y., Suzuki, T., and Suzuki, Y. (1990) *Biochem Biophys Res Comm*, **170**, 296-300.
18. Ishii, S., Sakuraba, H., Shimmoto, M., Minamikawa-Tachino, M.S., Suzuki, T., and Suzuki, Y. (1991) *Ann Neurol*, **29**, 560-564.
19. Yokoi, T., Shinoda, K., Ohno, I., Kato, K., Miyawaki, T. and Taniguchi, N. (1991) *Jpn J Hum Genet*, **36**, 245-250.

The genetic heterogeneity of Fabry disease is characteristic of many X-linked diseases (Mandel *et al.*, 1992).

1.3.7.1 Methods for mutation analysis

A number of methods for mutation analysis of RNA or genomic DNA, used in the study of other genes, were considered for identifying unknown, putative disease-causing mutations in the α-galactosidase A gene.

Direct sequencing is the only method for determining the precise nature and location of specific base alterations but manual methods (Sanger *et al.*, 1977; Ruano *et al.*, 1991; Sheffield *et al.*, 1993) are time-consuming and laborious, while automated sequencing requires expensive equipment. However, the approximate location of unknown mutations can be identified, by other relatively inexpensive mutation screening methods, (reviewed by Grompe, 1993), to reduce the portion of DNA for sequencing. These methods are described below.

1.3.7.1.1 Analysis of RNA and DNA

Genomic DNA is the preferred choice of template for sequence analysis because it is highly stable and so it can be handled and stored with little expertise. It contains all of the mutated sequence and it is present in almost all cell types.

RNA analysis requires more expertise and care in isolation due to its instability. It often requires reverse-transcription into cDNA for analysis. In addition, it does not always contain the mutation. Some mutations result in a decrease in the amount of RNA, for example a stop mutation (Zhang *et al.*, 1994) and may cause difficulty in obtaining enough RNA for analysis. Other mutations, such as a complete gene deletion (Brown *et al.*, 1981) do not produce RNA and could not be detected. However, despite these disadvantages, one advantage over DNA analysis is that mutations causing abnormal mRNA splicing are more easily identified. Mutation screening using genomic DNA often involves analysis of the protein-coding regions of DNA and avoids analysis of the introns, which may contain mutations that alter splicing. For example, an intronic mutation in a patient with Scheie syndrome created a splice site that would have been undetected by this strategy but was easily identified by RNA analysis (Moskowitz *et al.*, 1993).

1.3.7.1.2 Detection of large rearrangements

Large rearrangements can be conveniently detected by analysing total cell DNA directly, using Southern blotting techniques or by multiplex PCR methods. Southern blotting (Southern, 1975) involves the fragmentation of genomic DNA by digestion with restriction enzymes and separation by agarose gel electrophoresis. The DNA is transferred from the gel to a nylon membrane by Southern blotting and hybridised to a radiolabelled DNA probe to detect size alterations. It has been used to detect mutations in many disorders, including Hunter disease (Froissart *et al.*, 1993). This method does not require specific sequence information for analysis and detects the presence of all large rearrangements. Point mutations can be detected if they alter a restriction site (Bernstein *et al.*, 1989). Disadvantages are that it is not always informative for small rearrangements and it requires a large amount of undegraded DNA sample.

Multiplex PCR reactions amplify a complete gene sequence, in several large fragments, using genomic DNA as a template. Size alterations are then detected by electrophoretic separation through an agarose gel, (e.g. in Fabry disease, Kornreich *et al.*, 1993). Sequence information is required for analysis but it allows detection without requiring a gene-specific probe and with minimal template DNA. As with Southern blotting, small rearrangements might not be detected.

1.3.7.1.3 Detection of small sequence alterations by alterations in the DNA conformation

PCR amplification of specific areas of genomic DNA or reverse-transcribed RNA (cDNA), followed by analysis of their conformation, which is sequence dependent, is the basis of several mutation detection methods. In each of the following methods, sequence changes cause alterations in the electrophoretic migration of DNA, which are detected on an acrylamide-based gel system using autoradiography or non-isotopic visualisation. The wild-type DNA is analysed in an adjacent lane to the mutant sequence.

(a) Double-stranded conformation analysis, DSCA, (Saad *et al.*, 1994) detects altered migration of homoduplex DNA by electrophoresis though a non-denaturing polyacrylamide gel.

(b) In heteroduplex analysis (Keen *et al.*, 1991; White *et al.*, 1992) the DNA for analysis is melted and annealed with normal DNA so that the mutant alleles form incompletely paired DNA duplexes. These are separated by electrophoresis through a non-denaturing acrylamide gel and their migration is retarded relative to the homoduplexes.

(c) In single-strand conformation polymorphism, SSCP (Orita *et al.*, 1989; Hayashi, 1991) analysis, the DNA is denatured and alterations in secondary structure are detected by electrophoresis in a non-denaturing polyacrylamide gel.

All three methods require minimal DNA template. Heteroduplex analysis and SSCP analysis detect small rearrangements and point mutations with more than 80% sensitivity in DNA that is less than 500bp in length, using optimised conditions. The gel systems are simple and the DNA does not require modification, except by denaturation for SSCP detection and for the formation of heteroduplexes. The main disadvantage is that some sequence alterations are not detected.

1.3.7.1.4 Detection of small sequence changes by other methods

In denaturing gradient gel electrophoresis, DGGE, (reviewed by Cariello *et al.*, 1993; Fodde *et al.*, 1994) DNA is amplified by the PCR and is analysed by migration through a polyacrylamide gel that contains a linearly increasing gradient of denaturant, such as urea and formamide. The homoduplex DNA is progressively denatured until it reaches a concentration which is equivalent to the melting temperature of the lowest melting domain. At this point the migration is severely retarded. Since sequence variations in the highest melting temperature domains are often not resolved, a G/C tail is sometimes incorporated into the PCR product to create a high melting temperature domain, ensuring that mutations are in a lower melting domain. The sensitivity of the method is enhanced by mixing mutant and normal DNA to form heteroduplexes, in addition to homoduplex DNA (Sheffield *et al.*, 1989). The major disadvantage is that the preparation of the denaturing gel requires more expertise than for analysis of changes in the DNA conformation.

The RNase cleavage method (Myers *et al.*, 1985) uses a radio-labelled wild-type, RNA probe which is hybridised to the sample DNA, forming an RNA-DNA heteroduplex. Areas of mismatch are digested with RNase A, which cleaves single-

stranded RNA, and these are separated by electrophoresis on an agarose gel. The requirement for radioactively labelled RNA and its low detection efficiency, about 50%, have limited its use.

In chemical mismatch cleavage (Cotton *et al.*, 1988; Smooker *et al.*, 1993), DNA is amplified by PCR. The radio-labelled wild-type DNA is mixed with the either unlabelled or labelled sample (Forrest *et al.*, 1991) to form heteroduplexes and base changes are detected by cleavage of mismatches. DNA analysis is by electrophoretic separation through a polyacrylamide gel. Hydroxylamine and osmium tetroxide bind to unpaired cytosine and thymine bases, respectively, and the DNA can then be cleaved with piperidine. The method detects close to 100% of mutations, although T-G mismatches, which are the most stable, are not always successfully cleaved. Alternative mismatch-detection methods use a carbodiimide (CDI) to modify guanine and thymine bases. The DNA surrounding the base modified with CDI can be cleaved with *E. coli* ABC excinuclease for electrophoretic separation. DNA-bound CDI can also be detected by binding to antibodies. Alternatively and more usually, mis-matched bases that bind CDI can be detected by PCR amplification from a primer towards the modified base, in a 5' to 3' direction. The polymerase cannot extend beyond the modified base, thus creating a shortened PCR-product. The mutation can be localised, an advantage of this method over the methods exploiting changes in DNA conformation. Disadvantages are the toxicity of the chemical reagents and the requirement for mismatch modification steps prior to electrophoretic analysis.

1.3.7.1.5 Mutation screening strategy used in this study

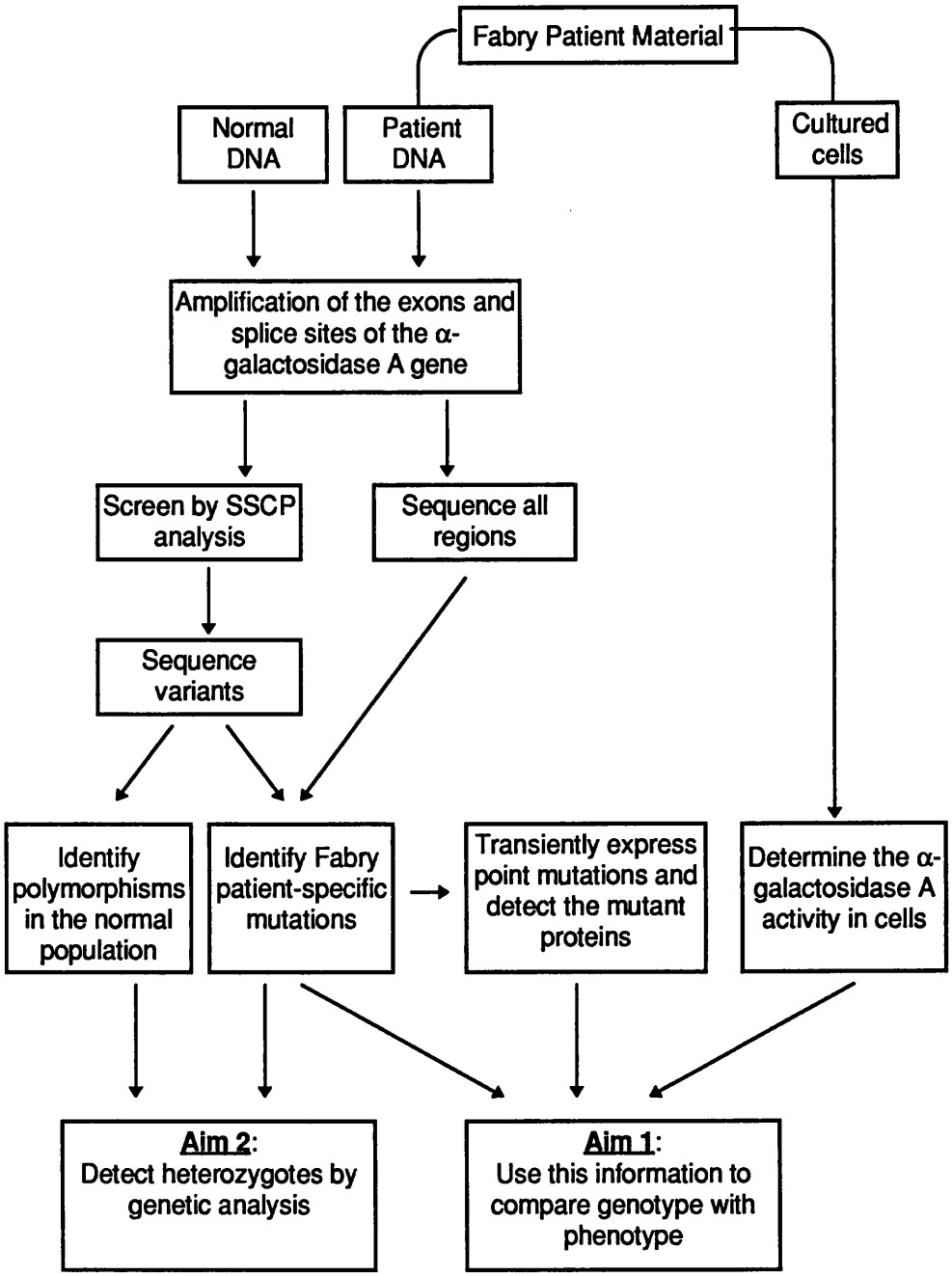
SSCP analysis was chosen as the method for detection of unknown mutations in the α-galactosidase A gene. Although it is less efficient than chemical mismatch detection and does not allow localisation of the mutation within the sequence analysed, the simplicity of SSCP analysis is a distinct advantage. Its sensitivity is higher than RNase cleavage and equivalent to heteroduplex analysis. DGGE is complicated by the use of a gel with a gradient of denaturant and requires expertise regarding the analysis of sequence melting domains and the use of GC-clamps. Genomic DNA was used as the template since most of the samples for analysis were in this form rather than RNA. As there are only 7 exons in the gene, PCR amplification of the protein-coding region in seven reactions is not a problem.

1.4 AIMS AND EXPERIMENTAL STRATEGY

The two aims of this study were: firstly, to find the disease-causing mutations in Fabry hemizygotes, allowing the genotype to be correlated with the observed clinical phenotype and the biochemical changes in the patients; and secondly, to devise an efficient genetic method for detecting female Fabry heterozygotes, whose detection by biochemical means is often inconclusive.

To achieve these aims the experimental strategy indicated in figure 5 was followed. Mutations were detected by screening genomic DNA by SSCP analysis and sequencing the localised region, or by sequencing the entire protein-coding region. To understand the effect of mutations on α-galactosidase A, some mutations were transiently expressed in COS-1 cells and the residual α-galactosidase A in cells from some patients was characterised.

Figure 5: Experimental Strategy



2. MATERIALS AND METHODS

2.1 MATERIALS AND BUFFERS

2.1.1 Material from Fabry families and patient information

DNA, extracted from blood or blood samples were collected from 28 families with at least one Fabry hemizygote. In a further 2 families, families 1 and 30, cultured fibroblasts were the only source of DNA. Cultured fibroblasts from hemizygotes were available for biochemical analysis in families 10, 13, 18, 22, and 28 and lymphoblastoid cells lines were obtained for families 3, 8, 12 and 27. Information regarding the clinical phenotype of the hemizygotes analysed in this study and that of their affected relatives (R) is shown in table 5.

The α-galactosidase A activities in leukocytes and plasma that were measured in the diagnostic laboratories at Institute of Child Health (ICH), information were kindly provided by Mrs E. Young. Enzyme activities quoted by other laboratories are also shown. The clinical symptoms listed in table 5 are: acroparesthesias and episodic pain crises (2); angiokeratoma (3), a characteristic skin rash; ocular involvement (4), including corneal dystrophy, lens abnormalities, conjunctival and retinal vascular lesions; hypohidrosis (5), reduced sweating; proteinuria (6); cardiac involvement (7), including abnormal electrocardiographic, ECG, abnormalities and angina; kidney involvement (8), including decreased creatinine clearance and various stages of renal failure; and neurological involvement (9), excluding pain. It includes cerebrovascular accidents (CVA, strokes), hemiplegia, tetraparesis and vertigo.

The phenotype was classified as atypical if the symptoms in hemizygotes were known to be limited in their range or were obviously milder. Otherwise, they were classified as typical, based on information from hemizygotes and/or heterozygotes, table 5 and below. Only one hemizygote, in family 1, was classified as atypical, because he lacked all clinical symptoms, except for neurological involvement, at age 26. This patient also had low residual α-galactosidase A activity. All other hemizygotes had negligible enzymic activity, as far as is known.

No detailed information regarding the clinical phenotype was available for affected males or females in families 7, 8, 13, 22, 25 and 30.

Family	Location in pedigree chart	Phenotype	α-galactosidase activity		Current Age	Onset age and clinical symptoms									
			Assays at ICH Leukocytes (plasma)	Assays by others		1	2	3	4	5	6	7	8	9	Other symptoms and comments
1*	-	variant		5% in leukocytes, 15% in fibroblasts	26	-	-	N	-	-	N	N	N	Y	Neurological involvement only.
2*	-	typical		low	30	7	Y	Y	-	N	-	N	-	N	-
3*	III-2	typical	1.3 (0.3)		d44	2	Y	Y	-	-	Y	-	Y	Y	Renal transplant. Died of stroke.
	II-5	typical (R)	1.5 (0.5)		46	-	Y	Y	-	-	-	-	Y	Y	-
4		typical		low	33	-	Y	Y	Y	-	-	N	N	-	-
5*	II-1	typical	2.7 (0.6)		49	15	Y	Y	-	Y	Y	-	Y	N	-
6*	II-4	typical	1.3 (0.2)		37	5	Y	Y	-	Y	N	-	N	N	-
	II-7	typical (R)	0.5 (0.16)		37	20	Y	Y	Y	N	N	Y	Y	N	Pedal oedema.
7*	II-1	unknown			No hemizygote information										
8		unknown		low, in plasma	No hemizygote information										
9*		typical			No hemizygote information										
10*	II-1	typical		low	39	-	-	Y	-	-	-	-	Y	Y	Spastic quadraplegia due to stroke
11*	III-5	typical		low, in plasma	15	8	Y	Y	-	Y	N	-	N	N	-
12*	II-1	typical	3.7 (0.54)		27	8	Y	Y	Y	-	N	N	-	-	-
13*	I-4	unknown			No hemizygote information										
14*		typical			No hemizygote information										
15*	III-6	typical	5.8			-	Y	-	-	-	-	N	-	-	
16*	III-2	typical		low, in plasma	10	8	Y	Y	-	Y	N	-	N	N	-
		typical (R)		-	29	-	Y	Y	-	Y	N	N	N	Y	Pedal oedema

Table 5: The phenotypes and enzyme activities of Fabry hemizygotes in 30 unrelated families

1. onset age; 2. acroparesthesias and pain crises; 3. angiokeratoma; 4. ocular involvement; 5. hypohidrosis; 6. proteinuria; 7. cardiac involvement; 8. kidney involvement; and 9. neurological involvement. 'Y' is yes, 'N' no and '-' is unknown for the symptoms and 'd' is the age of death. (R) is a relative of the patient analysed by sequencing. '*' = see text for more family details. Enzyme assays performed at ICH (results from Mrs E Young) were in units of nmol/hr/mg protein for leukocytes, with a normal range of 40-162. In plasma, activity was in nmol/hr/ml of plasma, with a normal range of 4.8-26.5.

Family	Location in pedigree chart	Phenotype	α-galactosidase activity		Current Age	Onset age and clinical symptoms									
			Assays at ICH leukocytes (plasma)	Assays by others		1	2	3	4	5	6	7	8	9	Other symptoms and comments
17		typical			31	-	Y	-	-	-	-	-	Y	-	End stage renal failure.
18*	III-2	typical		< 3% in leukocytes	24	18	Y	N	Y	-	N	Y	N	-	Severe skeletal dysplasia and delayed puberty.
	IV-2	typical		low	11		Y	-	Y	-	-	Y	N	-	Diffuse lack of skeletal calcium
19*	III-1	typical (R)	2.9 (1.4)		29	5	Y	Y	-	Y	N	-	N	N	
20*	IV-41	typical		low	38	9	Y	Y	Y	-	-	Y	N	-	See Hamers *et al.*, 1979
21		typical	3.3 (0.17)		54	9	Y	Y	-	-	Y	N	Y	N	Pedal oedema. Still able to work.
22		unknown	2 (0)		No hemizygote information										
23*		typical		low	11	-	Y	Y	N	-	-	-	-	N	-
24*	I-2	typical (R)			38	6	Y	Y	N	N	Y	Y	Y	N	Oedema of legs
25*		unknown			No hemizygote information										
26*	III-3	typical	1.2 (0.8)		33	22	Y	Y	Y	Y	N	N	Y	N	Pedal oedema
27*	III-2	typical	2 (0.55)		37	5	Y	Y	Y	Y	N	Y	N	Y	-
28*		typical		5%, in plasma	d45	32	Y	Y	Y	-	Y	Y	Y	N	Died of renal failure
29*	II-4	typical	2.9 (3.8)		15	5	Y	N	Y	N	-	-	-	N	Delayed puberty, retarded growth. and gastrointestinal abnormalities.
	II-5	typical (R)	4.0 (1.7)		15	5	Y	N	Y	N	-	-	-	N	Delayed puberty and retarded growth. (Flynn, 1972)
30*		unknown			No hemizygote information										

Table 5 (cont.): The phenotypes and enzyme activities of Fabry hemizygotes in 30 unrelated families

1. onset age; 2. acroparesthesias and pain crises; 3. angiokeratoma; 4. ocular involvement; 5. hypohidrosis; 6. proteinuria; 7. cardiac involvement; 8. kidney involvement; and 9. neurological involvement. 'Y' is yes, 'N' no and '-' is unknown for the symptoms and 'd' is the age of death. (R) is a relative of the patient analysed by sequencing. '*' = see text for more family details. Enzyme assays performed at ICH (results from Mrs E Young) were in units of nmol/hr/mg protein for leukocytes, with a normal range of 40-162. In plasma, activity was in nmol/hr/ml of plasma, with a normal range of 4.8-26.5.

The pedigrees and additional information on families studied

Identification of the family-specific mutation was carried out on DNA samples analysed from the individuals indicated by an arrow and/or indicated in table 5. The α-galactosidase activities in leukocytes (and plasma) of heterozygotes that were analysed by mutation detection in this study were measured in the enzyme laboratory. Activities in the normal range are underlined. The normal range for leukocytes is 40-162nmol/mg protein/hr and for plasma, 4.8-26.5nmol/ml plasma/hr.

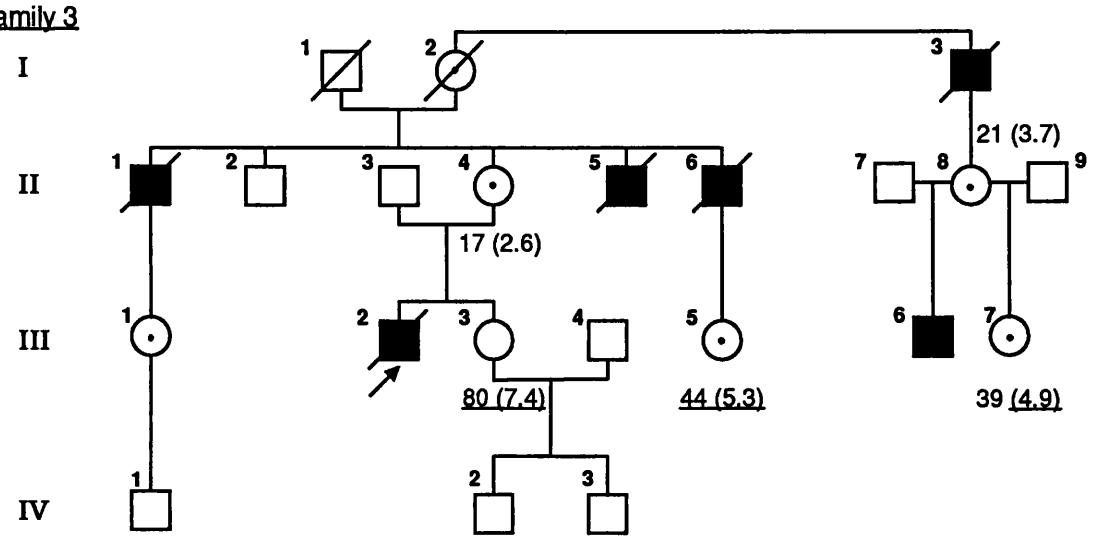
Family 1

The hemizygote analysed had a variant phenotype at age 26, with no angiokeratoma, joint swelling, cardiac or renal symptoms. Cerebral involvement was the major symptom, with hyperdense spots in the basal ganglia, thalamus, cerebellum and capsula internum. He suffered from reversible hemiparesis and recurrent attacks of fever. Both parents at ages 48 and 54 were unaffected.

Family 2

The mother of the hemizygote analysed was symptomatic, with acroparesthesias from age 9-17, which stopped and then returned at age 38.

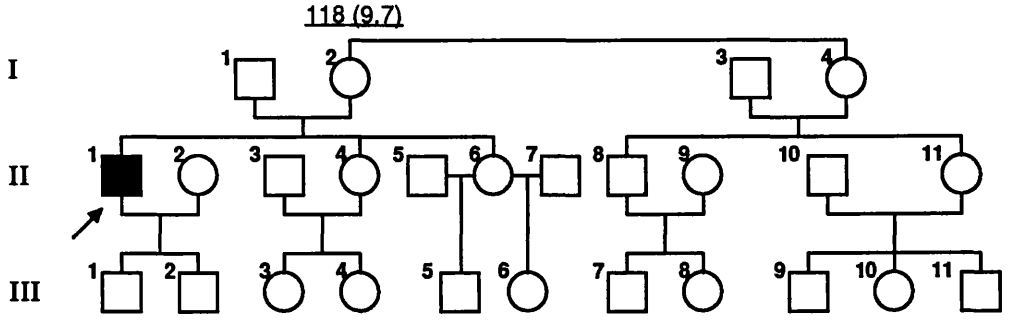
Family 3



Hemizygote II-1 died in his 40s and II-6 died at age 36, from renal failure. Three heterozygotes were symptomatic, one of whom, female II-4, had cardiac

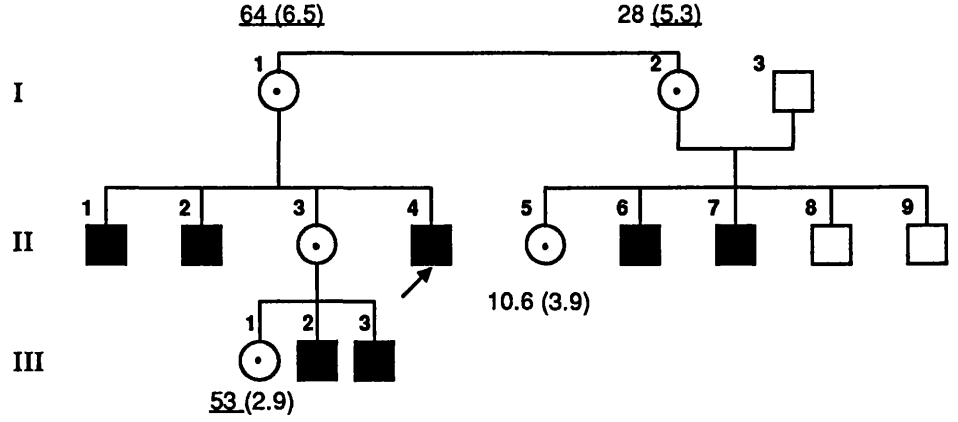
involvement from 55 years and abdominal pains. Female III-5 had a renal transplant in her 30s and III-1 had angiokeratoma and severe acroparesthesias.

Family 5



There was no family history of Fabry disease and hemizygote II-1 was the only male known to have the disease. His mother, I-2, two sisters, II-4 and II-6 and one cousin, II-11, had enzyme activities within the normal range. They had no ocular defects and were clinically normal.

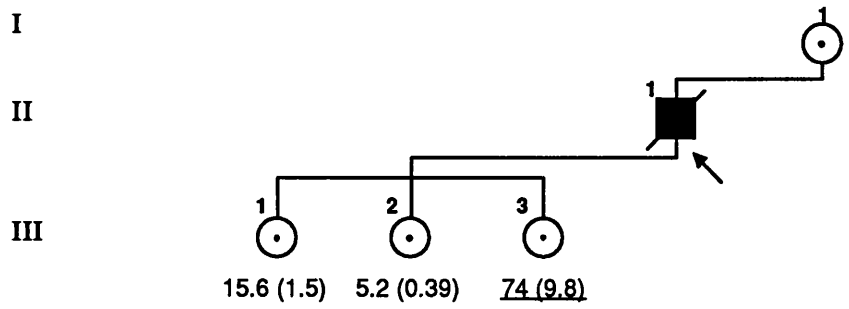
Family 6



Two hemizygotes, II-4 and II-7, both aged 37yr, had a range of symptoms that were typical of Fabry disease but were slightly milder than in other families. In particular, there was no renal involvement in II-4 and in patient II-7 the renal involvement was limited to a decrease in creatinine clearance. In addition, another hemizygote, II-6, at age ~39, continued to work full-time and was known to have acroparesthesia, but no other obvious symptoms. Further information on II-6 was unavailable. Hemizygote III-2, age 13, had episodic pain crises which were triggered by exercise or temperature changes and angiokeratoma was present on one wrist, but he had no other symptoms. At age 8, hemizygote III-3 was

asymptomatic. No Fabry males in this family were known to have had cerebrovascular accidents and none of the females were symptomatic.

Family 7

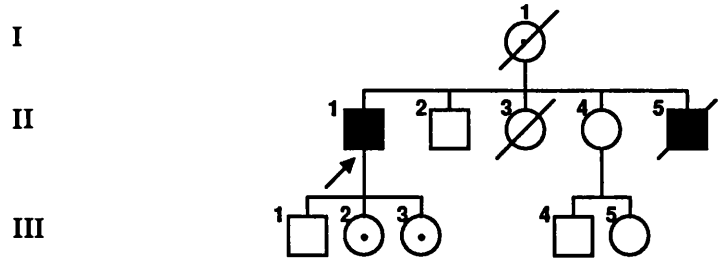


No clinical information was available for this family.

Family 9

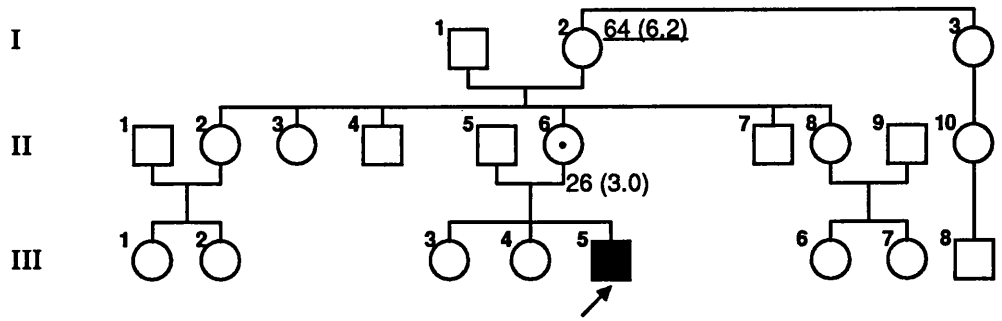
One 30 year old female had pain and ocular abnormalities and her mother had the same symptoms and cardiac involvement, at age 60. These symptoms were typical of Fabry disease. Both symptomatic heterozygotes had α-galactosidase activities in the normal range (communication by Dr M. Beck). No material or clinical information was available from affected males, although they were known to exist.

Family 10



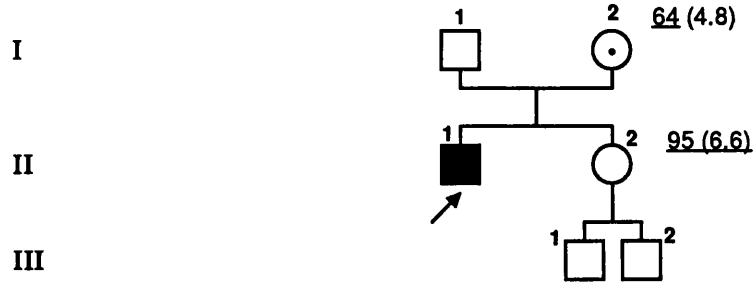
The phenotype of hemizygote II-1 was thought to be typical and his brother, II-5, had similar symptoms and died at age 27 from renal failure. His mother, I-1, had chronic heart failure and renal failure and died at age 34.

Family 11



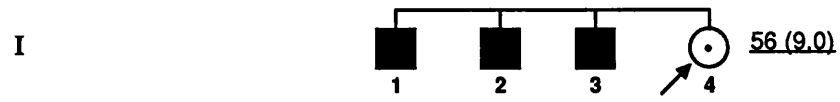
Hemizygote III-5 was the only affected male known to exist in this family. His mother, II-6, was diagnosed as a carrier by enzyme assay and by slit lamp examination, which showed ocular abnormalities.

Family 12



The symptoms of hemizygote II-1, were typical of classic Fabry disease. His mother, I-2, was asymptomatic, except for minor ocular abnormalities.

Family 13



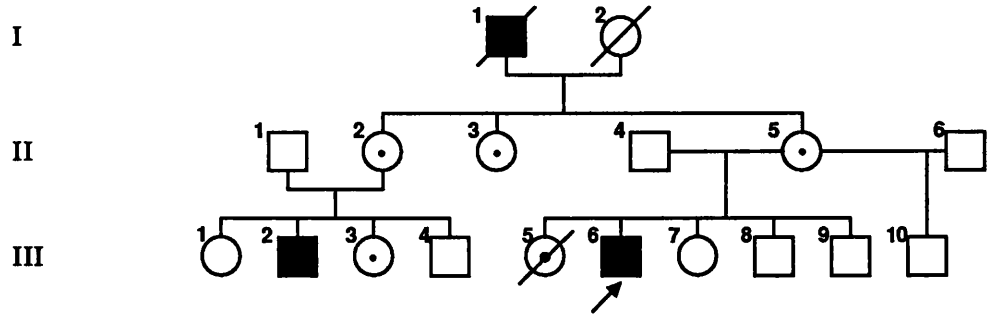
DNA was only available for analysis from a symptomatic female, I-4, whose three brothers had Fabry disease. Details of the phenotype were unknown.

Family 14

A 49 year old female had almost zero enzyme activity in her white blood cells, cardiomyopathy, vascular problems and limb pains. Her daughter, age 27, had severe renal problems but normal α-galactosidase activity in white blood cells. These symptoms were indicative of a typical phenotype for Fabry disease. No

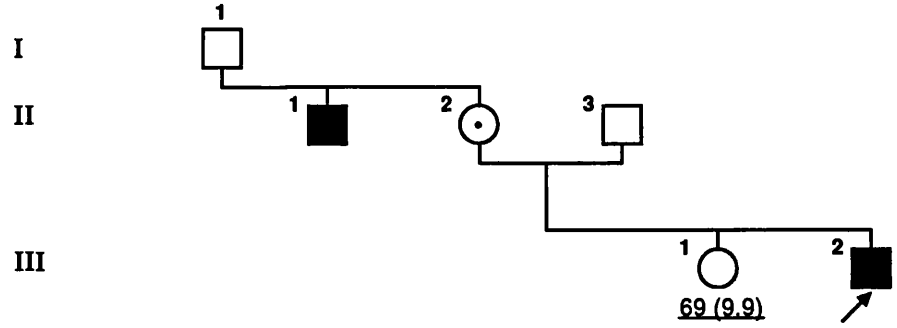
hemizygotes were known in this family and the father of the 49 year old female was unknown.

Family 15



The information regarding hemizygote III-6 analysed in this family was limited. However, a sister, III-5, died at 18 years due to a stroke, his mother, II-5, had nerve pains and an aunt, II-2, required a renal transplant. This indicated that the disease phenotype was typical.

Family 16

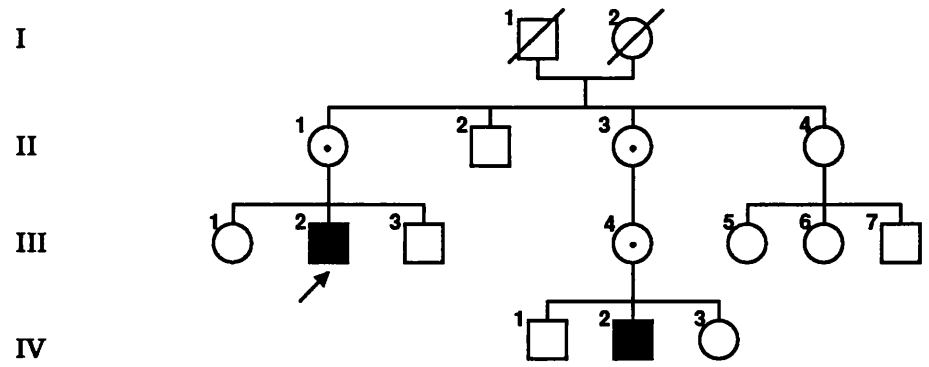


The hemizygote analysed, III-2 and his uncle, II-1 had typical symptoms.

Family 17

The information about this family was limited. The mother, age 42 and grandmother, age 70, of the hemizygote analysed, were asymptomatic. His sister, age 20, had corneal dystrophy. The clinical assessment of the hemizygote analysed was diagnostic of Fabry disease but no confirmatory enzyme assays were carried out. The phenotype was thought to be typical, although the details were not available.

Family 18



Hemizygote III-2 was unusual in that he had a skeletal dysplasia and delayed puberty, rare symptoms in Fabry disease. At 24 years old he did not have angiokeratoma but did have other typical symptoms, such as cardiac involvement and limb pain. A hemizygous relative, IV-2, did not have skeletal dyplasia but had a diffuse lack of skeletal calcium, at age 11. These males were classified as typical but severely affected.

Family 19

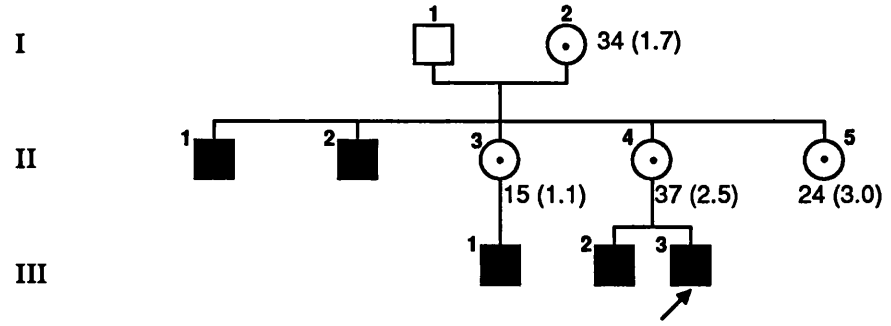


Table 5 gives the clinical details for hemizygote III-1, the cousin of an affected child, III-3, who was diagnosed enzymatically. Heterozygote II-3 had angiokeratoma and acroparesthesias and the child's heterozygous grandmother, I-2, had a renal transplant at age 46, indicating that renal abnormalities were a feature. The known hemizygotes in this family did not have renal involvement but were still young.

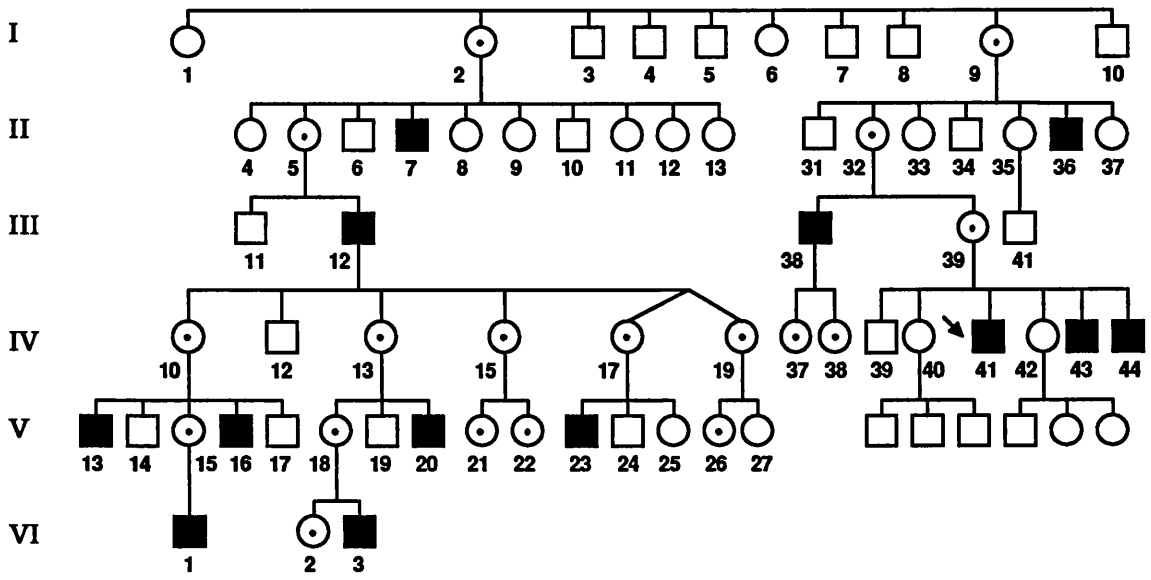
Family 20 (Johnston *et al.*, 1966; Hamers *et al.*, 1979)

The hemizygote analysed in this study, IV-41, was part of a family with Fabry disease inherited for at least 6 generations (Johnston *et al.*, 1966; Hamers *et al.*, 1979). This hemizygote was the oldest of 9 hemizygotes in generations IV-VI, figure 6. Six of these 9 hemizygotes had varying degrees of limb pain, angiokeratoma and corneal abnormalities. In addition, one, V-13, had slow growth, flexed fingers and oedema. Another had slow growth and one had haematuria. Two other hemizygotes were examined by Johnston *et al.*, 1966. Hemizygote III-12 had pains during childhood and developed extensive angiokeratoma. At age 42, kidney involvement was noted and he died at age 49 from chronic nephritis and cardiac failure. Hemizygote III-38 had pains from age 5 but was relatively mildly affected with almost no angiokeratoma. At age 48 he suffered from cerebral thrombosis but recovered and a further episode occurred at age 55 but he continued to work at age 59. He also had oedema of the legs and ocular abnormalities. All of the hemizygotes from generations III-V had blood groups A_1, except for III-2 and IV-44 who had groups O and A_2, respectively. The blood groups of V-23, VI-1 and VI-3 were unknown.

Thirteen heterozygotes were identified by hair root analysis and family history (figure 6). Ten of these heterozygotes, between the ages of 19 and 66, were clinically affected. One other heterozygote, IV-17, in this age range, age 36 was asymptomatic and two other heterozygotes, V-26 and VI-2 were unaffected at ages 9 and 3, respectively. No renal involvement was seen in the hemizygotes, except for III-38, or heterozygotes, except for IV-38 in whom nephrotic syndrome and proteinuria were observed from age 18 to 20 but did not reappear. In the study by Hamers, *et al.* (1979), analysis for cross-reacting immunological material (CRIM) in hemizygotes IV-41 and V-13 showed that there was no material in the urine that cross-reacted with antiserum raised against α-galactosidase A.

Later clinical investigations have found that female VI-2 had occasional pains in the fingers and toes but was otherwise well, at age 19. Hemizygote VI-3 had angiokeratoma from age 6 and pains in his legs. Other clinical symptoms, such as severe lymphoedema in the leg of one hemizygote, cardiomyopathy in another and a cerebrovascular accident in a 50 year old heterozygote, have been noted, although the positions of the affected individuals in the pedigree chart were unknown. Abdominal pains and indigestion were also a feature in some members of this family.

Figure 6: Pedigree and clinical features in family 20 (Johnston *et al.*, 1966; Hamers *et al.*, 1979)



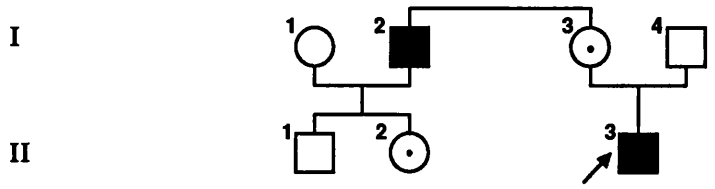
Location in pedigree chart	Age (yr)	Pain	Skin lesions	Abnormal ECG	Cornea verticillata	Other symptoms
Hemizygote						
IV-41	38	+	+	+	+	Slow growth, angina, phlebitis
IV-43	29	+	+	-	+	Diffuse lunules
IV-44	20	+	+	-	+	
V-13	32	+++	+	-	+	Slow growth, flexed fingers, hypohidrosis
V-16	25	+++	+	-	+	Hypohidrosis
V-20	17	++	+	-	+	Slow growth
V-23	15	+	+	-	+	Haematuria
VI-1	2	-	-	-	ND	
VI-3	1	ND	+	ND	ND	
Symptomatic heterozygote						
III-39	66	+	-	-	+	
IV-10	53	+	+	+	+	
IV-13	43	+	+	+	+	
IV-15	38	-	-	-	-	Tortuous retinal veins
IV-19	36	-	+	-	+	
IV-38	44	++	-	+	+	Fever, proteinuria, nodular vascularitis, oedema
V-15	28	+	-	-	ND	
V-18	22	-	+	-	+	
V-21	21	-	+	-	-	
V-22	19	++	+	-	+	

Severity of pain was expressed in categories varying from absent (-) to very severe (+++). Other clinical features were evaluated as either absent (-) or present (+). ECG is electrocardiogram. In the pedigree diagram the arrow marks the hemizygote analysed by DNA sequencing.

Family 23

The mother of the hemizygote analysed was asymptomatic. His maternal uncle was known to have died from Fabry disease.

Family 24

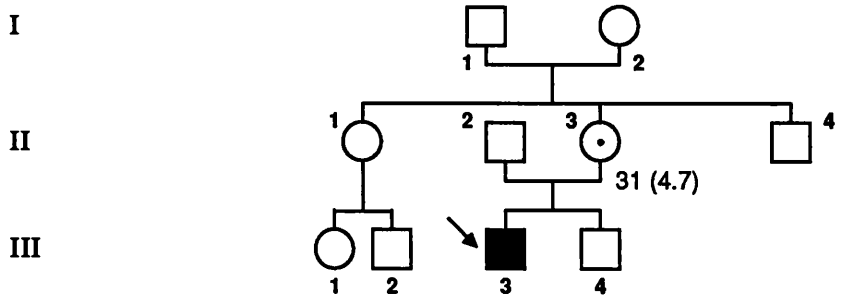


Hemizygote I-2 was diagnosed by histological examination of material from renal and skin biopsies. He had symptoms typical of Fabry disease. DNA from his nephew, II-3, age 19 was analysed. II-3 was clinically affected at age 7, with burning pains in his hands and feet. Two females, I-3 and II-2 were symptomatic for Fabry disease.

Family 25

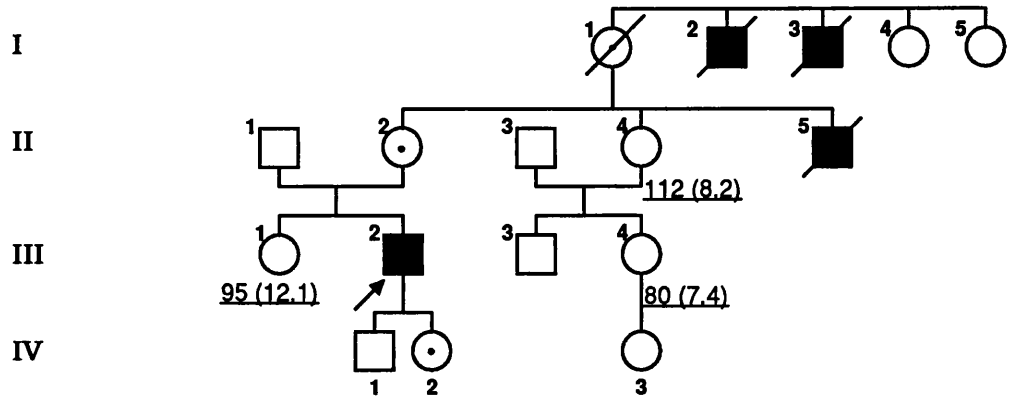
The phenotype of the hemizygote analysed, was unknown. However, two maternal uncles died in their 30's of kidney failure. The phenotype was suspected of being severe.

Family 26



One affected hemizygote, III-3 was known in this family and no other males or females were known to be clinically affected in this family.

Family 27

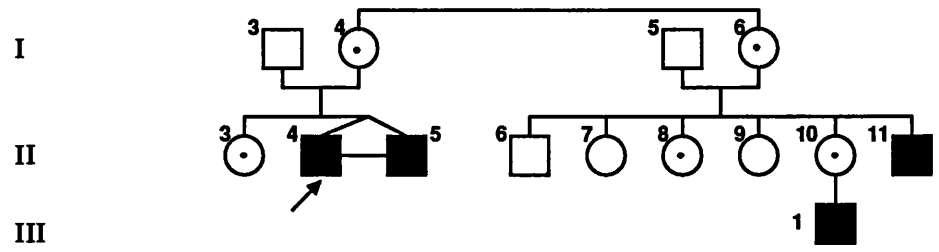


The patient material came from a 37 year old male, III-2, in the third generation of a Fabry family. He had 6 strokes between the ages of 18 and 35, heart murmur, gastrointestinal disturbances, such as diarrhoea, vomiting and abdominal pain but no renal involvement. In the first generation of this family, 2 males, I-2 and I-3, died in their 40s of kidney complications and nephritis. The maternal uncle, II-5, of the analysed patient, was known to have had angiokeratoma, neurological complications such as recurrent strokes, deafness, slurred speech and impaired renal function at age 53. His mother had a myocardial infarction at age 42 but no neurological complications.

Family 28

The hemizygote analysed was the only affected family member known to exist. No symptomatic heterozygotes were known.

Family 29 (Flynn *et al.*, 1972)



The clinical symptoms in the 15 year old, hemizygous monozygotic twins, II-4 and II-5 have been published, Flynn *et al.*, 1972. Heterozygotes were identified by α-galactosidase assays on leukocytes and chromatography of urinary ceramide hexosides. Females I-4, I-6, II-3, II-8 and II-10 were carriers, aged 41, 39, 17, 12 and 10, respectively. All of these, except for I-6, had ocular involvement. Female

II-3 had episodic leg pains at age 17, while female I-6 had leg pains from 12-23 years old.

Family 30

The maternal grandfather of the hemizygote analysed in this study was believed to have died from Fabry disease at age 37. The daughter of this hemizygote had limb pains and corneal dystrophy and his sister required a renal transplant. No other clinical details were available.

Ethnic Origin

Families 2-8, 10-13, 15-17, 20-21, 23 and 25-27 were believed to be of British origin, based on knowledge of their family names and because there was no information to suggest that they were from any other ethnic group. Family 19 lived in Britain but were known to be of Asian origin. The remaining families were referred by clinicians from other European countries and there was no information to suggest that their ethnic origin was different from the country in which they lived. Families 1, 9 and 14 were German, family 18 was Dutch and family 24 was Belgian.

2.1.2 Reagents

2.1.2.1 Chemical reagents

General chemical reagents were supplied by Sigma, British Drug Houses, or Fisherbrand (USA) and were of AnalaR grade or equivalent. All solutions were prepared in singly distilled water, except where specified otherwise. Mammalian cell culture reagents were from GibcoBRL or Sigma and ultrapure grade agarose was from GibcoBRL. Low melting point SeaPlaque agarose was supplied by Flowgen. Difco labs (USA) supplied Bacto-Agar. Qiagen plasmid preparation kits, the Sequenase II kit (USB) and Amplify™ solution were from Amersham International plc. Dynabeads (M-280) were from Dynal. Stratagene provided the Exo-*Pfu* cycle sequencing kit. Radiolabelled nucleotides were from ICN Biomedicals Ltd. AccuGel™ and ProtoGel™ were from National Diagnostics. Micro detergent was from International Products Corp. and Pharmacia provided the deoxynucleotides and dideoxynucleotides. Centricon 100 and 30 columns were obtained from Amicon. All solutions were autoclaved or filter-sterilised.

2.1.2.2 Enzymes

Restriction enzymes were from Northumbria Biologicals Limited, New England Biolabs, or CP Biolabs. Boehringer Mannheim and Promega supplied the calf intestinal phosphatase and *T4* DNA ligase was from New England Biolabs. *Taq* DNA polymerase was from Promega or Bioline and *Vent* DNA polymerase was from New England Biolabs.

2.1.2.3 Plasmids and size markers for protein and DNA

The 1kb ladder contained DNA fragments in the size range of 0.75kb-12.2kb and was from GibcoBRL. The plasmid vector pGEM-3Z was from Promega and pCMVβ was from Clontech. The [^{14}C] methylated rainbow markers (from Amersham) were radiolabelled and coloured proteins of known size in the range of 14.3-200kDa.

2.1.2.4 Others

Anti-(α-galactosidase A) serum (IgG-purified) and pGB6 (α-galactosidase A full length cDNA) were kindly provided by Dr RJ Desnick, Mount Sinai Medical Center, New York. Kodak X-OMAT AR-5, X-ray film was from Eastman Kodak Co. USA. *E. coli* strain DB10B™ cells were from GibcoBRL.

2.1.3 Common solutions

The following solutions were prepared in dH_2O.

(a) 5x TAE (Tris Acetate EDTA) buffer	200mM Tris base 5.71ml/l (v/v) of acetic acid 50mM EDTA (pH 8)
(b) 5x TBE (Tris Borate EDTA) buffer	45mM Tris base 45mM Boric acid 10mM EDTA (pH 8)
(c) TE (Tris EDTA) buffer at pH 8.0	10mM Tris-HCl (pH 8) 1mM EDTA (pH 8) Adjusted to pH 8 with NaOH

2.2 DETECTION OF SEQUENCE CHANGES IN GENOMIC DNA

2.2.1 Mammalian cell culture

All operations involving mammalian cells in culture were conducted in a sterile hood, using sterile implements and with maximum care to prevent infection (Freshney, 1994).

2.2.1.1 Lymphoblastoid cell culture

RPMI 1640 growth medium (Moore *et al.*, 1967) was prepared to contain 10% (v/v) fetal calf serum (GibcoBRL), 4mM glutamine solution and either 50μg/ml gentamycin sulphate (Sigma) or 100U/ml penicillin with 100μg/ml streptomycin (GibcoBRL).

Lymphoblastoid cells were grown in a 75cm^2 flask (Marathon), in 10-20ml medium and at 37°C, with 5% CO_2 in an IR-1500 automatic CO_2 incubator (Flow Laboratories). The cells form aggregates which are easily visible by eye and settle but do not adhere to the bottom of the flask. The medium was replaced every 2-3 days by careful aspiration, leaving the cells in a 1ml volume before adding fresh medium.

Once the cell aggregates covered the flask base the medium was removed by aspiration, leaving the cells in about 1ml of medium. The cells were either subdivided for further growth or they were harvested, as follows. They were transferred to a centrifuge tube and spun at 500 x *g* for 5 minutes. The supernatant was discarded and the cell pellet gently resuspended and washed twice in 0.9% (w/v) NaCl solution, using centrifugation to pellet the cells. The cell pellet was either used directly or stored at -70°C for DNA extraction (2.2.2). Alternatively, if the cells were required for enzyme assay they were resuspended in an appropriate buffer prior to direct use or storage at -70°C.

2.2.1.2 Fibroblasts

Growth medium was prepared as for the lymphoblastoid cells, except that Ham's-F10 medium (Ham, 1963) was sometimes used in place of RPMI 1640. Phosphate buffered saline (PBS) was prepared by dissolving 8g NaCl, 0.2g KCl, 1.44g Na_2HPO_4 and 0.24g KH_2PO_4 in a volume of 1 litre of dH_2O, adjusted to pH 7.4 with HCl.

The growth conditions were the same as for lymphoblastoid cells except that they grow as monolayers that adhere to the bottom of the flask. Their growth was checked by viewing under a light microscope and the medium was removed completely and replaced every 2-3 days.

The cells were grown to about 80% confluency before harvesting, as follows. The medium was removed completely by aspiration and the cell monolayer was washed with 2-3ml of either PBS or 0.1% (w/v) trypsin solution (GibcoBRL or Sigma). Trypsin solution was then added and the cells incubated at 37°C for about 5 minutes. The cells were dislodged into solution by hitting the side of the flask. They were then subdivided into new flasks with fresh medium for further growth. Alternatively, the cells were diluted in medium, centrifuged, washed with 0.9% NaCl and stored in the same manner as lymphoblastoid cells, prior to DNA extraction and enzyme assay.

2.2.1.3 Storage of cells in liquid nitrogen

One confluent flask (75cm^2) of lymphoblastoid cells or fibroblasts was sufficient for 2 ampoules of frozen cells. Lymphoblastoid cells, and trypsinised fibroblasts were centrifuged at 500 x *g* for 5 minutes. The cell pellet from 1 flask of cells was immediately resuspended in 1ml of ice-cold medium. An equal volume of ice-cold, 20% (v/v) DMSO (diluted in medium) was added to the cells and 1ml aliquots were placed in plastic ampoules. The ampoules were immediately wrapped in several layers of paper towels and placed inside a 2-3cm thick polystyrene box before leaving them at -70°C overnight, allowing them to cool at a rate of about 1°C per minute, thus limiting cell damage. They were transferred to liquid nitrogen for long-term storage (Lovelock *et al.*, 1959).

2.2.2 Isolation of genomic DNA

Genomic DNA was isolated from 5-10ml of whole human blood by a method using salt precipitation of protein (Miller, SA *et al.*, 1988). Standard safety procedures were followed, with initial operations performed in a safety hood. Nuclei lysis buffer contained 10mM Tris-HCl, 400mM NaCl, 2mM Na-EDTA and was stored at 4°C. Proteinase K solution was prepared freshly to contain 2mg/ml of proteinase K in 2mM Na_2-EDTA, 1% (w/v) SDS, or it was thawed from storage at -20°C.

Venous blood was collected in 10ml plastic tubes containing EDTA to prevent coagulation. These were stored at -20°C or -70°C, prior to DNA extraction. Frozen blood (10ml) was thawed quickly, avoiding a temperature rise above 4°C, and transferred to a 50ml Falcon tube. The red blood cells were lysed by adding ice-cold dH_2O to increase the volume to 50ml and the white cells, containing DNA, were separated by centrifuging at 1200 x *g* for 20 minutes, at 4°C. The leukocyte pellet was resuspended in 25ml of ice-cold 0.1% (v/v) Nonidet-P40, vortexing lightly, if necessary. Disrupted cell membranes were separated from intact nuclei by centrifuging at 1200 x *g* and 4°C for 20 minutes. The supernatant was carefully decanted and the tube inverted over a paper towel to remove excess liquid from the pellet of nuclei. The nuclei were lysed by suspension in 3ml of nuclei lysis buffer, which was vortexed lightly to disperse the pellet. 600µl of proteinase K solution and 200µl of 10% (w/v) SDS in dH_2O were added and mixed. Protein digestion was performed at 60°C for 1-2 hours or at 37°C, overnight.

Digested protein was precipitated with 1ml of saturated ammonium acetate (9.6mol/l). The sample was shaken vigorously for a few seconds and allowed to stand at room temperature for 10-15 minutes before centrifugation at 1200 x *g* and room temperature, for 20 minutes. The supernatant was carefully decanted into a new 50ml Falcon tube, avoiding transfer of the protein pellet. Two volumes of cold absolute ethanol was gently mixed with the supernatant and the precipitated DNA threads were spooled out on the tip of a heat-sealed, glass, Pasteur pipette. The DNA was redissolved in 0.5-1ml of TE buffer.

This method was scaled down accordingly when smaller volumes of blood were used. For extraction of genomic DNA from cultured fibroblasts or lymphoblastoid cell pellets, the first steps were omitted and the method began with the addition of nuclei lysis buffer. For approximately two confluent, 75cm^2 flasks of fibroblasts or

one 75cm² flask of lymphoblastoid cells, the volumes were scaled down by a factor of 0.2.

Extraction of genomic DNA from chorionic villus samples also began with the addition of nuclei lysis buffer, but the fibrous tissue was first cut into small pieces with a pair of scissors and then a pestle and mortar used to increase the tissue surface area sufficiently for complete digestion with proteinase K.

2.2.3 Preparation of the oligonucleotide primers

2.2.3.1 Design of the primers

The PRIMER program, provided by the HGMP Resource Centre, based at the MRC Clinical Research Centre, Harrow, UK, was used to design primers located in the introns of the α-galactosidase A gene. The 12kb Human α-galactosidase A gene sequence (accession number X14448) was obtained from the GenBank database and divided into sequence lengths of less than 5000 bases, the maximum length of sequence accepted by the PRIMER program.

The primers allowed amplification of each of the 7 exons in the gene and approximately 50-100 bases of flanking intron region, providing a sufficient distance between the primer and splice sites for sequencing. Selection parameters were set to find primer pairs with the following attributes: a primer length of 18-22 bases (optimally, 20 bases); melting temperatures within the range of 57-63°C (optimum at 60°C); guanine and cytosine base content of 20-80%; maximum number of 12 bases in one primer with self-complementarity and a maximum of 8 of these at the 3' end; and a pair of primers which do not form heterodimers.

These parameters helped select against primers which form strong self-pairing secondary structures or the formation of primer-dimers which decrease the product yield. It assisted in finding primers that gave a single product.

2.2.3.2 Synthesis of primers

Unless stated otherwise, all primers were synthesised on an Applied Biosystems 381A DNA synthesiser and 5'-biotinylation was carried out using DMT-biotin-C-6-phosphoamidite (CRB), according to the manufacturers instructions. I am very grateful to Paul Rutland, who carried out the syntheses.

Primers were precipitated from ammonia solution by adding 0.1 volumes of 3M sodium acetate solution (pH 7.0) and 3 volumes of absolute ethanol together. The mixture was cooled at -70°C for 0.5 hours and then centrifuged at 10000 x *g* for 15 minutes. The primer DNA pellet was washed with 70% (v/v) ethanol and centrifuged again, before drying in air for 20 minutes and resuspension in ddH_2O or TE buffer.

2.2.3.3 Estimation of the concentration and purity of the DNA primers

The concentration of DNA was estimated by measuring the optical density (OD) of the sample at a light wavelength of 260nm, using a Gene Quant (Pharmacia) spectophotometer. If the GC:AT ratio is approximately 1:1, an OD unit of 1 is estimated to represent a concentration ('X') of 37µg/ml of single stranded primer DNA. For double-stranded DNA, X = 50µg/ml (Sambrook *et al.*, 1989). Therefore,

$$\text{sample concentration (µg/ml)} = \text{OD (at 260nm)} \times \text{'X'}$$

The molar concentration of single-stranded primer DNA was calculated using the 'primer'program written by Paul Rutland and which used the following calculation:

$$\text{DNA (pmol/µl)} = [OD_{260nm} \times 1000]/[G(12010) + A(15200) + T(8400) + C(7050)]$$

where G, A, T and C are the number of guanine, adenine, thymine and cytosine residues in the primer sequence, respectively and each is multiplied by the molar extinction coefficient of its nucleotide.

The purity of the DNA in solution was estimated by measuring the optical density at 260nm and 280nm. A 260nm/280nm optical density ratio of 2 is estimated to indicate 100% purity.

2.2.4 Amplification of DNA by the polymerase chain reaction

DNA was amplified from a double-stranded DNA template using the polymerase chain reaction (PCR, Saiki *et al.*, 1985). A 50µl or 100µl reaction mixture was prepared in a 0.5ml Eppendorf tube and contained the following, except where stated otherwise: 100-500ng of genomic DNA template, 1-5 units of thermostable *Taq* DNA polymerase, 1x reaction buffer, 1.5-2.5mM $MgCl_2$, 0.2mM each of dGTP, dATP, dTTP and dCTP, and 5-100pmoles of each primer. The 10x reaction buffer from Bioline contained 0.1M Tris-HCl (pH 8.8), 0.5M KCl and 1% Triton X-100 and from Promega it contained 0.1M Tris-HCl (pH 9.0) and 0.5M KCl. The reaction mixture was overlaid with 50µl of mineral oil to prevent its evaporation. For each set of reactions a control reaction mixture, lacking template DNA, was prepared to check for contamination.

The PCR reaction was achieved by performing 30 cycles of: denaturation at 94°C, followed by annealing short oligonucleotide primers at a temperature dependent on the specific base composition of the primer and finally, an elongation step at 72°C. In addition, a 5 minute denaturation step at 94°C preceeded the cycling stage and a 10 minute step at 72°C was included at the end of 30 cycles, to ensure complete elongation of the newly synthesised product. The conditions were optimised to yield a specific product by varying the concentration of $MgCl_2$ and the primer annealing temperature.

The specific primer pairs and conditions indicated in Table 6 were used to amplify each of the seven exons and their flanking intron regions in the α-galactosidase A gene as separate products.

PCR amplification was also used to amplify the exonic regions of the α-galactosidase A gene in four fragments which contained exons 1, 2, 3-4 and 5-7. The conditions and primer pairs are indicated in table 7 and were provided by Dr R. J. Desnick, Mount Sinai Medical Center, New York. The annealing temperature was varied between 56-60°C, when necessary.

Exon	Sequence	Gene nucleotide position	Anneal Temp. (oC)
1	5'-GTCCCAGTTGCCAGAGAAAC-3' 5'-AAAGGGAAGGGAGTACCCAA-3'	1054-1073 1403-1384	58
2	5'-ATGGGAGGTACCTAAGTGTTCA-3' 5'-GTCCTCTGAATGAACAAGAACA-3'	5033-5054 5306-5285	59
3	5'-GATGATTTTGGGGGTTTGTG-3' 5'-GATTGGTTCTTTGGCTCAGC-3'	7123-7142 7522-7503	58
4	5'-ATAGCCCCAGCTGGAAATTC-3' 5'-CAGTTCTATTGGATTCTGGGC-3'	8262-8281 8445-8425	59
5	5'-AGAAGGCTACAAGTGCCTCC-3' 5'-GCAGGGTCTTGAACAAGGAG-3'	10042-10061 10327-10308	61
6	5'-GGATGCTGTGGAAAGTGGTT-3' 5'-GGCCCAAGACAAAGTTGGTA-3'	10413-10432 10769-10750	61
7	5'-GAATGCCAAACTAACAGGGC-3' 5'-GCATGAGCCACCTAGCCTT-3'	10912-10931 11352-11334	61

Table 6: Intronic primers for amplification of single exons in the α-galactosidase A gene

PCR amplification was with 30 cycles of: 94°C for 30 seconds, annealing for 30 seconds and 72°C for 1 minute in 0.5ml Eppendorf tubes in a Techne PCR machine (PHC-2 or PHC-3) or in a Biometra TRIO thermoblock. The reaction volume was 50 or 100µl with 5-50pmol of each primer, 1 unit of *Taq* DNA polymerase (Promega or Bioline) and buffer with 1.5mM $MgCl_2$.

Exons	Sequence (5'->3')	Gene nucleotide position	Anneal Temp (°C)
1	Bio- GGATCACTAAGGTGCCGC AACTGTTCCCGTTGAGACTC	926-943 1460-1441	60
2	Bio- CTTGTGATTACTACCACACT AACAAGCTTCTGTACAGAAGTGC	4968-4987 5334-5312	58
3-4	TAGCTCAGCAGAACTGGGGGATTT Bio- GTAAGTAACGTTGGACTTTG	7071-7094 8510-8491	56
5-7	GTTTAGACCTCCTTATGGAGA Bio- CAGGAAGTAGTAGTTGG	10000-10020 11301-11285	58
5-7	Bio- CTTACAAGGATGTTAGT CAGGAAGTAGTAGTTGG	10078-10094 11301-11285	58

Table 7: Intronic primers for amplification of exons 1, 2, 3-4 and 5-7 in the α-galactosidase A gene

PCR amplification was with 30 cycles of: 94°C for 30 seconds, annealing for 45 seconds and 72°C for 1 minute in 0.2ml thin-walled tubes in a Geneamp PCR system 9600 (Perkin Elmer). The reaction volume was 100µl with 1µM of each primer, 5 units of *Taq* DNA polymerase (Promega) and buffer with 2-2.5mM $MgCl_2$. Bio indicates that the primer was biotinylated at the 5' end.

2.2.5 Analysis of DNA by agarose gel electrophoresis

All PCR products, approximately 200-1500 bases, were analysed by electrophoresis on an agarose gel (Sambrook *et al.*, 1989) to verify the success and specificity of the reaction.

The gel contained 0.7-2% (w/v) agarose which was melted in 1x TBE and allowed to cool to about 50°C before addition of ethidium bromide solution to a concentration of 0.5µg/ml. The agarose was poured into a horizontal mould and allowed to set. Sample loading dye contained 50% (v/v) glycerol, 0.1mM EDTA, pH 8 and 0.2% (w/v) orange G dye or bromophenol blue.

DNA samples were mixed in a ratio of 4:1 with loading dye. The DNA was resolved by electrophoresis at 50-125V with not more than 5V per centimetre distance between the electrodes in 1x TBE buffer. The product size was compared with a known DNA size marker, a 1kb Ladder. The electrophoresis apparatus used was either a 14cm by 15cm Midigel system (NBL) or a 8cm by 10cm minigel GNA-100 apparatus (Pharmacia). Agarose gels were visualised under UV light and photographed using a Mitsubishi videocopy processor.

2.2.6 Single strand conformation polymorphism analysis

Single strand conformation polymorphism (SSCP) analysis (Orita *et al.*, 1989; Hayashi, 1991) was used to detect sequence changes that alter the migration of DNA through a non-denaturing polyacrylamide gel.

2.2.6.1 Radioactive SSCP detection of unknown sequence changes

2.2.6.1.1 Preparation of samples for SSCP analysis

Genomic DNA was amplified by the PCR (2.2.4), in the presence of radiolabelled nucleotides to allow detection by autoradiography. The PCR was carried out in a 50µl reaction volume, using 40pmol of each primer (table 6), with substitution of the unlabelled 0.2mM dCTP with 1µCi of [α-^{32}P]dCTP (0.1µl of [α-^{32}P]dCTP, 3000Ci/mmol, ICN) and 0.02mM of unlabelled dCTP.

The PCR-products were diluted by 4-8 fold in a solution of 0.1% (w/v) SDS and 10mM EDTA. For each diluted sample a volume of 4µl was mixed with an equal volume of formamide solution (95% (v/v) deionised formamide; 0.05% (w/v) bromophenol blue; 0.05% (w/v) xylene cyanol; and 20mM EDTA). This mixture was denatured for 2-3 minutes at 94°C before placing on ice, prior to electrophoresis.

2.2.6.1.2 Analysis of DNA by non-denaturing polyacrylamide gel electrophoresis

A non-denaturing polyacrylamide gel mixture was prepared using a stock solution of liquid acrylamide which contained 30% (w/v) acrylamide and 0.8% (w/v) methylene bisacrylamide in dH_2O (ProtoGel™). The stock solution was diluted five fold to contain 6% (w/v) acrylamide, 5% (v/v) glycerol and 1x TBE buffer, in distilled water. This mixture could be stored at 4°C in the dark for several weeks.

Electrophoresis was carried out in a model S2 sequencing gel system (GibcoBRL), which holds a gel 30cm wide x 40cm long and 0.4mm thick. First, the glass plates were washed with dilute detergent (Micro, International Products Corp.) and rinsed with tap water before washing with 70% (v/v) IMS to dry the plates. The smaller of the two plates was then siliconised, in a fume hood, by covering one side with either Repelcote (BDH) or Sigmacote (Sigma) solution and allowing it to dry for about 10 minutes. The plates, separated by 0.4mm spacers arranged at the sides and one at the bottom were clamped together with bulldog clips to form the gel mould.

Polymerisation of 70ml of polyacrylamide gel mixture was induced by adding 0.5ml of fresh 10% (w/v) ammonium persulphate in dH_2O, to promote free radical formation, and 60µl of TEMED to catalyse the cross-linking. Care was taken not to shake the gel mixture too vigorously before pouring, since oxygen inhibits polymerisation. The gel was poured immediately and combs with 1cm width wells were inserted at the top. It was allowed to set for at least 1-2 hours before use.

Denatured DNA samples (6-8µl), section 2.2.6.1.1, were loaded onto the gel and electrophoresed overnight at 320-400V, at room temperature. For PCR products of less than 250 bases (exons 2 and 4) the gel was run until the xylene cyanol had migrated about 20cm down the gel, while larger fragments of up to 450 bases (exons 1, 3, 5, 6 and 7) were run until the xylene cyanol was about 30cm down the gel.

2.2.6.1.3 Detection of radiolabelled DNA by autoradiography

Polyacrylamide gels were placed on 3MM Whatman paper, covered in cling film and dried under vacuum at 80°C for 0.5-2 hours on an ATTA gel dryer (model AE3700). The cling film was removed from the dried gel. The gel was placed in a light-proof cassette and exposed to X-ray film, Kodak X-OMAT AR-5, for a suitable length of time (1-7 days), at -70°C. Intensifying screens were not included. The film was processed using a Fuji Photo Film Ltd. developer.

2.2.6.2 Detection of polymorphisms by non-radioactive SSCP

PCR amplification (2.2.4) was used to produce DNA for exons 1 and 3, in a 50µl reaction volume with 40pmol of each primer (table 6).

Polyacrylamide gels, 16cm wide x 20cm long and 1mm thick, were prepared for electrophoresis using the Midigel Protean II apparatus (BioRad). A vertical gel mould was formed by two glass plates which were held apart by 2 side-spacers and with a rubber seal at the gel base. The glass plates were washed twice with dilute detergent (Micro) and then with 70% (v/v) IMS to dry the plates, before use.

The non-denaturing polyacrylamide gel mixture was prepared as described (2.2.6.1.2) but the concentrations of glycerol and polyacrylamide were varied over the ranges 0-5% (v/v) and 6-18% (w/v), respectively, to determine the optimum concentrations and to maximise the differences in migration of the variant DNA. The gel mixture was polymerised by adding a x1/100 volume of 10% (w/v) APS and x1/1000 volume of TEMED, poured into the gel mould and allowed to set for at least 30 minutes.

Each PCR sample was mixed 3:1 with loading buffer, containing 15% (w/v) sucrose, 0.05% (w/v) xylene cyanol in dH_2O. The sample was denaturing for 3 minutes at 94°C and electrophoresis was in 1x TBE buffer, overnight at 15°C, with water-cooling. Exon 1 products were separated at 100V so that the xylene cyanol dye was at the bottom of the gel and exon 3 products at 200V. A DNA marker, 1kb ladder, was also run for comparison. The DNA was visualised by immersion of the gels in 0.5-1μg/ml ethidium bromide solution for 15 minutes and subsequently viewed under a UV light.

Exon 1 polymorphisms were detected on a 6% (w/v) polyacrylamide gel containing 5% (v/v) glycerol. The polymorphism close to exon 3 was detected on a 12% (w/v) polyacrylamide gel without glycerol.

2.2.7 Sequencing by the dideoxy chain termination method

Sequencing was performed using protocols based on the chain termination method of Sanger *et al.*, 1977. All sequencing protocols used a modified form of DNA polymerase, with the normal 3' to 5' exonuclease activity removed.

Genomic DNA was amplified by the PCR (2.2.4). All of the products were sequenced using Dynabead strand separation and the Sequenase II kit (USB), except when exons 1 and 4 were amplified using the primers in table 6, which were analysed by cycle sequencing.

2.2.7.1 The primers used for sequencing

Exons that were PCR-amplified as a single product (as in table 6) were sequenced using the same non-biotinylated primer as that used for the PCR. The four products containing exons 1, 2, 3-4 and 5-7, amplified as in table 7, were sequenced using the primers indicated in table 8, obtained from Dr CM Eng, Mount Sinai Medical Center, New York.

Table 8: Internal, intronic primers for sequencing the α-galactosidase A gene

Exon	+ or - strand	Sequencing primer (5'->3')	Gene nucleotide nos. (Kornreich *et al.*, 1989)
1	-	CTGGCTCTTCCTGGCAGTCAA	1360-1340
3	+	TGGTTCTCTCTTTCTGCTACC	7198-7218
4	+	TATAGCCCCAGCTGGAAATT	8261-8280
5	-	GCATCCTGCTCTAAGTACTCT	10418-10398
6	-	AGATTTAGGCCCAAGAC	10776-10760
6	+	GGGTCATCTAGGTAACTTTAAG	10442-10663
7	+	ATGAATGCCAAACTAAC	10910-10916

2.2.7.2 Sequenase II sequencing

2.2.7.2.1 Strand separation using Dynabeads

DNA was amplified (2.2.4) using 5-20pmol of each primer, one of which was biotinylated at the 5' end. The biotinylated product was bound to metallic beads coated with streptavidin, allowing them and the attached DNA to be concentrated and washed by attraction to a magnet, as described below (Hultman *et al.*, 1989). This process was carried out in a 0.5ml Eppendorf tube and at each step in which a solution was removed or the beads washed, the tube was placed in a Dynal MPC-E magnet for 30 seconds and the supernatant discarded. When a solution was added, the beads were resuspended fully. TES contained 100mM NaCl, 10mM Tris-HCl at pH 8 and 1mM EDTA at pH 8.

For each PCR product, 30µl of Dynabeads™ M-280 streptavidin beads were withdrawn, the liquid storage solution was removed and the beads were washed twice in 100µl of TES, to remove traces of sodium azide. PCR product, 50µl out of a 100µl reaction, was added to the washed Dynabeads and left for 15-30 minutes to allow the formation of biotin-streptavidin bonds. The beads were then washed once with TES before adding 100µl of freshly prepared 0.1-0.15M NaOH solution and incubating for 10-20 minutes to denature the DNA. The non-biotinylated strand was released into the supernatant and discarded. The pellet was washed with 100µl of NaOH solution to ensure removal of the non-biotinylated strand. Finally,

the single-stranded DNA, biotinylated DNA, which was attached to the beads, was washed once with TES and then once with TE buffer before resuspension of the pellet in 7µl of dH_2O for sequencing.

2.2.7.2.2 The sequencing reaction

The protocol described in the Sequenase II kit (USB) was used to sequence single-stranded DNA (2.2.7.2.1) and consisted of three steps: primer annealing; extension by DNA polymerase, in the absence of dideoxynucleotides; and termination of sequence by incorporation of dideoxynucleotides.

The following solutions were prepared. (a) 5x reaction buffer contained 200mM Tris-HCl at pH 8.0, 100mM $MgCl_2$ and 250mM NaCl; (b) enzyme dilution buffer contained 10mM Tris-HCl at pH 7.5, 5mM DTT and 0.5mg/ml BSA. It was used to dilute the Sequenase II enzyme 8 fold, immediately prior to use; (c) labelling mixture contained 7.5µM dGTP; 7.5µM dATP, 7.5µM dTTP; 7.5µM dCTP and was diluted 5-20 fold; (d) four dideoxytermination mixtures, each of which contained 50mM NaCl, 80µM dGTP, 80µM dATP, 80µM dTTP, 80µM dCTP and 8µM of one ddNTP, either ddGTP, ddATP, ddCTP or ddCTP; (e) stop solution contained 95% (v/v) deionised formamide, 20mM EDTA at pH 8, 0.05% (w/v) bromophenol blue and 0.05% (w/v) xylene cyanol.

Single stranded template, 5pmol of PCR product in a volume of 7µl, was mixed with 2µl of 5x reaction buffer and 1µl of sequencing primer which was complementary to the 3' end of the template. The mixture was heated at 65°C for 2 minutes and cooled to room temperature for 20-30 minutes to anneal the primer.

Extension of primers by DNA polymerase was conducted in the presence of radiolabelled [α-^{35}S] dATP, to allow detection by autoradiography. The annealed primer-template was mixed with 1µl of 0.1M DTT, 2µl of diluted labelling mixture and 0.5µl of [α-^{35}S] dATP (1000Ci/mmol. Polymerisation was initiated by adding 2µl of diluted Sequenase II enzyme and the mixture was incubated for 2-3 minutes at room temperature.

Chain-termination occurred in the presence of dideoxynucleotides. For one sequencing reaction, four tubes, each containing 2.5µl of one of the four dideoxytermination mixtures, were preheated to 37-42°C for 5 minutes. Four 3.5µl

aliquots of the polymerase extension mixture were withdrawn and added to each of the dideoxytermination mixtures. The chain-termination reactions were incubated at 37-42°C for 2-5 minutes and then stopped with 4µl of stop solution. All reactions were stored at -20°C.

For sequencing close to the primer, less than 50 bases, 20 fold dilutions of labelling mixture and short incubation times were used for the extension and termination steps. The chain-termination reaction was usually conducted at 40°C to help decrease formation of secondary structure in the template and failure of the polymerase to extend the sequence in this region of DNA.

2.2.7.3 Exo-*Pfu* cycle sequencing

The protocol described in the Exo-*Pfu* cycle-sequencing kit was used to sequence double-stranded DNA. Exo-*Pfu* DNA polymerase is thermostable, allowing sequencing using a cycling reaction analagous to the polymerase chain reaction. 200fmol of DNA template (~60ng of a 500bp product) was used in each reaction. The 10x sequencing buffer contained 200mM Tris-HCl at pH 8.8, 100mM KCl, 10mM $MgSO_4$, 100µM $(NH_4)_2SO_4$, 1% (v/v) Triton X-100, 1mg/ml BSA, 20µM of dATP and 50µM each of dCTP, dGTP and dTTP. The stop solution contained 80% (v/v) formamide, 50mM Tris-HCl at pH 8.3, 1mM EDTA at pH 8, 0.1% (w/v) bromophenol blue and 0.1% (w/v) xylene cyanol

PCR products were amplified (2.2.4) with 40pmol of each primer (table 6), in a 50µl reaction volume. The resulting product was diluted 5 fold and a sample of 1-5µl was withdrawn. 1pmol of primer was added to the sample and the volume was made up to 13µl with ddH_2O before placing on ice. The following solutions were added to form the reaction mixture: 4µl of 10x sequencing buffer, 1µl (10µCi) of [α-^{35}S] dATP (1000Ci/mmol), 2.5 units of Exo-*Pfu* polymerase and lastly, to avoid its precipitation, 4µl of DMSO.

The reaction mixture for each sample was divided into four aliquots of 5µl and each was added to 5µl of one of the four dideoxynucleotide solutions, which contained 1.5mM of: ddGTP, ddATP, ddTTP or ddCTP. Mineral oil, 15µl, was added to prevent sample evaporation. The same temperature cycles were used to amplify the original PCR product were used for sequencing, as described in table 6. The reactions were stopped with 5µl of stop solution.

2.2.7.4 Electrophoresis in a denaturing polyacrylamide gel

A polyacrylamide gel was prepared and electrophoresis was conducted as described previously (2.2.6.1.2) but using a different gel mixture. The stock polyacrylamide solution contained a different proportion of crosslinker molecule, N, N'-methylene-bis-acrylamide, with a ratio of 2% (w/v) methylene bisacrylamide to 38% (w/v) acrylamide in dH_2O (AccuGel 40™). The gel mixture was prepared by mixing 15ml of AccuGel 40™, 42g of urea and 20ml of 5x TBE buffer, increasing the volume to 100ml with dH_2O before filtering through 3MM Whatman paper (Sambrook *et al.*, 1989). This mixture could be stored at 4°C in the dark for several weeks.

The sequencing reactions (from sections 2.2.7.2.2 and 2.2.7.3) were denatured at above 85°C for 2-3 minutes and 2.5-4µl of each sample was loaded into 0.5cm-wide wells. Electrophoresis was at 55-65W (approx. 1kV) in 1x TBE buffer. Resolution of the first 150 bases was achieved by running the bromophenol blue marker dye to the bottom (40cm) of the gel (approx. 2 hours). A further 100-200 bases could be separated by running the xylene cyanol marker to the gel base (approx. 4 hours).

The sequencing ladders were detected by autoradiography (2.2.6.1.3).

2.2.8 Detection of known mutations by restriction enzyme digestion

Restriction enzymes cleave double stranded DNA at specific palindromic sequences. All digestions of DNA were conducted, following the manufacturers instructions. The reactions were incubated at 37°C, except for *Bst*N I and *Bst*X I which had optimal activity at 60°C and 55°C respectively.

Genomic DNA containing the sequence variant was amplified by the PCR in a volume of 50µl (2.2.4, table 6). 10-15µl of the product was diluted with dH_2O to give a final volume of 20µl, including 5-10 units enzyme, so that there was at least 1 enzyme unit per µg of DNA, and its appropriate buffer. The DNA was digested for 1-2 hours. Care was taken not to add enzyme to a volume of more than 10% of the total because the glycerol buffer, in which it is stored, can inhibit the action of some enzymes.

The digested DNA was analysed by electrophoresis in 1x TBE buffer in an agarose gel of suitable concentration (2.2.5). For gels of greater than 2% (w/v), the gel was prepared by mixing the ultrapure grade agarose to 2% (w/v) and increasing the concentration with low melting point SeaPlaque agarose.

2.3 EXPRESSION OF ALPHA-GALACTOSIDASE A MUTANTS

2.3.1 Biochemical analyses

2.3.1.1 Preparation of cell lysates

Lysis buffer contained 150mM NaCl, 50mm NaH_2PO_4 at pH 6.9, 1mM EDTA at pH 8, 1% (v/v) Nonidet-P40 and 0.02% (w/v) sodium azide and was stored at 4°C. A x1/1000 volume of 100mM PMSF in dH_2O (stored at -20°C) was added to lysis buffer and kept on ice, for use within 30 minutes (Sambrook *et al.*, 1989).

Cell pellets were resuspended in lysis buffer, containing PMSF, a proteinase inhibitor and centrifuged at 10000 x *g* for 15 minutes at 4°C. The supernatants were used for enzyme and protein assays.

2.3.1.2 Fluorescamine protein assay

Protein assays were carried out using the fluorescamine assay (Bohlen *et al.*, 1973). Cell lysate, containing approximately 50µg of protein, was mixed with 1ml of 0.2M sodium borate in dH_2O (pH 9.0) in a 75mm x 10mm borosilicate tube. While vortexing, 350µl of 0.3mg/ml fluorescamine, in anhydrous acetonitrile, was added. All samples were prepared in duplicate, and duplicate blank tubes, replacing cell lysate for lysis buffer were included for each set of assays.

The fluorescence was measured using a Turner 111 fluorimeter (Farand), with a primary excitation filter (Corning 7-60) at 365nm and secondary emission filters (Wratten) at 465nm. The fluorimeter was calibrated with a Fluram filter set.

2.3.1.3 Enzyme assay for human α-galactosidase A

Alpha-galactosidase A assays were carried out using a modified version of the methods by Kint, 1970 and Mayes *et al.*, 1980. Citrate-phosphate buffer was prepared by mixing 0.1M citric acid in dH_2O and 0.2M Na_2HPO_4 in dH_2O, to form a solution at pH 4.6. The substrate solution for measurement of both α-galactosidase A and B contained 5mM 4-methylumbelliferyl-α-D-galactopyranoside

(4MU-α-gal), dissolved in citrate-phosphate buffer, pH 4.6, by sonication for 3 minutes. Alpha-galactosidase A activity was measured by adding an inhibitor of α-galactosidase B, N-acetylgalactosamine (GalNAc), to a concentration of 0.1M in the substrate solution. This has been shown to inhibit the B form by 94% Mayes *et al.*, 1980. Alpha-galactosidase A activity was defined as the activity remaining after inhibition of the B form with GalNAc.

Cell lysate was made up to a volume of 20µl with cell lysis buffer. A blank, with no lysate, was prepared and all samples were in duplicate, in 75mm x 10mm borosilicate tubes. Then 150µl of substrate solution, with or without GalNAc inhibitor, was added. The samples were incubated at 37°C for an appropriate length of time before stopping the reaction with 2.375ml of 0.1mM (6.68ml/l) of ethylene diamine. The fluorescence was measured as before, (2.3.1.2) but the fluorimeter was calibrated with the 4MU filter set.

2.3.1.4 Enzyme assay for *E. coli* β-galactosidase

The β-galactosidase assay was carried out by a modified version of the method described by the manufacturers, Clontech, for measurement of activity produced by expression from the plasmid vector pCMVβ. Z buffer, at pH 7.0, contained 60mM Na_2HPO_4, 40mM NaH_2PO_4, 10mM KCl, 1mM $MgSO_4$ and 0.02% (w/v) sodium azide. The substrate solution contained 3mM 4-methylumbelliferyl-β-D-galactoside, which was dissolved in Z-buffer by heating at 80°C for about 1 minute, immediately prior to use.

Cell lysate, was increased to 140µl with Z buffer. All samples and a blank, with no cell lysate (Z buffer only), were prepared in duplicate in 75mm x 10mm borosilicate tubes. Substrate solution, 35µl, was added and the samples incubated at 37°C for an appropriate length of time. The reaction was stopped with 2.375ml of 0.1mM (6.68ml/l) ethylene diamine in dH_2O and the fluorescence was measured as described (2.3.1.2) but with calibration of the fluorimeter using the 4MU filter set.

2.3.2 Determination of α-galactosidase A activity in cultured cells

Cells from one flask (75cm²) of lymphoblastoid cells or 2 confluent flasks (75cm²) of fibroblasts were harvested and washed in 0.9% (w/v) NaCl (2.2.1) before preparation of cell lysates in 1ml of lysis buffer and biochemical analyses (2.3.2). For protein assays on lymphoblastoid cells and fibroblasts, 2µl and 10µl of the cell lysate was used, respectively. For all α-galactosidase assays, the substrate included 0.1M GalNAc. For assays on both fibroblasts and lymphoblastoid cells, 10µl of cell lysate containing approximately 100-200µg and 30-70µg of total cell protein, respectively, was used and the incubation time was 20 minutes.

2.3.3 Site-directed mutagenesis

2.3.3.1 The megaprimer method of mutagenesis

The megaprimer method of site-directed mutagenesis (Sarkar *et al.*, 1990; Barik, 1993) relies on the production of mutant sequences in a two-step PCR amplification process: amplification from the mutant primer to one end of the cDNA to form the megaprimer, followed by extension from the megaprimer to the opposite end of the cDNA, to produce the full-length mutant cDNA.

In the first step, megaprimers were produced by PCR amplification (2.2.4), using the primers and conditions indicated in table 9. For each amplification, the primer pair consisted of either an antisense (-) mutant and T7 (+) primer, or a sense (+) mutant and SP6 primer. In each reaction, 50ng of template DNA, plasmid pGB6 (figure 7 (A)), containing the full length α-galactosidase A cDNA, was used. *Vent* polymerase was used with a buffer that contained 100mM KCl, 10mM Tris-HCl (pH7.4) and 50% glycerol.

Each double-stranded megaprimer product was purified, with removal of the unincorporated primers, using a Centricon column, following the manufacturers recommendations (Amicon). Megaprimers formed using the R112H, V269A, Q279E, R301Q, V316E, Q327K and G361R primers were purified on Centricon 100 columns which remove DNA with a molecular weight of less than 100000, approximately 150bp of double-stranded DNA. The PCR product was added to the column with 2ml of ddH_2O and centrifuged at 4500 x *g* for 10 minutes. The Centricon column was washed twice more with ddH_2O and centrifuged.

Primer	Sequence (5'->3') [(+) and (-) are sense & antisense sequences, respectively]	cDNA nucleotide nos.(ref)	$MgCl_2$ (mM)	*Anneal Temp (°C)
G35R	CCTTGCCAATC**T**ATTGTCCAG (-)	114-94	2	50
R49L	GCACATGAAG**A**GCTCCCAGTG (-)	156-136	2	50
R112H	TGAGGAAAG**T**GCTGAGGGT (-)	344-326	2	46
V269A	GATATGTTAG**C**GATTGGCAAC (+)	796-816	4	46
Q279E	CAGCTGGAAT**G**AGCAAGTAAC (+)	825-845	4	46
R301Q	ATGACCTCC**A**ACACATCAG (+)	893-911	4	46
V316E	GATAAGGACG**A**AATTGCCATC (+)	937-957	2	46
Q327K	TTGGGCAAG**A**AAGGGTACCA (+)	970-989	2	46
G361R	GGAGATTGGT**A**GACCTCGCTC (+)	1071-1091	4	46
T7	TAATACGACTCACTATAGGG (+)	Vector	-------	-------
SP6	GATTTAGGTGACACTATAG (-)	Vector	-------	-------

Table 9: Primers for producing mutant megaprimers

Primers were designed by inspection, with the mutation (underlined and in bold) near the centre, flanked by at least 9 bases and with an estimated melting temperature of 50-60°C. Megaprimers were produced by 30 cycles of PCR amplification (2.2.4) at: 94°C for 30 seconds, annealing for 30 seconds and 72°C for 1 minute, in a 100µl reaction volume, using 2 units of *Vent* DNA polymerase (NEB). Each cDNA mutagenesis primer was used with one pGB6 vector primer, T7 or SP6.

Figure 7: Plasmid vectors pGEM-3Z and pGB6 used for site-directed mutagenesis

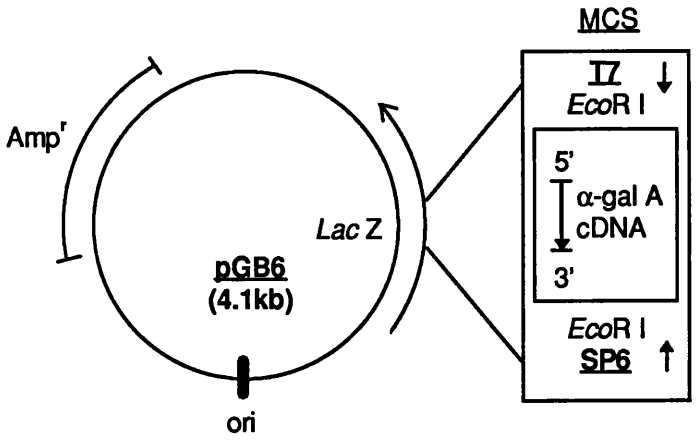


Figure 7 (A): pGB6, containing the α-galactosidase A cDNA

This vector was a gift from Dr RJ Desnick and Dr Y Iounnou, Mount Sinai Medical Center, New York. pGB6 is derived from pGEM-3Z (Figure 7(B)). It contains the α-galactosidase A cDNA cloned into the *Eco*R I site of the multicloning site, MCS, interrupting the α-peptide LacZ gene sequence. Amp^r encodes β-lactamase, conferring ampicillin resistance to the host bacterium and ori is the origin of replication. The primers located at SP6 and T7 promoter sites were used for PCR amplification in the direction indicated by the arrows.

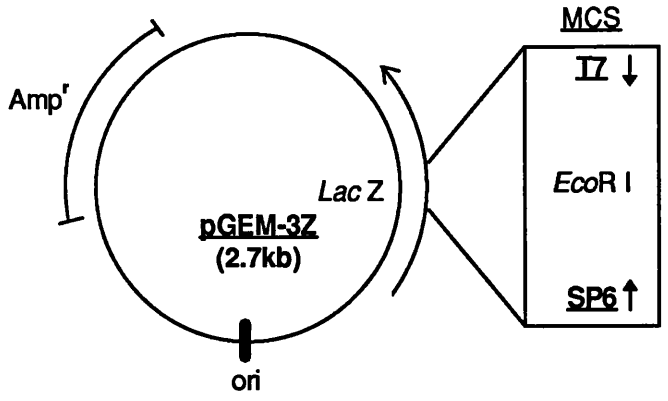


Figure 7 (B): pGEM-3Z subcloning vector (Promega)

LacZ is the gene encoding the LacZ α-peptide, allowing blue/ white colour selection of recombinants. Amp^r encodes β-lactamase, conferring ampicillin resistance to the host bacterium and ori is the origin of replication. T7 and SP6 are transcription promoters and the multicloning site (MCS) contains a unique *Eco*R I site for subcloning.

It was then inverted over a collection tube and centrifuged to recover the sample in about 40µl of liquid. For the smaller megaprimers, formed with G35R and R49L, a similar purification method was used, except that Centricon 30 columns were used to remove DNA with a molecular weight of less than 30000, and centrifugation steps were for 30-40 minutes.

Full length cDNA was produced in a second PCR amplification using half of the purified megaprimers and a vector primer, SP6 or T7, at the opposite end from that used to produce the megaprimers. The 30 amplification cycles were at 94°C for 30 seconds, 52°C for 30 seconds and 72°C for 1.5 minutes. The reaction mixture contained 1µM primers, 4mM $MgCl_2$ and 1µl of *Vent* DNA polymerase. The PCR-amplified mutant cDNA was digested (2.2.8) with *Eco*R I, for subcloning.

2.3.3.2 Purification of mutant cDNA by electrophoresis

The whole PCR product that contained the mutant cDNA was subjected to electrophoresis through a 0.7% (w/v) agarose gel (2.2.5) and the portion of gel containing the cDNA, approximately 1.4kb in size, was excised while viewing under long wavelength UV light. The gel slice was placed in a 1.5ml Eppendorf tube and liquified by crushing it using a 1ml syringe plunger. The volume was increased to at least 350µl with dH_2O and then 0.5 volumes of 7.5M ammonium acetate and one volume of phenol (equilibrated with 0.5M Tris-HCl to pH 8) were added. The gel mixture was vortexed vigorously, frozen at -70°C for 5 minutes, thawed at 37°C for 5 minutes and then centrifuged at 10000 x *g* for 5 minutes. The top aqueous phase was transferred to a new tube and an equal volume of phenol: chloroform (1:1) was added to it. The mixture was again vortexed, centrifuged for 5 minutes and the top aqueous layer removed. If the solution did not appear clear at this stage, the phenol:chloroform step was repeated. Otherwise, 1 volume of chloroform was added and the tube vortexed before centrifugation at 10000 x *g* for 5 minutes, to remove traces of phenol. The upper phase was transferred to a separate tube and DNA was precipitated from it by adding 2.5 volumes of ethanol. The tube was placed at -70°C for 15 minutes and centrifuged at 10000 x *g* for 15 minutes. The pellet was washed with 70 % (w/v) ethanol and resuspended in dH_2O or TE. The DNA concentration was estimated by electrophoresis on a 0.8% (w/v) agarose gel (2.2.5) and comparison of the band intensity with that of a known amount of DNA marker.

2.3.4 Subcloning of mutated cDNA

2.3.4.1 Bacterial cell culture

2.3.4.1.1 Growth of E. coli on agar plates

Luria-Bertani (LB) growth medium was prepared, containing 10g/l tryptone, 10g/l NaCl and 5g/l yeast extract in dH_2O. Liquid agar was prepared by autoclaving 15g of agar in 1 litre of LB medium, which was allowed to cool to room temperature. The liquid agar was poured into sterile petri dishes with a diameter of 15cm, to a depth of about 0.5cm and the plates allowed to set for about 30 minutes. The plates were stored at 4°C.

An appropriate volume of *E. coli* in LB medium was used to allow the growth of discrete bacterial clones that were not confluent. A maximum volume of 200µl was spread over the surface of an agar plate, using a sterile glass spreader. The plates were inverted and incubated at 37°C for 12-16 hours for cell growth (Sambrook *et al.*, 1989).

2.3.4.1.2 Growth of bacteria in liquid medium and storage

A single clone of *E. coli* was picked from an LB-agar plate (2.3.4.1.1), using a sterile plastic loop. It was transferred into 5-50ml of LB medium and the cells were grown to saturation at 37°C for 12-16 hours, in a shaking incubator (model G26, New Brunswick Scientific Co. Inc.). For growth of larger volumes of cultured bacteria, this saturated culture was diluted by 1 in 100 and regrown for 12-16 hours.

The bacteria were stored as a glycerol stock by mixing 900µl of cultured cells with 600µl of 50% (v/v) glycerol, to give a final concentration of 20% (v/v) glycerol. They were stored at -20°C or -70°C, indefinitely.

2.3.4.1.3 Preparation of competent cells for transformation

Competent cells were prepared using a method based on that by Cohen *et al.*, 1972 and Oishi *et al.*, 1972. A single colony of *E. coli* cells, strain DB10B™, was grown overnight in 20ml of LB medium (2.3.4.1.2). 0.5ml of the culture was diluted in 100ml of LB and grown for a further 2-3 hours to an optical density of 0.4-0.6 at 600nm. The culture was chilled on ice and spun at 7000 x *g* for 5 minutes at 4°C to separate the cells.

The cells were resuspended in 20ml of ice-cold 0.1M $MgCl_2$, left on ice for 20 minutes and centrifuged at 7000 x *g* and 4°C for 5 minutes. The pellet was then resuspended in 2ml of 0.1M $CaCl_2$ and left on ice for 1 hour to establish competency. Competent cells were stored by adding x1/4 volume of sterile 75% (v/v) glycerol and freezing at -70°C in 0.1ml aliquots.

2.3.4.2 Subcloning of mutant cDNA into pGEM-3Z

Preparation of competent cells and bacterial cell culture was as described (2.3.4.1.3).

Digestion of 10µg of pGEM-3Z vector (figure 7(B)) with *Eco*R I was performed in a 20µl volume (2.2.8). The vector was then dephosphorylated at its 5' ends (Sambrook *et al.*, 1989). At least 1 unit of calf intestinal phosphatase per 100pmol of 5' phosphate DNA ends was added and the reaction volume was increased to 100µl by adding 25µl of 4mM $ZnCl_2$ and dH_2O. The mixture was incubated for 30 minutes at 37°C. This treatment prevented self-ligation, catalysed by DNA ligase. The plasmid was purified (2.3.3.2).

Purified mutant cDNA (2.3.3.2) and 15-50ng of dephosphorylated pGEM-3Z vector, both of which had been treated with *Eco*R I, were ligated in a molar ratio of 1:1-3, using *T4* DNA ligase. Ligation was carried out in a 10-20µl volume, with 400U of *T4* DNA ligase and an appropriate buffer (provided by the manufacturer), at 16°C for 1-16 hours.

Half of each ligation mixture was mixed with 100-200µl of DB10B™ competent cells. Transformation was carried out by incubation on ice for 30 minutes, heating at 42°C for 90 seconds and back on ice for a further 2 minutes. The cells were

grown in 2ml of LB medium by incubation at 37°C, with shaking for 1-1.5 hours. This allowed the transformed bacteria to gain antibiotic resistance.

The transformed cells were spread over an agar plate with 20µl of 0.1M IPTG in dH_2O, 100µl of X-Gal (20mg/ml in dimethylformamide) and 100µl of ampicillin (50mg/ml in dH_2O), to allow colour and antibiotic selection. The cells were grown overnight. White colonies, which contained plasmid and the ligated cDNA insert, were scraped from the agar into 10ml of LB medium, grown overnight and stored as glycerol stocks. The remaining culture was used for sequence analysis.

2.3.4.3 Sequence analysis of mutant cDNA in pGEM-3Z

2.3.4.3.1 Isolation of plasmid DNA from bacteria

A method based on the boiling lysis protocol described by Wang, L-M *et al.*, 1988, was used for isolating plasmid containing the mutated cDNAs, for restriction enzyme digestion and sequencing. Approximately 15-20µg of pGEM-3Z vector could be isolated from 3ml of bacterial cell culture, with plasmid amplified in *E. coli*, strain DB10B™.

STET solution was prepared to contain 8% (w/v) glucose, 0.5% (v/v) Triton-X 100, 50mM EDTA at pH 8 and 10mM Tris-HCl at pH 8. A solution of lysozyme (1mg/ml) in STET was made up fresh or, alternatively, was stored as a stock solution (5mg/ml) at -20°C, which was thawed once only before use. All centrifugation steps were at 10000 x *g*.

Three ml of each cell culture, containing the mutant cDNA in pGEM-3Z (2.3.4.2), was divided into two 1.5ml aliquots in Eppendorf tubes. The cells were centrifuged for 30 seconds, the supernatants discarded and the Eppendorf tubes inverted over a paper towel to remove all liquid. The two pellets were recombined by resuspending them in 350µl of STET. Then 40µl of lysozyme solution (1mg/ml in STET) was added to digest bacterial cell walls. The mixture was vortexed before boiling for 0.5-1 minute. It was centrifuged for 10-15 minutes and the protein/ bacterial DNA pellet removed with a sterile toothpick and discarded.

An optional step, for digestion of RNA, was carried out by adding 1µl of RNase solution (10mg/ml in dH_2O) and leaving at room temperature for 5 minutes. The

plasmid was then precipitated by adding 0.5 volumes of 7.5M ammonium acetate and one volume of isopropanol. The mixture was centrifuged for 30 minutes before discarding the supernatant. The pellet was washed with ethanol (70%, v/v), air-dried and redissolved in 50µl of TE or dH_2O. The DNA (1-2µl) was analysed by electrophoresis through a 0.8% agarose (w/v) gel (2.2.5), in 1x TAE buffer, which allows better resolution of supercoiled vector than TBE buffer.

2.3.4.3.2 Sequencing of the mutant cDNA constructs

The mutant cDNAs in pGEM-3Z plasmid were made single-stranded, using the method described by the manufacturers of the Sequenase II kit (USB). Denaturing solution was prepared fresh, by mixing 5M NaOH, 100mM EDTA at pH 8 and dH_2O in a ratio of 20:1:79, respectively. Approximately 5-10µg of isolated plasmid (2.3.4.3.1) was mixed with 5µl of denaturing solution and made up to 20µl with dH_2O. The mixture was incubated for 30 minutes at 37°C and then neutralised with 2µl of 3M sodium acetate (pH 4.5-5.5). The single-stranded DNA was precipitated by the addition of 2.5 volumes of ethanol and incubation at -70°C for 15 minutes, followed by centrifugation at 10000 x *g* for 15 minutes. The DNA pellet was washed with 70% (v/v) ethanol and dried before either storage at -20°C or resuspension in dH_2O for immediate use.

The presence of each mutation was verified by sequencing, using Sequenase II (2.2.7.2.2) and the primers indicated in table 10. A region of 161bp-263bp, between the two flanking, unique restriction enzyme recognition sites indicated in table 10, was also fully sequenced to ensure that no other mutations were present.

Minimum cDNA nucleotides sequenced	Region contains mutations	Sequencing primers (5'->3') (+) and (-) are on the cDNA sense & antisense strands, respectiveiy.	Primer No.	Primer cDNA nucleotide nos.(ref)
-34 to 127 *Sac*II-*Nco*I	G35R	GCACATGAAGAGCTCCCAGTG (-) TGAGGAAAGTGCTGAGGGT (-)	1 2	156-136 344-326
122 to 362 *Nco*I-*Nhe*I	R49L R112H	TGAGGAAAGTGCTGAGGGT (-) TTTAGCAGAACTACTCCCCAG (-)	3 4	344-326 483-503
775-1037 *Bst*XI- *Eco*NI	V269A Q279E R301Q V316E Q327K	CTCCTGTGAGTGGCCTC (+) ATGACCTCCAACACATCAG (+)	5 6	600-616 893-911
1027-1237 *Eco*NI- *Alw*NI	G361R	GATAAGGACGAAATTGCCATC (+) GGAGATTGGTAGACCTCGCTC (+)	7 8	937- 957 1071-1091

Table 10: Primers for sequencing mutant cDNA constructs

Plasmids were sequenced 2.2.7.2, using the primers indicated, to cover the region between the two flanking unique restriction sites indicated and containing the mutation. cDNA nucleotide numbers were as in Bishop *et al.*, 1991.

2.3.4.4 Subcloning a portion of the mutant cDNA into pGB6

A small, sequenced portion of each mutant cDNA that had been subcloned into pGEM-3Z was subcloned into the wild-type cDNA of plasmid pGB6 (figure 7 (A)), as follows. Restriction enzyme digestions and DNA purification were carried out as described (2.2.8 and 2.3.3.2, respectively), except where specified otherwise.

The mutant cDNAs in pGEM-3Z were digested with the restriction enzymes indicated in table 10 and the released, mutant cassette was purified. Ten μg of pGB6, plasmid containing the wild-type α-galactosidase A cDNA, was cut with the same pair of enzymes used to create each mutant cassette. Digestions with the *Sac* II and *Nco* I, *Nco* I and *Nhe* I, and *BstX* I and *Eco*N I, pairs of enzymes were as described previously and the vector was then purified. Each of these enzymes cuts at a unique site in the α-galactosidase A cDNA and did not cut the remaining vector.

The method used to digest pGB6 with *Eco*N I and *Alw*N I was modified slightly because although *Eco*N I cuts the cDNA only, *Alw*N I cuts both the cDNA and the vector once. To obtain pGB6 that was cut with *Eco*N I and *Alw*N I in the cDNA only, pGB6 was digested completely with *Eco*N I and then divided into aliquots. Aliquots containing 1μg of plasmid were digested with 0.3-5 units of *Alw*N I, for 10 minutes at 37°C, to produce partially digested vector. The digestion mixture was immediately electrophoresed in a 0.7% (w/v) agarose gel to purify the correct vector fragment.

The remaining subcloning steps were as described in section 2.3.4.2. All of the cut, pGB6 vectors were dephosphorylated and ligated to the mutant cassette that was cut with the same restriction enzymes. Competent DB10B™ cells were transformed and the cells were grown on agar the presence of ampicillin, but in the absence of IPTG and X-Gal. Glycerol stocks were prepared. The presence of the mutation was verified by sequencing (2.3.4.3).

2.3.5 Transfection and expression of cDNA mutants

2.3.5.1 Subcloning of mutant α-galactosidase A cDNAs into expression vector pMT2

Subcloning, transformation and preparation of glycerol stocks was performed as described for cloning into pGEM-3Z (2.3.4.2). The full length, mutant cDNAs and the wild-type cDNA, in pGB6 were digested with *Eco*R I, purified and subcloned into the expression vector pMT2 (figure 8) which was also digested with *Eco*R I. The transformed cells were grown on agar with ampicillin but without IPTG and X-Gal. Vector was isolated (2.3.4.3.1) and the orientation of the insert in pMT2 was determined by restriction analysis using *Bgl* II, which cuts pMT2 once and the cDNA once. cDNAs that were successfully subcloned into pMT2 in the correct orientation, were sequenced (2.3.4.3.2) to verify the presence of the mutation.

2.3.5.2 Midi-preparations of plasmid DNA using the Qiagen kit

Extraction of plasmid DNA on a large scale was carried out by alkaline lysis, using the method of Birnboim *et al.*, 1979 and purification of plasmid DNA by adsorption and washing on a Qiagen-tip 100 column, following the manufacturers instructions. This allowed isolation of about 100μg of plasmid that was suitable for transfection of COS-1 cells.

The following buffers were used. Buffer P1 contained 100mg/ml RNase A, 50mM Tris-HCl at pH 8 and 10mM EDTA at pH 8. P2 contained 200mM NaOH and 1% (w/v) SDS and P3 contained 2.55M potassium acetate at pH 4.8. QBT buffer, pH 7, contained 750mM NaCl, 50mM MOPS, 15% (v/v) ethanol and 0.15% (v/v) Triton-X 100. QC buffer, pH 7, contained 1M NaCl, 50mM MOPS and 15% (v/v) ethanol and QF buffer, pH 8.2, contained 1.5M NaCl, 50mM MOPS and 15% (v/v) ethanol.

Bacterial cultures were grown in 50-100ml of LB medium overnight (2.3.4.1.2). The cultured cells were pelleted by centrifugation at 5000 x *g* for 5 minutes at 4°C. The supernatant was discarded and the tube inverted over a paper towel to remove excess liquid. The cell pellet was resuspended in 5ml of P1 then 5ml of P2 was added and mixed gently by inversion before incubation at room temperature for 5 minutes. Five ml of ice-cold P3 buffer, was added to neutralise the solution. The mixture was placed on ice for 15-20 minutes to allow renaturation and solubilisation of the plasmid and precipitation of bacterial DNA, protein and SDS.

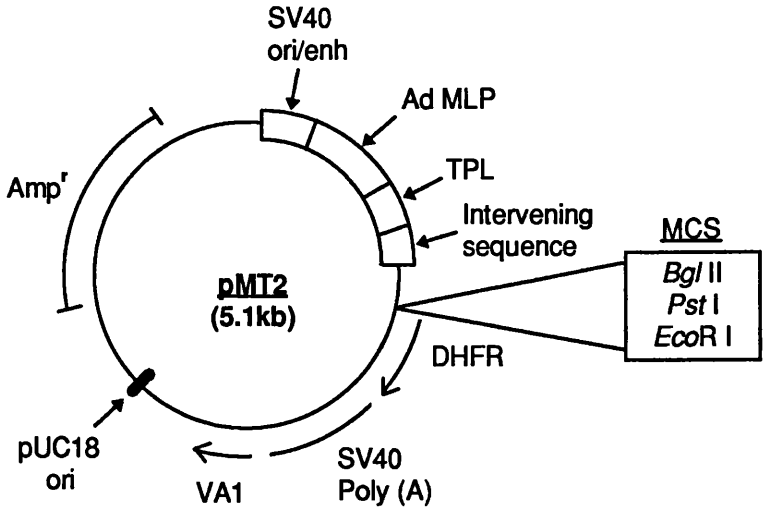


Figure 8: pMT2 mammalian expression vector (Kaufman *et al.*, 1989, 1990)

pMT2 is an expression vector which contains the pUC 18 (Stratagene) origin, ori, of replication and the Amp^r gene, allowing amplification of the vector and selection by ampicillin resistance in *E. coli*. Other elements are important for high level expression in mammalian cells. The **SV40 ori** binds SV40 large T antigen, produced by the host COS-1 cells, which are African green monkey kidney CV-1 cells transformed with origin defective mutant SV40 virus, for replication of pMT2. The **SV40 enh**, enhances transcription from the adenovirus major late promoter, **Ad MLP**. The tripartite leader sequence, **TPL**, enhances the translation efficiency and forms part of an mRNA that encodes **DHFR**, dihydrofolate reductase, allowing selection during stable transfection. The gene of interest is inserted into the multicloning site, **MCS**, and is transcribed to form a bicistronic mRNA with DHFR, terminating transcription at the **SV40 poly (A)** site. Adenovirus **VA1** produces RNA, which potentiates translation.

The mixture was centrifuged at 30000 x *g* for 30 minutes at 4°C and the supernatant was immediately transferred into a fresh tube. The pellet was discarded.

A Qiagen 100 column was equilibrated with 4ml of QBT buffer. The supernatant from the alkaline lysis, containing the plasmid DNA, was poured into the equilibrated column. Once the solution had passed through, allowing the DNA to adsorb to the silica-based column surface, the column was washed twice with 10ml of QC buffer. Purified DNA was then eluted with 5ml of QF buffer. Plasmid DNA was precipitated from the eluate with 0.7 volumes of isopropanol, followed by centrifugation at 15000 x *g* and 4°C for 30 minutes. The pellet was washed with 70% (v/v) ethanol and the supernatant discarded. The pellet was air-dried for 20-30 minutes before resuspension in 500µl of sterile ddH_2O. The plasmid concentration was measured spectrophotometrically (2.2.3.3).

2.3.5.3 Transfection of COS-1 cells by electroporation

The COS-1 cell line is derived from monkey kidney cells transformed with SV40, Gluzman, 1981. They grow as an adherent monolayer and were cultured and harvested in the same manner as fibroblasts (2.2.1.2), except that the growth medium was DMEM (Dulbecco *et al.*, 1959), containing 10% (v/v) fetal calf serum, 4mM glutamine, 100 units/ml penicillin and 100µg/ml streptomycin. An almost confluent flask (75cm²) of COS-1 cells was split into 5 new flasks (75cm²) and grown overnight. Colcemid™ (colchicine, a 10µg/ml solution) was then added to the medium to give a concentration of 0.1µg/ml colchicine and the cells were grown for 12-24 hours, causing them to arrest in metaphase. The cells were trypsinised and centrifuged at 500 x *g* for 5 minutes at room temperature. Cell pellets from 1-2 flasks were resuspended in 0.7ml of DMEM medium in 4mm electroporation cuvettes (Invitrogen). 10µg of pMT2 plasmid containing α-galactosidase A cDNA (figure 8) and 0-3µg of pCMVβ plasmid, for coexpression of *E. coli* β-galactosidase (Clontech, figure 9), were added and mixed. These plasmids had been prepared using the Qiagen method (2.3.5.2). The mixture of cells and plasmids was electroporated (Neurmann *et al.*, 1982; Chu *et al.*, 1987), using a BioRad Gene Pulser, at 250V with a capacitance of 500µF and set to the extend mode. After electroporation, the cells were transferred into 10ml of DMEM medium in 15cm diameter petri dishes and incubated at 37°C in 5% CO_2 for 24-72 hours.

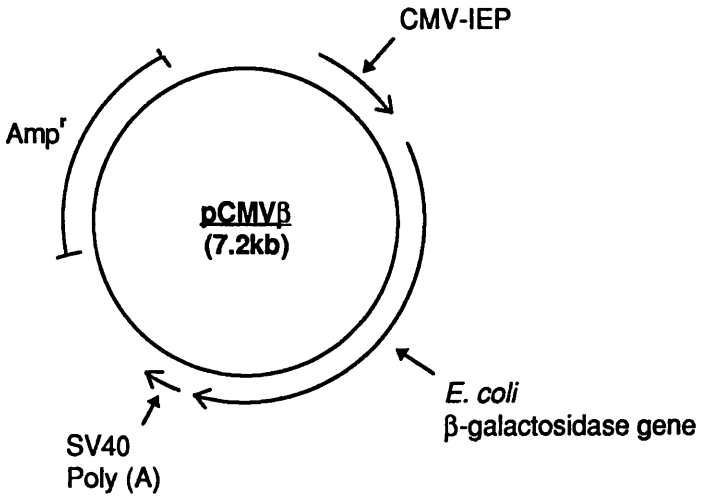


Figure 9: pCMVβ mammalian expression vector (Clontech)

pCMVβ plasmid (Clontech) expresses the *E. coli* β-galactosidase gene. The human cytomegalovirus immediate early promoter/ enhancer (CMV-IEP) controls this and transcription is terminated at the **SV40 poly (A)** site. The Amp^r gene allowes amplification of the vector and selection by ampicillin resistance in *E. coli.*

2.3.6 Analysis of expression products

2.3.6.1 Measurement of the expressed enzyme activities

All biochemical assays were performed as described, section 2.3.1. Medium from one dish of electroporated COS-1 cells was removed and 10-20µl of this was assayed for α-galactosidase activity, using substrate with, or without GalNAc. The remaining adherent cell monolayer was washed three times with 0.9% (w/v) NaCl, removing all liquid. The cells were lysed by addition of 1ml of lysis buffer with PMSF.

The cell lysate was assayed for protein, α-galactosidase A activity and *E. coli* β-galactosidase activity, as appropriate. For the fluorescamine assay, 10µl of cell lysate was used. For measurement of α-galactosidase, 10-20µl of cell lysate was incubated with substrate, with or without GalNAc, for 10-30 minutes. The *E. coli* β-galactosidase assays were carried out with 2-20µl of cell lysate, incubated for 10-20 minutes.

2.3.6.2 Detection of α-galactosidase A protein

2.3.6.2.1 Radioactive labelling in COS-1 cells

COS-1 cells were transfected with α-galactosidase A constructs only, without cotransfection of pCMVβ using electroporation (2.3.5.3) . They were incubated for 24-48 hours in 15cm petri dishes before labelling (Sambrook *et al.*, 1989). The medium was replaced with 5ml of methionine-free medium, containing 10% (v/v) dialysed fetal calf serum. The cells were incubated at 37°C for 30 minutes to allow depletion of the essential amino acid methionine.

The medium was discarded and replaced with fresh methionine-free medium. Cell labelling grade methionine-L-^{35}S, 4µl, (133µCi) was added and the cells labelled for 6 hours at 37°C.

The medium was transferred to Falcon tubes and cell debris was removed by centrifuging at 500 x *g* for 5 minutes. Cells were lysed with 1ml of lysis buffer (2.3.1.1).

2.3.6.2.2 Immunoprecipitation

Half of the medium and of the cell extract obtained from COS-1 cell labelling was used for immunoprecipitation (Sambrook *et al.*, 1989). IgG-purified and titrated anti-(α-galactosidase A) serum was used for immunoprecipitation. NET buffer contained 50mM Tris-HCl at pH 8, 150mM NaCl, 0.1% (v/v) Nonidet-P40, 0.25% (w/v) gelatin and 0.02% (w/v) sodium azide. The 5x SDS loading buffer contained 200mM Tris-HCl (pH 8), 10% (w/v) SDS, 50% (v/v) glycerol and 0.5% (w/v) bromophenol blue.

A volume of 3µl of anti-(α-galactosidase A) serum was added to the cell lysate and medium, to bind to the enzyme, and this was incubated at 4°C, with rocking, for 1 hour. Protein A-Sepharose, 20µl, which binds IgG and therefore α-galactosidase A antibody, was added and this was incubated at 4°C for 1 hour, with rocking.

The Sepharose-protein A-antibody-enzyme complex was precipitated by centrifugation. It was then washed sequentially by adding 1ml each of: NET with 0.5M NaCl; then NET with 0.1% (w/v) SDS; and finally NET which contained 10mM Tris-HCl, pH 7.5 and 0.1% (v/v) Nonidet-P40. At each step the pellet was resuspended completely and then pelleted by centrifuging at 10000 x *g* for 20 seconds.

The immunoprecipitates were mixed with 1x SDS loading buffer, which contained a 0.1x volume of freshly added β-mercaptoethanol. These were boiled for 5 minutes to denature the protein, centrifuged at 10000 x *g* for 5 minutes and the supernatant kept for analysis. Samples were stored at 4°C for 1-2 weeks.

2.3.6.2.3 SDS-polyacrylamide gel electrophoresis

SDS-polyacrylamide gel electrophoresis was conducted by the method of Laemmli, 1970, using the Protean II vertical minigel system and a 10cm x 10cm x 1.5mm thick gel size. A 30:0.8% (w/v) acrylamide to bisacrylamide stock polyacrylamide solution was used to prepare two gel mixtures. The stacking gel contained 5% (w/v) acrylamide, 0.25M Tris-HCl at pH 6.8 and 0.1%(w/v) SDS. The separating gel contained 10% (w/v) acrylamide, 0.38M Tris-HCl at pH 8.8, and 0.1% (w/v) SDS.

Each gel was polymerised with a x1/1000 volume TEMED and a x1/100 volume of APS.

First the separating gel was poured into the mini-gel mould (Protean II minigel system), filling to 0.5cm below the level of the comb base and overlaid with butanol to ensure that the surface set flat. Once set, the butanol was removed and the stacking gel was poured in. The comb was inserted and the gel allowed to set for about 30 minutes.

Electrophoresis was conducted at 4°C and using a buffer that contained 25mM Tris base, 250mM Glycine and 0.1% (w/v) SDS. The samples were loaded alongside a ^{35}S-labelled rainbow marker (Amersham) and electrophoresed through the stacking and separating gels at 100V and 150V respectively. Once the bromophenol blue dye had just reached the gel base, the protein gel was removed and fixed in 20% (v/v) methanol and 10% (v/v) acetic acid for 5-10 minutes. It was then placed in Amplify™ solution (Amersham) for 15-30 minutes before drying. The protein was detected by autoradiography as described, as in section 2.2.6.1.3, except that the gels were dried by slowly heating to 80°C, maintaining the 80°C temperature for 90 minutes and then slowly cooling, to prevent the gel from cracking. Also, intensifier screens were included during exposure of the film.

3. RESULTS

3.1 SEQUENCE CHANGES IN THE ALPHA-GALACTOSIDASE A GENE

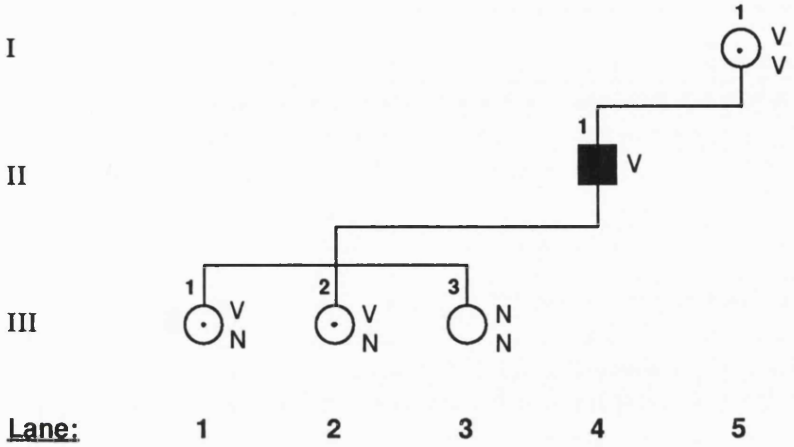
Genetic analysis of the α-galactosidase A gene was conducted for two reasons: firstly, to determine the mutations that cause Fabry disease, allowing the relationship between genotype and the clinical phenotype to be studied; and secondly, to evaluate this as a method of carrier detection. At the start of this study it was hoped that phenotype information would be obtained in all families analysed for mutations. However, no information was available for 6 families, numbers 7, 8, 13, 22, 25 and 30.

Two experimental strategies were used to identify unknown mutations. One method was to sequence the entire protein-coding region of the gene (seven exons) and the splice-site consensus sequences. Alternatively, each exon and its flanking splice-site sequences were initially screened by SSCP analysis and only those which showed an altered band pattern were sequenced.

3.1.1 Polymorphisms in the α-galactosidase A gene

3.1.1.1 Detection of a variant SSCP analysis pattern in exon 1

Radioactive SSCP analysis in a hemizygote from Fabry family 7 was carried out to locate regions with sequence changes. The first analysis, of exon 1, revealed a variant band pattern in exon 1. The segregation of this band alteration was investigated by SSCP analysis of this family (figure 10) and showed two unexpected findings. Firstly, the mother (I-1) of hemizygote II-1 was homozygous for the pattern alteration (lane 5) but was still living at age 72 and so she was not thought to be homozygous for a disease-causing mutation. This indicated that the altered band pattern represented a polymorphic sequence change. Secondly, although two daughters, III-1 and III-2, lanes 1-2, had inherited the variant pattern from their affected father and presumably the normal pattern from the mother (not analysed), the third daughter III-3, lane 3, did not carry the altered bands. This can be explained by one of three possibilities. The sequence variant in hemizygote II-1 may have been altered during inheritance to give the normal pattern, or the daughter is a case of non-paternity, or she has inherited two X chromosomes from her mother which give the normal pattern.



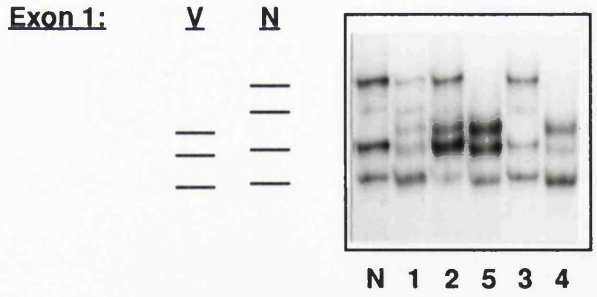


Figure 10: Radioactive SSCP analysis of exon 1 in family 7

Exon 1 was amplified by PCR and subjected to radioactive SSCP analysis. Lanes 1-5 are as indicated in the pedigree and lane N is the normal pattern. Segregation of the normal pattern, N, and the variant, V, is marked on the pedigree diagram.

3.1.1.2 Radio-labelled SSCP analysis of all exons in the normal α-galactosidase A gene

The possibility that polymorphisms may be present in the α-galactosidase A gene prompted a study of DNA from the normal population. They were of interest for 2 reasons. Firstly, they may useful for carrier detection and secondly, they needed to be distinguished from putative disease-causing mutations.

Radio-labelled SSCP analysis of at least 100 chromosomes from a sample of 61 male and 21 female, non-Fabry, British individuals identified 6 polymorphic, variant band patterns. These were named according to the exon with which they were associated and the order in which they were identified. Thus, the allele designation 1P3 indicates that the sequence variation was either within or close to exon 1 and that it was the third polymorphic variant found in this DNA fragment. Three polymorphic variants, named 1P1, 1P2 and 1P3 were found to be associated with exon 1 and one, 3P1, was with exon 3 (figure 11). In exon 1 there was no alteration in the homoduplex DNA but in exon 3 the homoduplex DNA of 3P1 migrated faster and may include a sequence deletion. Clear heteroduplex bands were also observed in a female with the common pattern and 3P1. Two other polymorphic variants, 5P1 and 7P1, were associated with exons 5 and 7 respectively (see figures 20 and 21, section 3.1.2.1.1).

Sixty males and 19 females were fully analysed for all 7 exons. The frequency of the 6 variant band patterns is indicated in table 11(A) and the haplotypes are shown in table 11(B). One or more polymorphism(s) was found in 10% of chromosomes in males. The phase of the polymorphisms in the females was not established and so the number of polymorphic chromosomes was unknown. DNA samples for one male and two females were incompletely analysed because of lack of sufficient material. The frequency of these polymorphic patterns indicated that they may be useful for carrier detection in some Fabry families.

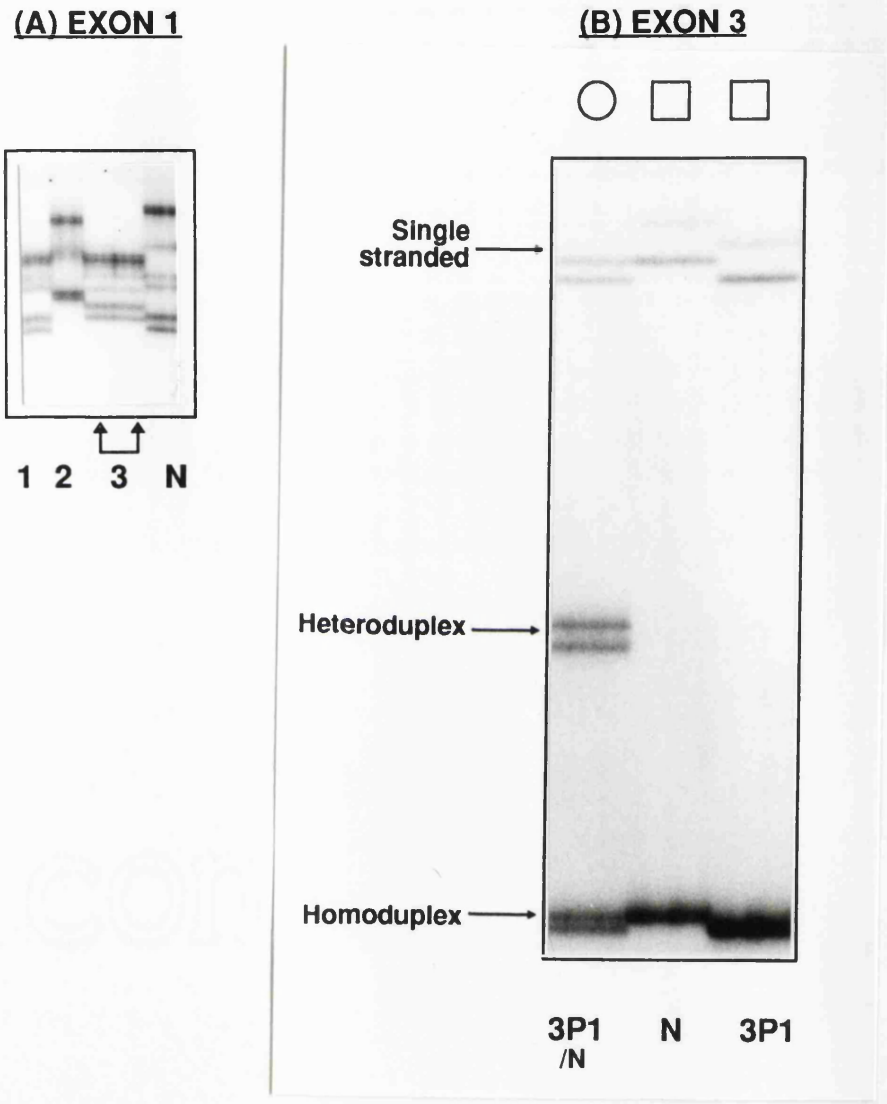


Figure 11: Radioactive SSCP analysis of four polymorphisms

SSCP analysis by electrophoresis on 6% polyacrylamide gels containing 5% glycerol, at room temperature and at 340V for ~16hr. PCR products including exons 1 and 3 (2.2.6.1) were analysed. In **(A)**, DNA from males is analysed. 1, 2 and 3 are bands from single-stranded DNA containing polymorphisms 1P1, 1P2 and 1P3 respectively. N is the most common pattern. In **(B)**, N is the most common pattern, 3P1 is the deletion polymorphism.

(A) The frequency of polymorphic variants

Exon analysed	SSCP Variant	Frequency/ No. chromosomes		Total frequency
		Males	Females	
1	1P1	1/61	2/40	3%
1	1P2	2/61	3/40	5%
1	1P3	1/61	1/40	2%
3	3P1	5/61	6/40	11%
5	5P1	4/60	5/40	9%
7	7P1	5/60	8/40	13%

(B) The observed haplotypes

Males/ females	Frequency	Cosegregation of the normal (N) and polymorphic variant patterns in exons:			
		1	3	5	7
Males	2	**1P2**	**3P1**	**5P1**	**7P1**
	2	N	**3P1**	**5P1**	**7P1**
	1	**1P1**	N	N	**7P1**
	1	**1P3**	N	N	N
	1	N	**3P1**	-	-
	54	N	N	N	N
Females	1	**1P2** / N	**3P1** / N	**5P1** / N	**7P1** / N
	3	N / N	**3P1** / N	**5P1** / N	**7P1** / N
	1	**1P2** / N	N / N	N / N	**7P1** / N
	1	**1P1** / N	N / N	N / N	**7P1** / N
	1	N / N	N / N	**5P1** / N	**7P1** / N
	1	**1P1** / **1P3**	N / N	N / N	**7P1** / N
	1	N / N	**3P1** / N	N / N	N / N
	1	**1P2** / N	**3P1** / N	-	-
	1	-	-	N	N
	10	N / N	N / N	N / N	N / N

Table 11: The frequency of polymorphic variants and haplotypes in 61 males and 21 females

Radioactive SSCP analysis was used to identify polymorphic variants in non-Fabry, British individuals. 'N' is the normal pattern and in **(B)**, '-' indicates that the DNA was not analysed.

3.1.1.3 Sequence analysis of polymorphic changes in the α-galactosidase A gene

For each distinct band pattern that was produced by SSCP analysis, DNA from a single normal male was sequenced at least twice (2.2.7) to identify the specific alterations. Each PCR product was fully sequenced and contained a single change. One was a 5 base pair deletion and five others were single base substitutions (table 12 and figures 12 13, 14 and 15). Polymorphisms 1P1 and 1P3 were predicted to change a *Msp* I and a *Sac* II restriction enzyme recognition site, respectively (figure 12). The sequencing results for 1P1 and 1P3 were confirmed by enzyme digestion, which offers an alternative strategy to SSCP analysis for detection of the polymorphisms and their potential use in identifying the pattern of inheritance.

For polymorphism 3P1, the deletion may have been produced by slipped mispairing (Krawczak *et al.*, 1991) and this mechanism could potentially generate several deletion sizes. To investigate whether the 3P1 SSCP analysis variant was caused by independent mutagenic events, giving different deletion sizes, DNA was sequenced in four non-Fabry males and in an affected hemizygote in Fabry family 16, who also carried this polymorphism. In all five individuals the deletion was identical and therefore gave no evidence that it was a recurrent mutation.

3.1.1.4 Detection of polymorphisms for tracing the pattern of inheritance

The usefulness of genetic methods to detect polymorphisms and their application as an indirect method of carrier detection in Fabry families was investigated. Three methods were used to detect the polymorphisms within the α-galactosidase A gene, allowing the inheritance pattern to be followed.

3.1.1.4.1 Radio-labelled SSCP analysis

In Fabry family 16 (figure 16), two males, II-1 and III-2, were diagnosed clinically as typical Fabry hemizygotes but the carrier status of the clinically normal female III-1 was in question. In female III-1, the α-galactosidase activity (E. Young, personal communication) in leukocytes and plasma was 69nmol/mg/hr and 9.9nmol/mg/hr, respectively, within the normal ranges of 40-162nmol/mg/hr and 4.8-26.5nmol/mg/ml. The possibility that a skewed pattern of X-chromosome

Location	Allele	Sequence change	Gene No. (cDNA No.)	Restriction site altered
Exon 1	1P1	G -> A	1068 (-12)	-*Msp* I
Exon 1	1P2	C -> T	1050 (-10)	
Exon 1	1P3	G -> T	1070 (-30)	-*Sac* II
Intron 2	3P1	7184-cccccagcc-7192 -> -cccc-	7187del5	
Intron 4	5P1	a -> g	10115	
Intron 6	7P1	c -> t	10956	

Table 12: Polymorphic sequence changes in the α-galactosidase A gene of non-Fabry individuals

Polymorphisms were identified in normal individuals by screening the PCR products for each exon using radioactive SSCP analysis (2.2.6.1) and then by sequencing those with altered band patterns (2.2.7). Gene numbering was as in Kornreich *et al.* (1989) and cDNA nucleotides were numbered as in Bishop *et al.* (1991).

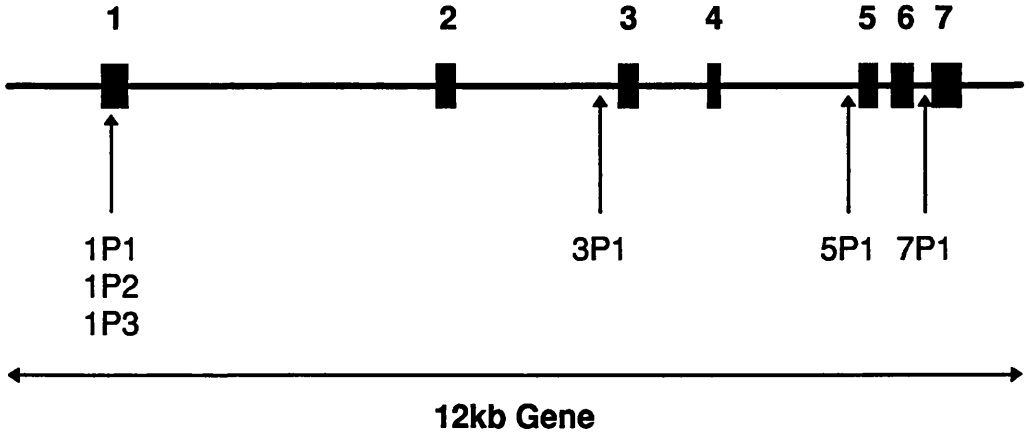
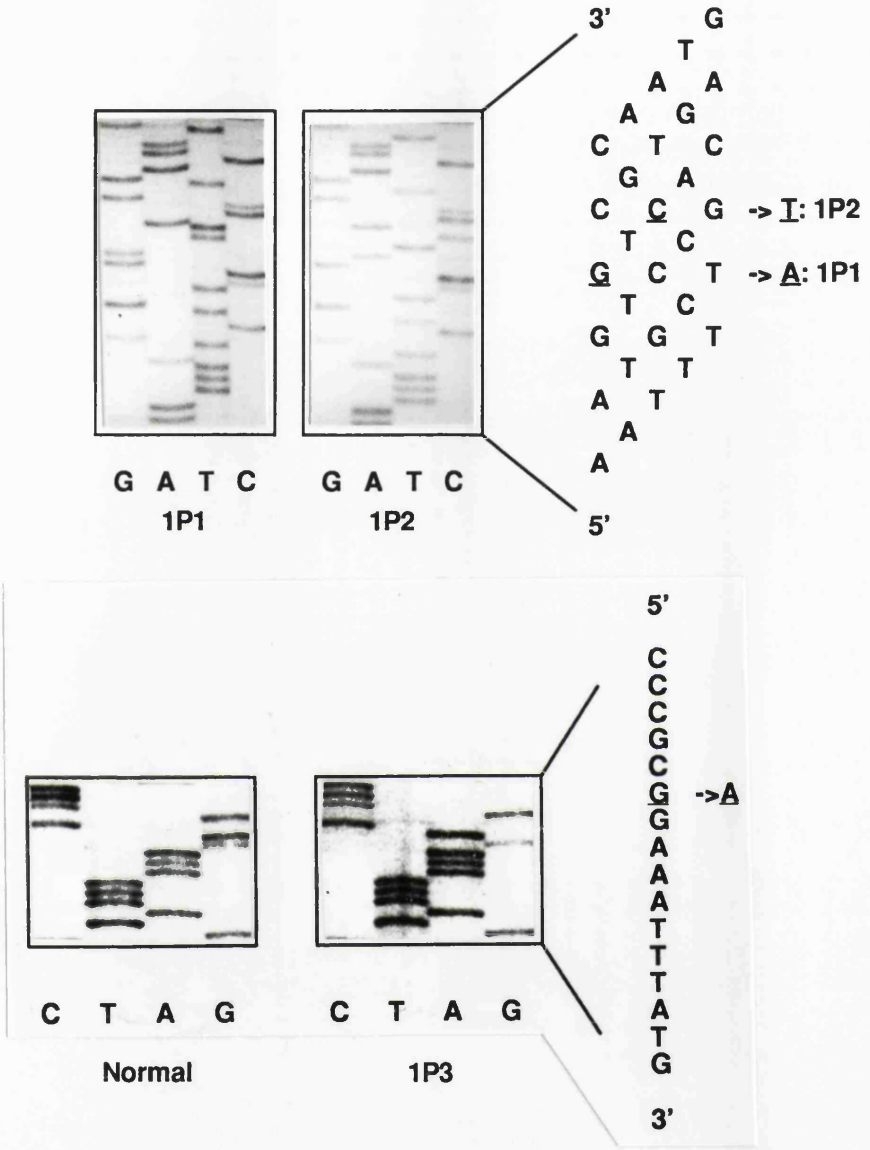


Figure 12: The Location of polymorphisms in the α-galactosidase A gene

Exons are indicated by the filled boxes and the numbers are indicated above. Polymorphisms 1P1, 1P2 and 1P3 are in the 5' untranslated region of exon 1. All others are in the introns.

Figure 13: Polymorphisms in the 5' untranslated region of exon 1



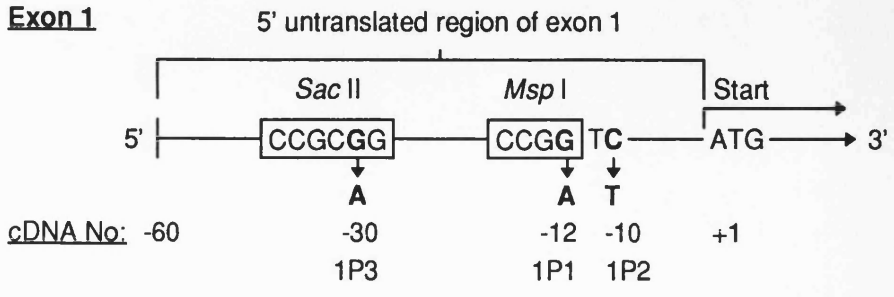


Figure 14: Polymorphism 3P1 in intron 2 of the α-galactosidase A gene

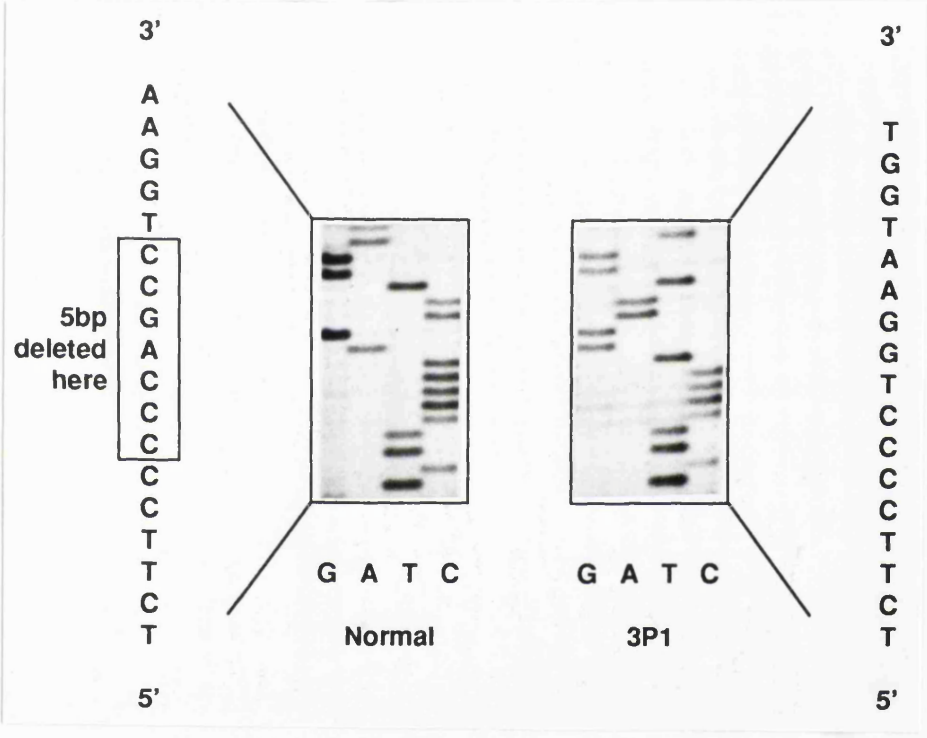
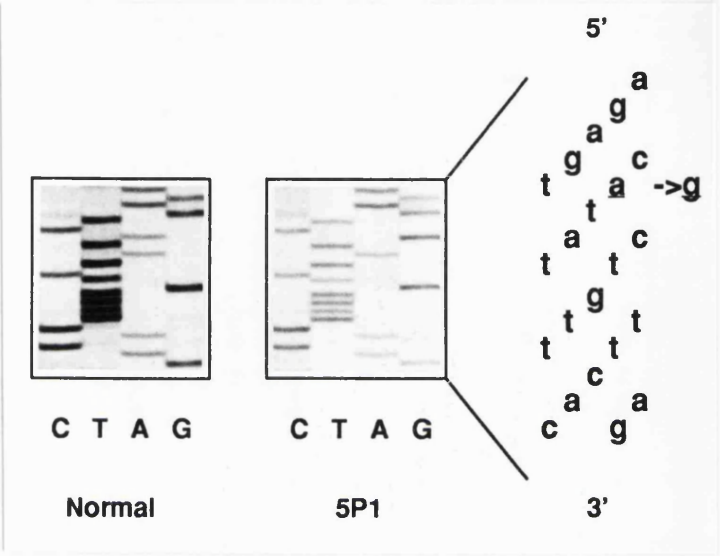
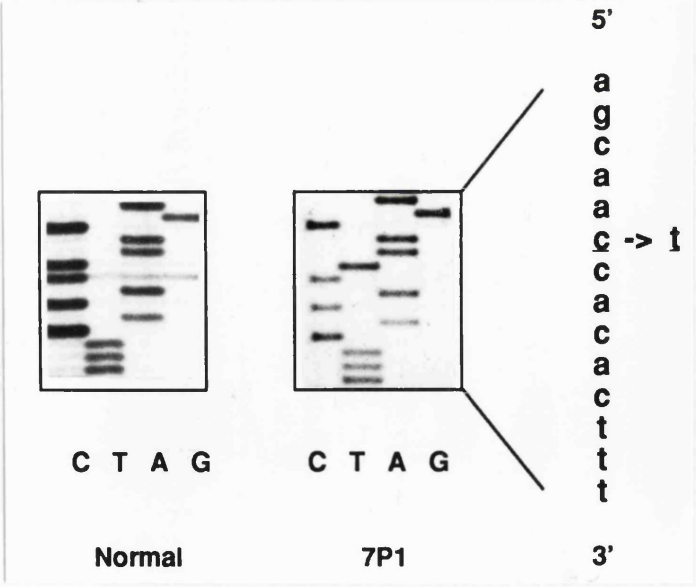


Figure 15: Polymorphism 5P1 and 7P1 in introns 4 and 6 of the α-galactosidase A gene





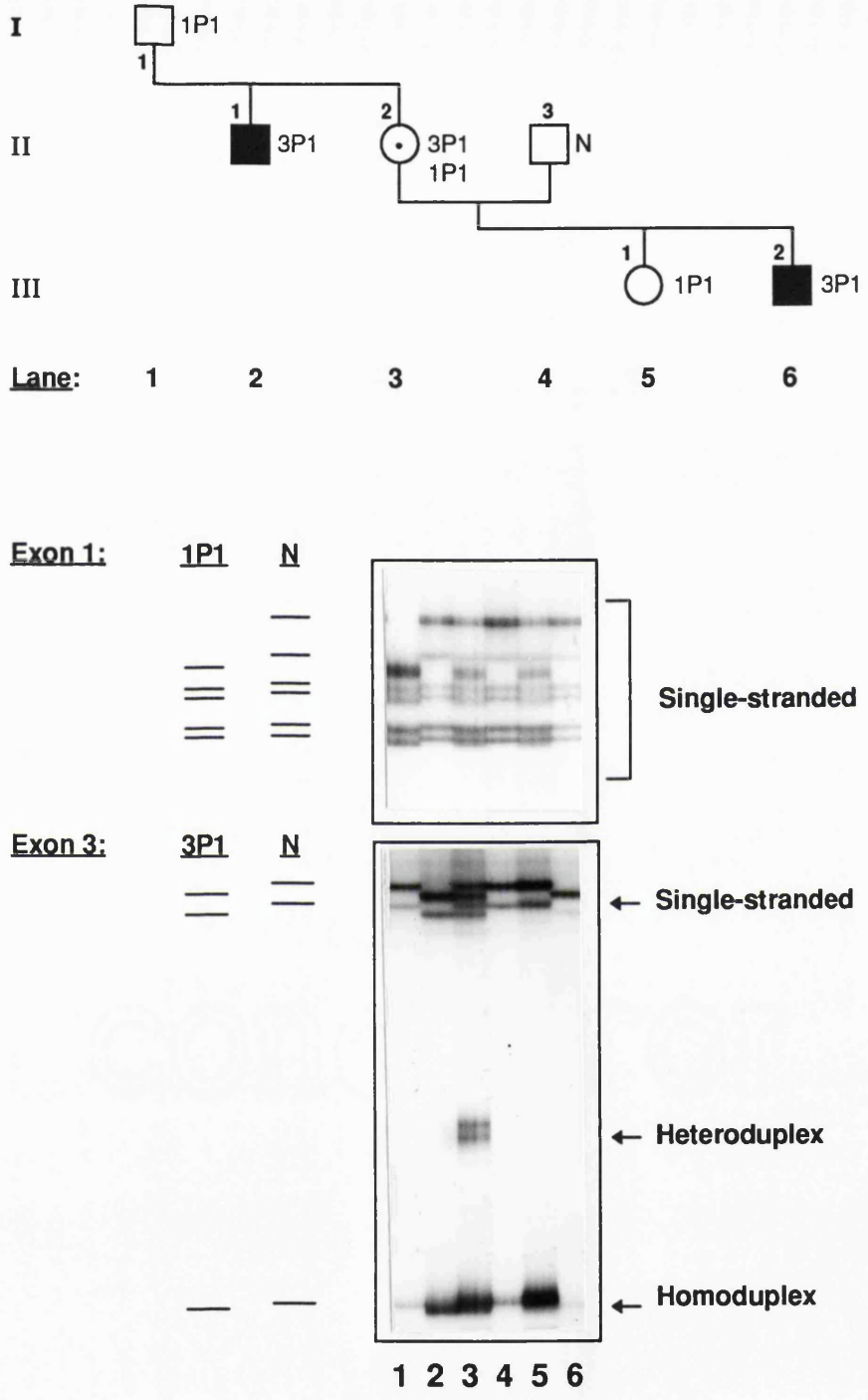


Figure 16: Radioactive SSCP analysis to show segregation of 1P1 and 3P1 in Fabry family 16

The pedigree of a Fabry family is shown (1-6) and the inherited polymorphisms are indicated. The band patterns of the common sequence and the polymorphisms are indicated on the left-hand side of the SSCP analysis gel, where N is the common pattern. Open squares and circles are unaffected males and females. Filled squares and circles with a dot are hemizygotes and heterozygotes for Fabry disease.

inactivation could produce an apparently normal activity in a carrier could not be determined biochemically.

Polymorphism analysis in exons 1 and 3 of members of family 16, by radio-labelled SSCP analysis, detected the presence of 1P1 and 3P1 (figure 16). The 3P1 band pattern cosegregated with the Fabry disease phenotype. The 1P1 pattern was associated with the normal chromosome and was inherited from male I-1, lane 1. In female III-1, lane 5, inheritance of 1P1 and not 3P1 from her obligate heterozygote mother, female II-2, lane 3, supported her diagnosis as a non-carrier. Therefore, indirect genetic analysis by this method has proved useful in this family. This result was confirmed by identification of the gene defect and demonstration of its absence, in female III-1 (3.1.2.2, figure 23(C)).

3.1.1.4.2 Restriction enzyme analysis in Fabry family 7 and a normal family

Although radio-labelled SSCP analysis can detect all of the polymorphisms, two of them, 1P1 and 1P3 alter a restriction site. Since analysis by restriction enzyme digestion does not require detection by radio-labelling and the preparation of gels for electrophoretic detection it is simpler than SSCP analysis. The use of this technique was tested by polymorphism analysis in two families, whose polymorphisms had previously been detected by SSCP analysis.

In Fabry family 7 the disease phenotype segregated with 1P1. Digestion of the 350bp PCR product for exon 1 with *Msp* I yields fragments of 238bp and 112bp in the absence of 1P1 but is uncut in its presence. This difference could be detected by electrophoresis of the DNA in a 3% agarose gel (figure 17(A)). These results confirmed the SSCP analysis, 3.1.1.1, in which the obligate Fabry heterozygote I-1, lane 5, was found to be homozygous for this polymorphism. Her son II-1, lane 4, had 1P1 and Fabry disease and the polymorphism had been inherited in 2 of 3 daughters. The third daughter failed to inherit the intragenic polymorphism. Confirmation that the polymorphism segregated with the disease phenotype was shown by direct analysis of the gene defect (3.1.2.3, table 16).

In figure 17(B), segregation of the 1P3 polymorphism is shown in a non-Fabry family. Exon 1 was amplified by PCR to give a 350bp product which was digested with *Sac* II to give a 254bp and 96bp fragment in non-polymorphic DNA and was

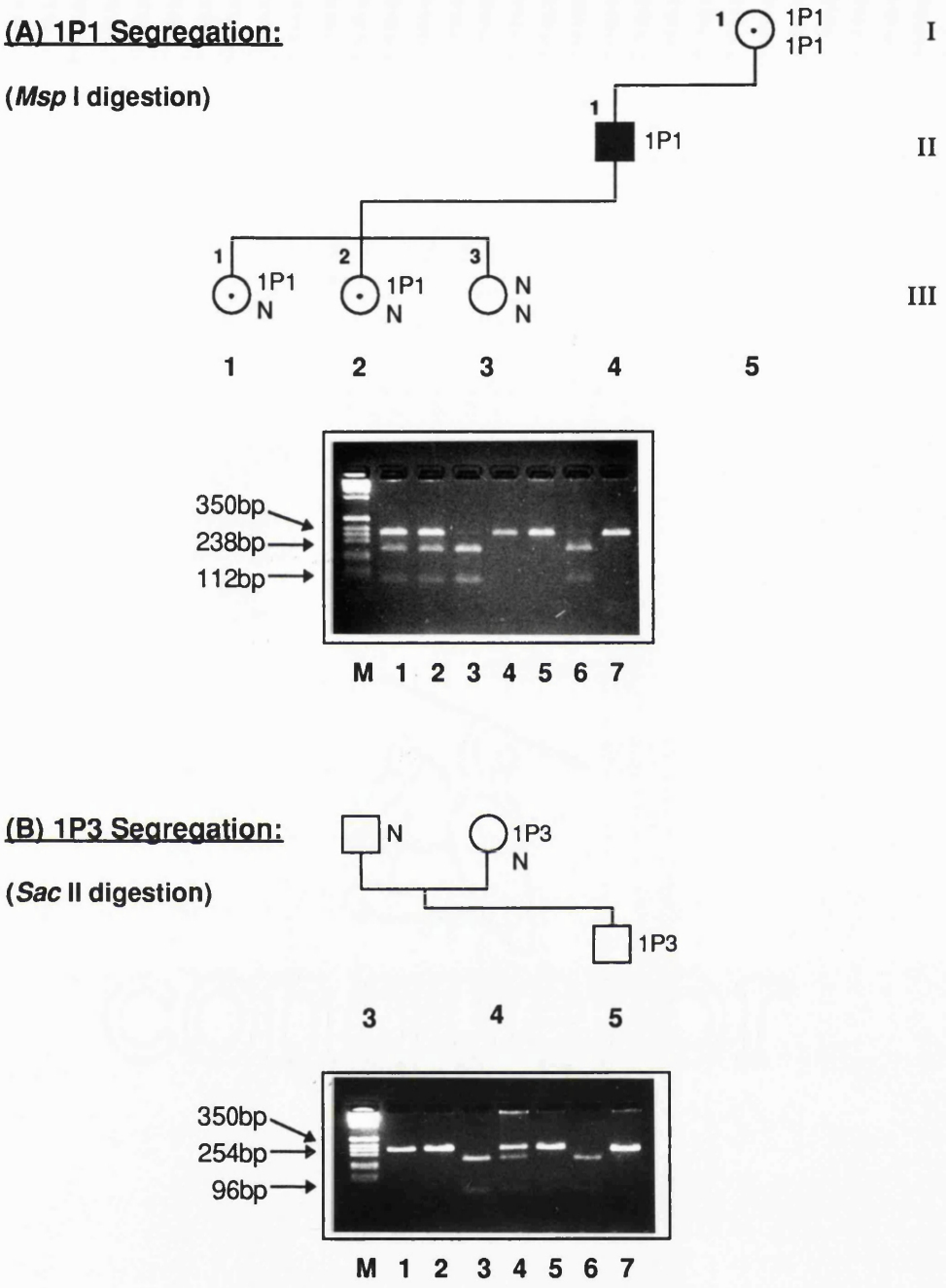


Figure 17: Restriction enzyme digestion analysis of 1P1 and 1P3

(A) Analysis of 1P1 segregation by *Msp* I digestion and separation on a 4% agarose gel. Lanes 1-5 are as in the pedigree. **(B)** Analysis of 1P3 segregation by *Sac* II digestion and separation on a 3% agarose gel. Lanes 3-5 are as in the pedigree, 1 is undigested product and 2 is an unrelated male with the 1P3 polymorphism.
In both cases, 6 is the common pattern, 7 is undigested product and M is the 1kb ladder DNA marker.

uncut when 1P3 was present. 1P3 was detected by electrophoretic analysis in a 4% agarose gel and Mendelian inheritance was observed.

The resolution of the restriction fragment length patterns due to the polymorphisms is good and the speed and simplicity of the method is an improvement on radio-labelled SSCP analysis. However, it is limited to the detection of 1P1 and 1P3.

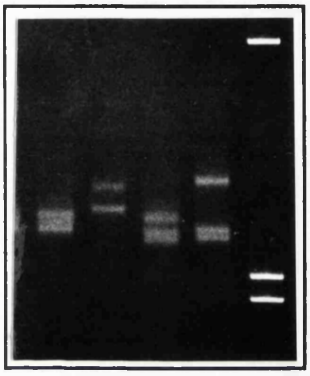
3.1.1.4.3 Detection of polymorphisms by non-radio-labelled SSCP analysis

A third method of detection, using a simple gel system (figure 18), was developed to allow simultaneous detection of the three polymorphic variants in exon 1 and detection of the 3P1 polymorphism. In this method, small (16cm x 20cm x 1mm) non-denaturing polyacrylamide gels were used for electrophoresis and the gel constituents and conditions were modified so that all of the polymorphisms could be distinguished. Electrophoresis was carried out at about 15°C and 100V for 16 hours. The PCR products that contained the polymorphisms in exon 1 were denatured and electrophoresed in a gel that contained 6% polyacrylamide and 5% glycerol (figure 18(A)), while 3P1 was detected in a 12% polyacrylamide gel with no glycerol (figure 18(B)). Analysis could be performed within 24 hours and the band shifts were visualised by staining with ethidium bromide (Yap *et al.*, 1993). This detection method was faster and the gels were easier to prepare than for radio-labelled SSCP analysis. In addition, it allowed more polymorphisms to be detected than by restriction analysis.

3.1.2 Detection of family-specific mutations in Fabry patients

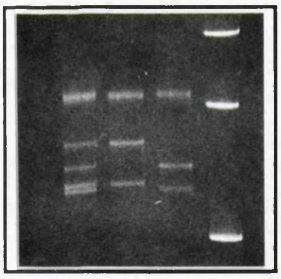
Individuals from 24 unrelated Fabry families were screened for unknown mutations by SSCP analysis prior to sequencing the altered portion of DNA. A further 6 British families were analysed in collaboration with Dr RJ Desnick and Dr CM Eng in New York but they were not screened by SSCP analysis. Mutations in these families were detected by sequencing the entire protein-coding region and splice sites. In addition, DNA that was screened by SSCP analysis and did not show an abnormal band pattern was also fully sequenced.

(A) EXON 1



3 2 1 N M

(B) EXON 3



2 1 N M

Figure 18: Non-radioactive SSCP analysis of 1P1, 1P2, 1P3 and 3P1

DNA was analysed by SSCP (2.2.6.2) and detected by ethidium bromide staining. M was the DNA marker (1kb ladder, GibcoBRL) and N was the most common pattern. In **(A)**, 1, 2 and 3 were males with polymorphisms 1P1, 1P2 and 1P3, respectively. In **(B)**, 1 is a male with 3P1 and 2 is a female with 3P1 and the common pattern.

3.1.2.1 Identification of mutations in 30 families

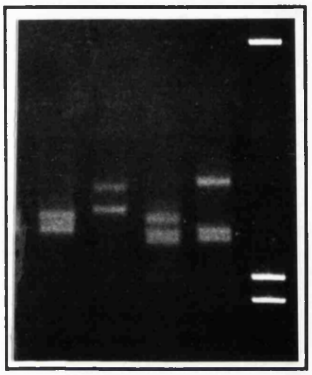
3.1.2.1.1 Mutation detection by radio-labelled SSCP analysis and sequencing

Twenty one hemizygotes, from unrelated Fabry families 1, 3, 4, 6-8, 10, 12, 16-24, 26, 28-30 and three heterozygotes from families 9, 13, 14 were screened for unknown mutations in the 7 exons and splice sites, using radio-labelled SSCP analysis (2.2.6.1). DNA that contained known polymorphisms was electrophoresed beside the patient DNA. This allowed identical, polymorphic band patterns to be distinguished from unique patterns, due to mutations.

In twenty of the 24 samples, one altered, non-polymorphic band pattern of either the single-stranded and/or the double-stranded DNA was observed in just one of the 7 exons. Thirteen of these are shown in figures 19, 20 and 21. All but one of the 20 mutant patterns showed an alteration in the migration of single-stranded DNA, the exception being in family 8 (figure 20(B), lane 1). The hemizygote in family 8 was detected by a shift in the mobility of the homoduplex DNA. No altered band patterns were observed for individuals in families 3, 6, 29 and 30. In all cases, analysis of PCR products containing exons 2-5 produced two major bands of single-stranded DNA for each X-chromosome, representing the two complementary sequences and a shift in this pattern of single-stranded DNA indicated the presence of a mutation. However, the PCR products from exons 1 and 6 always gave more than one band for each single strand, indicating that several stable conformations existed during the electrophoresis. The simplicity of the band patterns contributed to the relative ease with which mutation screening was performed, as compared with sequencing.

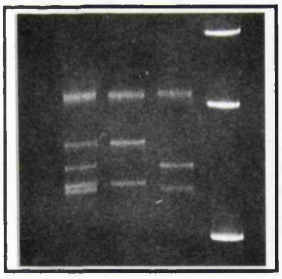
One or more polymorphic band patterns were found in 4 hemizygotes, in addition to the non-polymorphic band pattern changes. The band patterns matched the 1P3, 1P1 and 7P1, 3P1 and 5P1, and 1P3 polymorphisms in hemizygotes from families 4, 7, 16 and 28, respectively. The presence of the 1P3 and 1P1 polymorphisms were confirmed by *Sac* II and *Msp* I digestion respectively and 3P1 and 7P1 were confirmed by sequencing (2.2.7).

(A) EXON 1



3 2 1 N M

(B) EXON 3



2 1 N M

Figure 18: Non-radioactive SSCP analysis of 1P1, 1P2, 1P3 and 3P1

DNA was analysed by SSCP (2.2.6.2) and detected by ethidium bromide staining. M was the DNA marker (1kb ladder, GibcoBRL) and N was the most common pattern. In **(A)**, 1, 2 and 3 were males with polymorphisms 1P1, 1P2 and 1P3, respectively. In **(B)**, 1 is a male with 3P1 and 2 is a female with 3P1 and the common pattern.

3.1.2.1 Identification of mutations in 30 families

3.1.2.1.1 Mutation detection by radio-labelled SSCP analysis and sequencing

Twenty one hemizygotes, from unrelated Fabry families 1, 3, 4, 6-8, 10, 12, 16-24, 26, 28-30 and three heterozygotes from families 9, 13, 14 were screened for unknown mutations in the 7 exons and splice sites, using radio-labelled SSCP analysis (2.2.6.1). DNA that contained known polymorphisms was electrophoresed beside the patient DNA. This allowed identical, polymorphic band patterns to be distinguished from unique patterns, due to mutations.

In twenty of the 24 samples, one altered, non-polymorphic band pattern of either the single-stranded and/or the double-stranded DNA was observed in just one of the 7 exons. Thirteen of these are shown in figures 19, 20 and 21. All but one of the 20 mutant patterns showed an alteration in the migration of single-stranded DNA, the exception being in family 8 (figure 20(B), lane 1). The hemizygote in family 8 was detected by a shift in the mobility of the homoduplex DNA. No altered band patterns were observed for individuals in families 3, 6, 29 and 30. In all cases, analysis of PCR products containing exons 2-5 produced two major bands of single-stranded DNA for each X-chromosome, representing the two complementary sequences and a shift in this pattern of single-stranded DNA indicated the presence of a mutation. However, the PCR products from exons 1 and 6 always gave more than one band for each single strand, indicating that several stable conformations existed during the electrophoresis. The simplicity of the band patterns contributed to the relative ease with which mutation screening was performed, as compared with sequencing.

One or more polymorphic band patterns were found in 4 hemizygotes, in addition to the non-polymorphic band pattern changes. The band patterns matched the 1P3, 1P1 and 7P1, 3P1 and 5P1, and 1P3 polymorphisms in hemizygotes from families 4, 7, 16 and 28, respectively. The presence of the 1P3 and 1P1 polymorphisms were confirmed by *Sac* II and *Msp* I digestion respectively and 3P1 and 7P1 were confirmed by sequencing (2.2.7).

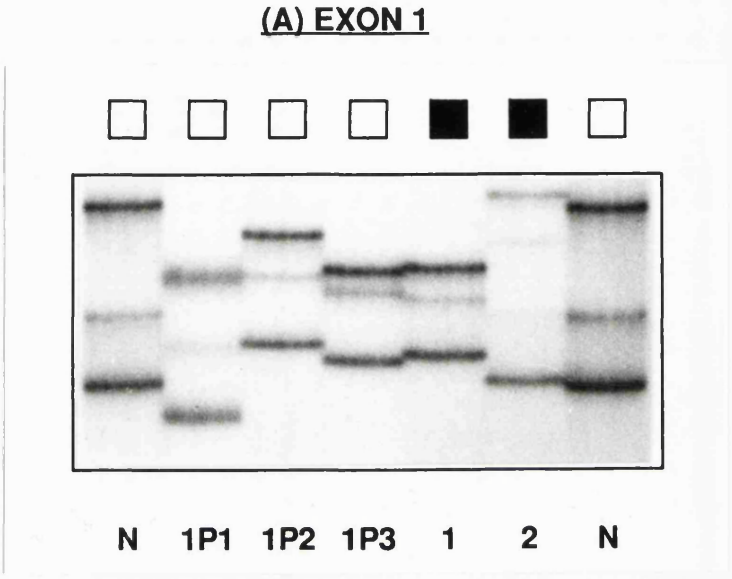


Figure 19: SSCP analysis of Fabry patient mutations in exons 1, 2 and 4

Radioactive SSCP analysis of Fabry mutations by electrophoresis through 6% polyacrylamide gels with 5% glycerol, at room temperature and at 320-360V (2.2.6.1). In each case, males were analysed and 'N' was the most common, normal pattern, for the exon indicated. In **(A)**, polymorphic patterns are indicated and lanes 1 and 2 are from hemizygotes in families 4 and 1, respectively. In **(B)**, and **(C)**, lane 1 shows the result of hemizygote analysis in families 17 and 28, respectively.

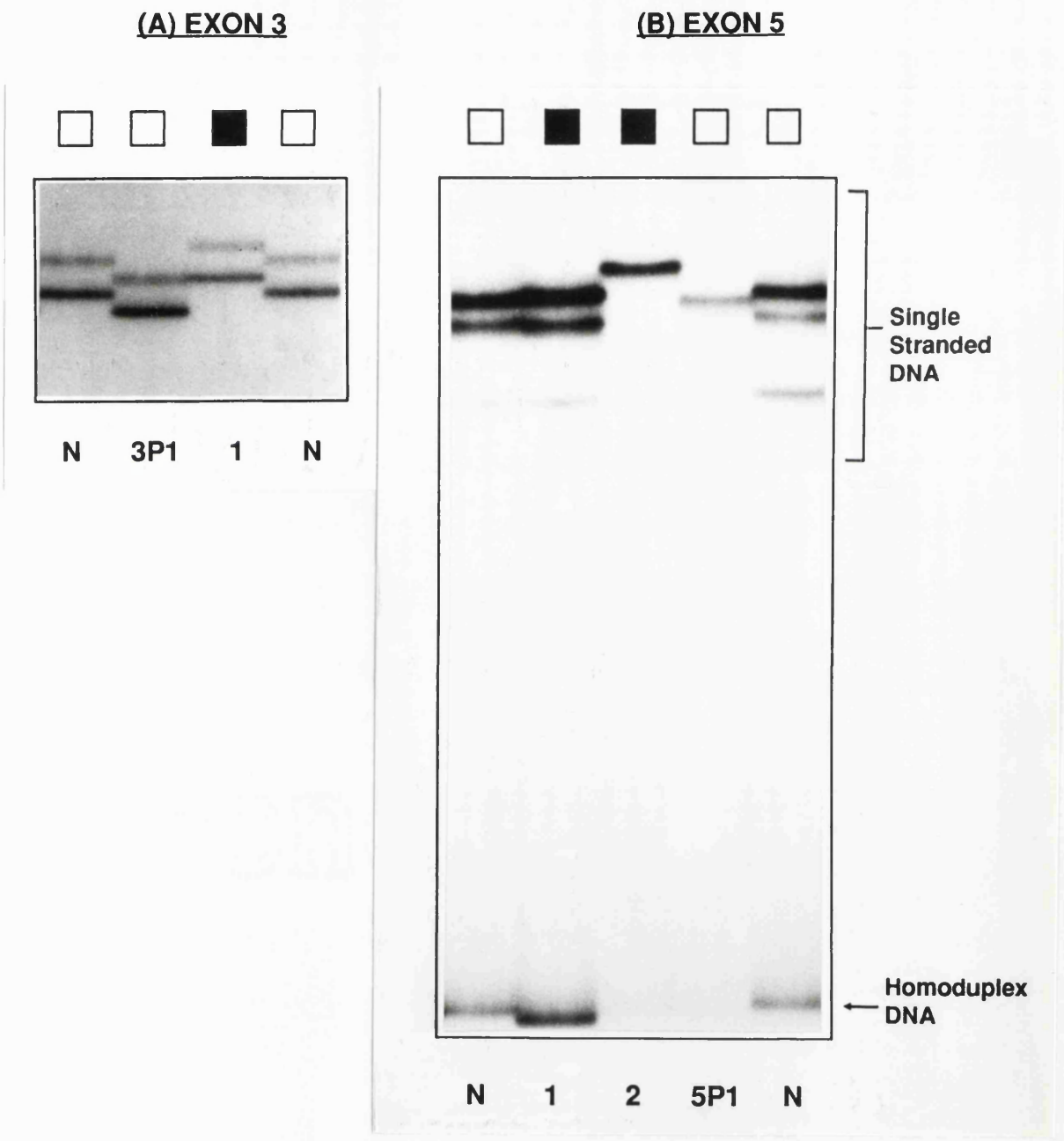


Figure 20: SSCP analysis of Fabry patient mutations in exons 3 and 5

Radioactive SSCP analysis of Fabry mutations by electrophoresis through 6% polyacrylamide gels with 5% glycerol, at room temperature and at 320-360V (2.2.6.1). In each case, males were analysed, 'N' was the most common, normal pattern and polymorphic patterns are indicated. In **(A)**, lane 1 shows analysis of exon 3 from a hemizygote in family 7. In **(B)**, lanes 1 and 2 show analysis of exon 5 for hemizygotes in families 8 and 20, respectively.

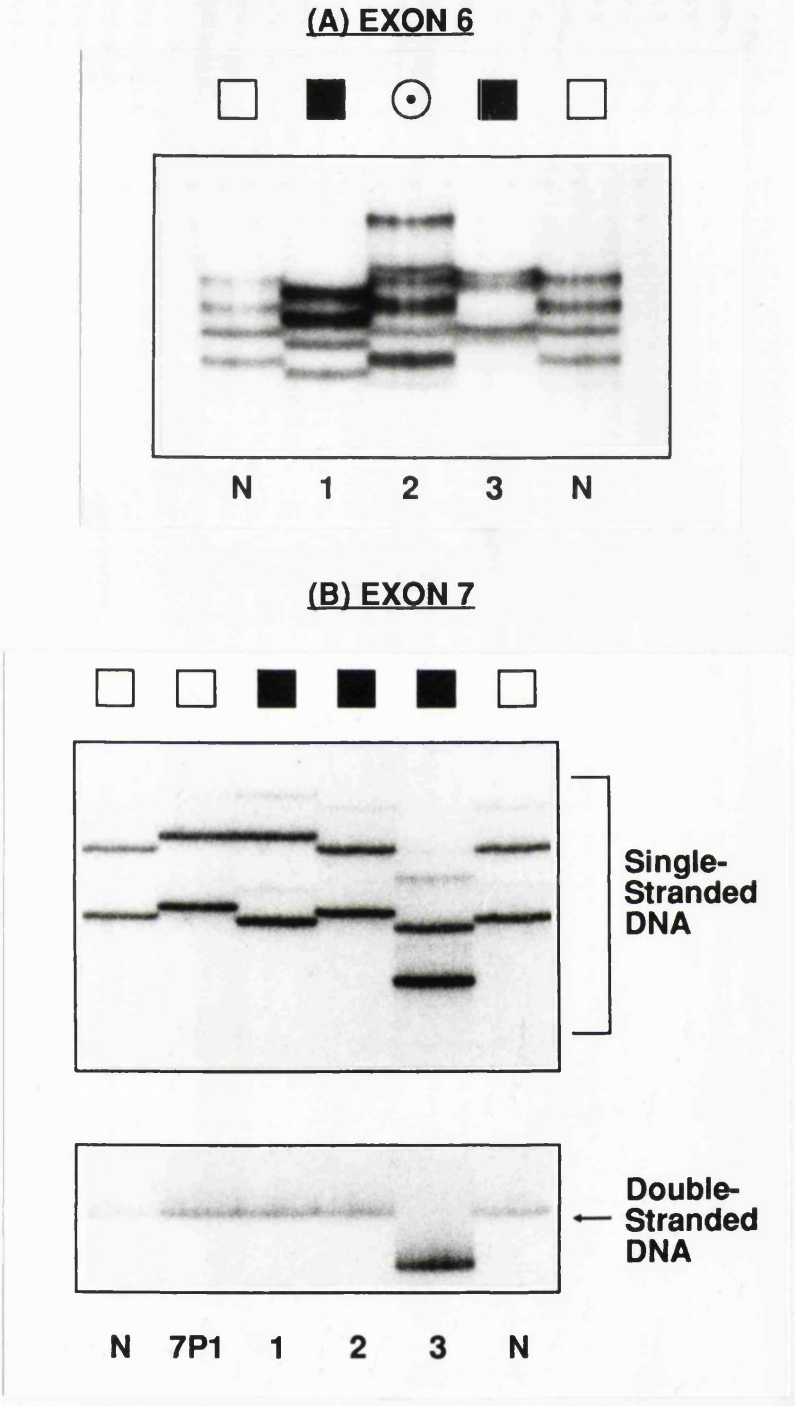


Figure 21: SSCP analysis of Fabry patient mutations in exons 6 and 7

Radioactive SSCP analysis of Fabry mutations by electrophoresis through 6% polyacrylamide gels with 5% glycerol, at room temperature and at 320-360V (2.2.6.1). In each case, males were analysed, except where indicated, and 'N' was the most common, normal pattern. In **(A)**, lanes 1, 2 and 3 show analysis of exon 6 from a hemizygote in family 10, a heterozygote in family 13 and a hemizygote in family 21, respectively. In **(B)**, the 7P1 polymorphic pattern is shown and lanes 1, 2 and 3 show analysis of exon 7 for hemizygotes in families 22, 16 and 26, respectively.

DNA containing the unknown mutations that produced the 20 unique band patterns observed by SSCP analysis, was sequenced. For exons 1 and 4, PCR-amplified DNA was sequenced by cycle-sequencing (2.2.7.3). For all other exons the DNA was PCR-amplified using one biotinylated primer and one non-biotinylated primer and the biotinylated strand was isolated and sequenced (2.2.7.2). Each sample was sequenced at least twice.

Mutations were found in all 20 patients with one being found twice in unrelated individuals (table 13, appendix, (A)). The phenotypes of patients with these mutations were known (2.1.1, tables 5 and figure 6), except for the heterozygote with Q327K and the hemizygotes with the D165V, N215S and R342X mutations. All were typical except for the hemizygote with G35R. The types of mutation included 15 different base substitutions in the protein-coding region, 3 small rearrangements in the protein-coding region and one splice site mutation. None of these had been previously described. The allele designations were as described in table 4, section 1.3.7.

All of the mutations were predicted to alter the amino acid sequence of the protein. Six point mutations, one of which, R227X, was found twice in unrelated hemizygotes, were predicted to substitute a stop codon for an amino acid and ten to cause substitution of an amino acid (table 13(A)). An adenine to thymine base substitution at the 3' intron 3 splice site (table 13(B)) was predicted to affect the mRNA splicing and therefore the protein sequence. All 3 frameshift mutations (table 13(C)) were predicted to create a premature stop codon. The insertion, 716ins1, in codon 239, replaces amino acid codons 241-249 (SILDWTSFN) with EYLGLDIF-stop and the deletion 717del2 (in codon 238) replaces amino acid codons 240-248 (KSILDWTSF) with EYLGLDIF-stop. In the 1010del19 mutant, codons 338-341 (EVWE) are replaced by SQA-stop.

Evidence for the disease-causing nature of these mutations is supported by the prediction that all can alter the amino acid sequence of the protein. In addition, the mutations had not been detected by SSCP analysis in at least 100 normal chromosomes and no other non-polymorphic SSCP changes were seen in the remaining 6 exons. Also, the exon containing the unique change was fully sequenced and did not contain any other mutations. Therefore, in 20 out of 24 samples, a single, putative disease-causing mutation was successfully identified by screening the exons using SSCP analysis, followed by sequencing.

Table 13: Mutations identified in 20 Families screened by SSCP analysis of all 7 exons of the α-galactosidase A gene

(A) Point mutations in the exons

Family No.	Pheno-type	Exon	Aileie (& poly-morphisms)	Nucleotide		Nucleotide Substitution	Nation -aiity
				cDNA	Gene		
1	variant	1	G 35 R	103	1282	GGA->AGA	British
4	typical		R 49 L (1P3)	146	1325	CGC->CTC	British
17	typical	2	Q 107 X	319	5218	CAG ->TAG	British
7	unknown	3	D 165 V (1P1, 7P1)	494	7393	GAT->GTT	British
8	unknown	5	N 215 S	644	10135	AAT->AGT	British
18	typical		W 226 X	677	10168	TGG->TAG	Dutch
19	typical		R 227 X	679	10170	CGA->TGA	British
20	typical		R 227 X	679	10170	CGA->TGA	British
9	typical		W 236 C	708	10199	TGG->TGC	German
10	typical	6	V 269 A	806	10514	GTG->GCG	British
21	typical		W 287 X	861	10569	TGG->TGA	British
12	typical		V 316 E	947	10655	GTA->GAA	British
13	unknown		Q 327 K	979	10687	CAA->AAA	British
14	typical	7	W 340 R	1018	10996	TGG->CGG	German
22	unknown		R 342 X	1024	11002	CGA->TGA	Greek
16	typical		G 361 R (3P1, 5P1))	1081	11059	GGA->AGA	British

(B) Splice site mutations in the introns

Family No.	Pheno -type	Splice Site	Gene Nucleotide	Nucleotide Substitution	Linked Polymor -phisms	Nationality
28	typical	3' Intron 3 (IVS3-1,t)	8321	ag-exon 4 >at	1P3	Belgian

(C) Rearrangements in the exons

Family No.	Pheno -type	Exon	Allele	cDNA Sequence change Normal -> mutant	Rearranged gene (& cDNA) nucleotides	Nation-ality
23	typical	5	716ins1	714-TA$\underline{I}$AAAG-720 -> TA$\underline{II}$AAAG	between 10207- 10208 (716-717)	British
24	typical	5	717del2	714-TAT$\underline{AAA}$G-720 -> TAT$\underline{A}$G	between 10208-10210 (717-719)	Belgian
26	typical	7	1010del19	1008-CT$\underline{I}$T-CC$\underline{I}$CT-1031 -> CT$\underline{I}$-CT	between 10988-11007 (1010-1029)	British

Table 13: Mutations identified in 20 Families screened by SSCP analysis of all 7 exons of the α-galactosidase A gene

Putative disease-causing mutations were identified in the α-galactosidase A gene of hemizygotes from Fabry families 1, 4, 7, 8, 10, 12, 16-24, 26 and 28 and also in heterozygotes from families 9, 13 and 14. Detection was by screening the 7 exons using radioactive SSCP analysis 2.2.6.1) and then sequencing the altered exon (2.2.7).

Underlining in **(A)**, indicates the base altered for point mutations and in **(C)** the repeated sequences involved in deletions and insertions. Known polymorphisms were detected by radioactive SSCP analysis. Gene numbering was as in Kornreich *et al.* (1989) and cDNA nucleotides were numbered as in Bishop, *et al.* (1991).

In four families, SSCP analysis did not detect a sequence change and they required further analysis.

3.1.2.1.2 Mutation detection by sequencing all 7 exons and splice sites

SSCP analysis of DNA from hemizygotes in families 3, 6, 29 and 30 had failed to detect any sequence alterations in the protein-coding region or the splice sites. An alternative strategy to detect their mutations was to sequence the same region to determine whether any sequence changes were present in this region. DNA containing the 7 exons and splice sites was PCR-amplified from these four patients, using one biotinylated and a non-biotinylated primer (2.2.4, table 7). The biotinylated strand was isolated and sequenced using Sequenase II (2.2.7.2). DNA from a further 6 hemizygotes, in families 2, 5, 11, 15, 25 and 27, was also fully sequenced in this manner. These 6 samples had not been screened by SSCP analysis and were obtained in collaboration with Dr RJ Desnick in New York.

Eight mutations, none of which had been reported in other studies, were detected in 8 of the 10 samples analysed (table 14, appendix (B)). All were found in Fabry families with the typical symptoms (2.1.1), except for 717del2, which was found in a hemizygote with an unknown phenotype. Of these, 2 mutations, R49S and D92H, had previously been undetected by SSCP analysis. Mutations were also found in DNA from all 6 hemizygotes that had not been analysed by SSCP analysis. No mutation was found in hemizygotes from families 29 and 30. In these 2 individuals, all seven exons and splice sites were sequenced but some parts of the of the gene, in particular the GC-rich areas in exons 1 and 3, were difficult to read and analysis of the whole region needs to be repeated for these two hemizygotes. No further genetic analysis has been conducted in these two families.

The eight mutations consisted of 6 base substitutions, each of which was predicted to alter an amino acid, and 2 small deletions within the exons. One of the deletions, 717del2, in British family 25, had also been detected in the Belgian family, 24, (table 13) by SSCP analysis and its predicted effect has been described. The predicted effect of the 1087del1 mutation was to replace amino acids 363-389 and to create a premature stop codon at codon 390.

(A) Point Mutatations in the exons

Family No.	Pheno -type	Exon	Allele	Nucleotide		Nucleotide Substitution	Nation -ality
				cDNA	Gene		
2	typical	1	M 42 V	124	1303	ATG->GTG	British
3	typical		R 49 S*	145	1324	CGC->AGC	British
5	typical		C 56 Y	167	1346	TGC->TAC	British
6	typical	2	D 92 H*	274	5173	GAT->CAT	British
11	typical	6	W 287 G	859	10567	TGG->GGG	British
15	typical	7	R 342 Q	1025	11003	CGA->CAA	British

(B) Rearrangements in the exons

Family No.	Pheno-type	Exon	Allele	cDNA sequence change Normal -> mutant	Rearranged gene (& cDNA) nucleotides	Nation -ality
25	unknown	5	717del2	714-TATAAAG-720 -> TATAG	between 10208-10210 (717-719)	British
27	typical	7	1087del1	1085-CTCGCTC-1091 -> CTC-CTC	11066 (1088)	British

Table 14: Mutations identified in 8 families by sequencing all 7 exons of the α-galactosidase A gene

Putative disease-causing mutations were identified in the α-galactosidase A gene of hemizygotes from Fabry families 2-4, 6, 11, 15, 25 and 27 by sequencing all seven exons (2.2.7.2). Gene numbering was as in Kornreich et al. (1989) and cDNA nucleotides were numbered as in Bishop, et al. (1991).

In **(A)**, * = undetectable by SSCP analysis but detected by sequencing.
In **(B)**, underlined bases are repeated nucleotide sequences.

Evidence for the disease-causing nature of these mutations comes from the prediction that they all alter the amino acid sequence of the mutant protein. Also, complete sequencing of the 7 exons and splice sites has shown conclusively that no other mutations are present in this region. However, the presence of other potentially disease-causing mutations outside of this region was not excluded. Further evidence that they were disease-causing mutations was provided by SSCP analysis. DNA containing the M42V, C56Y, W287G, R342Q and 1087del1 mutations, which had not been analysed by SSCP screening, was amplified and subjected to SSCP analysis. All 5 mutations gave an altered band pattern that was not observed in at least 100 non-Fabry DNA samples, providing evidence that they were not polymorphisms. Detection of these mutations by SSCP analysis is potentially useful for carrier detection.

The two samples from families, 29 and 30 require further analysis as neither SSCP screening nor sequencing of the same region detected a mutation.

3.1.2.2 Carrier detection in families with an identified mutation

Three methods for carrier detection, other than direct sequencing, were used to detect carriers of mutations in the 28 families in which a putative disease-causing mutation had been identified.

The standard SSCP analysis method was potentially useful for mutation detection in 26 families. The C56Y, D165V, W287G, V316E, Q327K, R227X, 1010del19 and 1087del1 mutations were detected in carriers, using this method (table 16, pg 149). SSCP analysis also detected the presence of N215S, Q107X, W226X, R342X and 716ins in the mothers of the hemizygotes analysed in families 8, 17, 18, 22 and 23, respectively. Thus, the method proved useful for analysis of 13 mutations. Another mutation, D92H, was detected by modifying the standard method. Digestion of the exon 2 PCR product with *Hinf* I, prior to radio-labelled SSCP analysis (figure 22 and table 16), allowed its identification. The undigested 274bp product did not give a band shift by SSCP analysis but the smaller digestion product, 168bp, which contained the D92H mutation, did show an SSCP change.

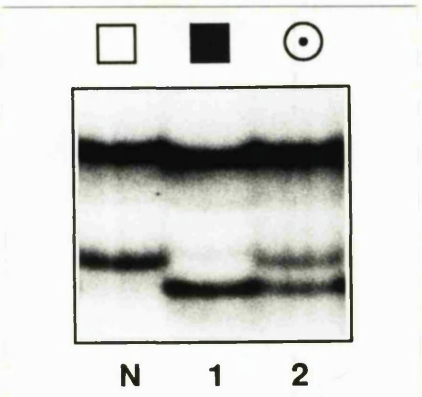


Figure 22: Detection of the D92H allele by *Hinf* I digestion and SSCP analysis

The full-length 274bp PCR-product for exon 2 was digested with *Hinf* I (2.2.8), yielding fragments of 168bp, 47bp and 59bp, prior to electrophoresis on a non-denaturing gel (2.2.6.1.2). The D92H mutation was in the 168bp fragment.

N is the normal pattern, 1 and 2 show the analysis in a Fabry hemizygote and heterozygote, for D92H, respectively.

In addition to SSCP analysis, 10 mutations were predicted to alter the recognition site for a restriction enzyme (table 15), which could also be used for heterozygote detection. Four of these, R49S, R49L, M42V and G361R were tested, using a faster and simpler gel electrophoresis system than for SSCP analysis (figure 23 and table 16). Since one of these, R49S, could not be detected by SSCP analysis, restriction digestion provided an alternative convenient method for its detection (figure 23(A)). The R49L mutation produced an identical band pattern to R49S as it deleted the same restriction recognition site. Analysis of the M42V mutation by digestion with *Nco* I (figure 23(B)), showed that while the affected hemizygote III-1 and his mother II-1, carried the mutation, his grandmother I-1, did not and since the grandfather was not known to have Fabry disease, the mutation must have arisen *de novo* in either I-1 or II-1, unless II-1 was a case of non-paternity. Analysis of the G361R mutation in family 16 (figure 23(C)) confirmed that female III-1 was not a carrier, consistent with the results from polymorphism analysis, section 3.1.1.4.2.

3.1.2.3 Correlation of the carrier status with α-galactosidase A activity

The genetic diagnosis of carriers in females was compared with biochemical diagnosis by enzyme assay. This analysis enabled the range of enzyme activity produced by genetically defined carriers to be established.

Designation of genotype in female members of families with Fabry disease by measurement of α-galactosidase activity is problematic because a heterozygote can have activity ranging from zero to full activity, due to X-chromosome inactivation. The α-galactosidase activity previously measured in the plasma and leukocytes of the families studied was compared with their genotype, determined by molecular methods (table 16). Measurement of the α-galactosidase activity in both leukocytes and plasma successfully identified 10/18 (56%) of the genetically defined carriers (table 17). In a further 4 heterozygotes, one of the two assays identified the heterozygotes. In the remaining 4/18 heterozygotes (22%) biochemical diagnosis failed to identify carriers by measurement of activities in both plasma and leukocytes. Genetic analysis showed that all females in Fabry families with activity below the known normal range of values, were heterozygotes. It also showed that all genetically normal females in the families had activity in the normal range. However, some genetically defined heterozygotes had α-galactosidase activity in the normal range, emphasising the need for molecular genetic detection of heterozygotes for counselling.

Family	Allele	Restriction enzyme		Cut fragment size (bp)		Normal PCR product size (bp)
		Name	Recognition site	Normal	Mutant	
2	M42V	-*Nco* I	C / CATGG	248+102	350	350
4 3	R49S R49L	-*Hha* I	GCG / C	163+101+78 +6+2	163+179+6 +2	350
7	D165V	-*Bgl* II	A / GATCT	269+131	400	400
8	N215S	-*Tsp*E I	/ AATT	92+194	286	286
18	W226X	+*Rma* I +*Mae* I	C / TAG	286	125+161	286
21	W287X	-*Bgl* I	$GCCN_4$ / NGCC	155+202	357	357
12	V316E	-*Mae* II	A / CGT	240+117	357	357
16	G361R	-*Sau*96 I	G / GNCC	17+131+293	17+424	441
26	1010d19	-*Eco*N I	$CCTN_2$ / N_3AGG	98+343	422	441

Table 15: Alterations in the size of DNA fragments produced by restriction enzyme digestion of PCR products containing mutations

'/' indicates the cut site in the restriction enzyme recognition sites and N is any base. The - and + indicate the loss and gain of a restriction site, respectively.

Figure 23: Detection of mutations in Fabry patients by restriction enzyme digestion analysis

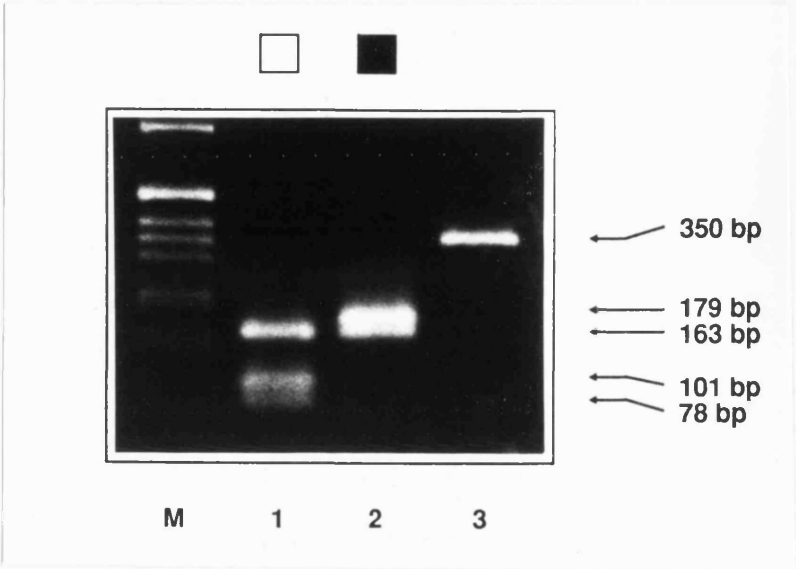


Figure 23 (A): Detection of the R49S mutation by *Hha* I restriction enzyme analysis

The PCR product for exon 1 (2.2.4) was digested with *Hha* I(2.2.8) and electrophoresed on a 3% agarose gel (2.2.5). M is the 1kb ladder DNA marker (Gibco BRL), 1 is the normal digestion pattern, yielding fragments of 163bp, 101bp, and 78bp, 2 is the pattern in the R49S mutant, with 163bp and 179bp fragment sizes and 3 is an undigested sample.

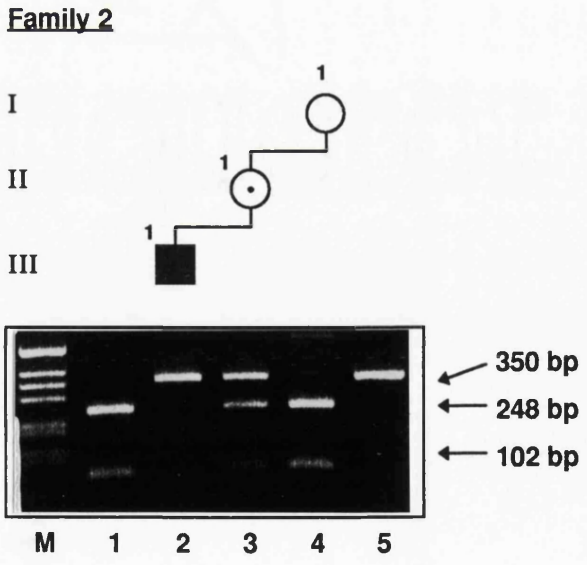


Figure 23 (B): Detection of the M42V mutation by *Nco* I restriction enzyme analysis

The PCR product for exon 1 (2.2.4) was digested with *Nco* I (2.2.8) and electrophoresed on a 2% agarose gel (2.2.5). M is the 1kb ladder DNA marker (Gibco BRL), 1 is the normal digestion pattern, yielding fragments of 248bp and 102bp, 2, 3, and 4 are indicated on the pedigree diagram as III-1, II-1 and I-1. 5 is undigested sample.

Figure 23, cont.: Detection of mutations in Fabry patients by restriction enzyme digestion analysis

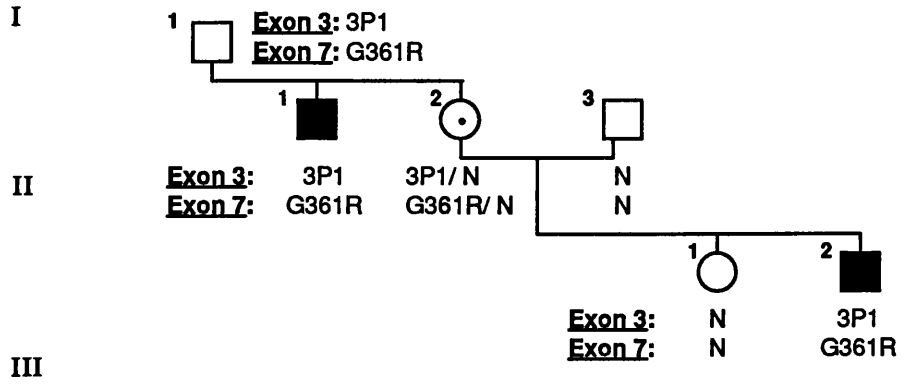


Figure 23 (C): Detection of the G361R mutation by *Sau*96 I restriction enzyme analysis

The PCR product for exon 7 (2.2.4) was digested with *Sau*96 I (2.2.8) and electrophoresed on a 2% agarose gel (2.2.5), to detect the G361R mutation that deletes this restriction enzyme site.

.

Family No.	Mutation	Location in pedigree chart *	α-galactosidase activity in			Genetic analysis
			leukocytes (nmol/ hr/ mg protein) Normal range: 40-162	plasma (nmol/ hr/ ml plasma) Normal range: 4.8-26.5	Activities are within the range for normals (N) or below the normal range (C)	
3	R49S	II-4 II-8 III-3 III-5 III-7	17 21 80 44 39	2.6 3.7 7.4 5.3 4.9	C C N N CN	Carrier Carrier Normal Carrier Carrier
5	C56Y	I-2	118	9.7	N	Normal
6	D92H	I-1 I-2 II-5 III-1	64 28 10.6 53	6.5 5.3 3.9 2.9	N CN C CN	Carrier Carrier Carrier Carrier
7	D165V	III-1 III-2 III-3	15.6 5.2 74	1.5 0.39 9.8	C C N	Carrier Carrier Normal
11	W287G	II-6 I-2	26 64	3.0 6.2	C N	Carrier Normal
12	V316E	I-2 II-2	67 104	4.8 6.8	N N	Carrier Normal
13	Q327K	I-4	56	9.0	N	Carrier
16	G361R	III-2	69	9.9	N	Normal
19	R227X	II-4	37	2.5	C	Carrier
20	R227X	IV-13 V-18 VI-2	21.8 59 37	1.5 3.8 3.0	C CN C	Carrier Carrier Carrier
26	1010d19	II-3	31	4.7	C	Carrier
27	1087d1	II-4 III-4 III-1	112 80 95	8.2 7.4 12.2	N N N	Normal Normal Normal

Table 16: Comparison of heterozygote analysis by enzyme assay and genetic analysis

The pedigrees, '*', are shown in section 2.1.1. All α-galactosidase activity measurements were carried out by Mrs E. Young in the Enzyme laboratory, Institute of Child Health.

Alpha-galactosidase activity (C= enzyme activity below normal range, N= activity in normal range)		Frequency in genetically diagnosed females	
Leukocytes	**Plasma**	**Carriers (18)**	**Normals (9)**
N	N	4	9
N	C	2	0
C	N	2	0
C	C	10	0

Table 17: Comparison of enzyme diagnosis and genetic diagnosis in females

The results from table 16 were summarised. Biochemical diagnosis was by measurement of the α-galactosidase activity in both leukocytes and plasma and genetic diagnosis was by detection of family-specific mutations.

The extent of overlap between the ranges of activities for genetically-defined heterozygotes and for normals is shown in figure 24.

This genetic study also showed that the family history is not always a reliable indicator of carrier status. Figure 25(A) shows that female III-3, a presumed obligate carrier, has not inherited the D165V mutation from her affected father in family 7. This is consistent with the pattern of inheritance of the polymorphism 1P1 and suggests that II-1 is not her father. In family 5 (figure 25(B)), female I-2 is the mother of a Fabry hemizygote but she does not carry the mutation in the DNA sample analysed, in agreement with enzyme activity in the normal range. She is either a mosaic for the mutation, C56Y, or the mutation has arisen *de novo* in her son. In family 7 (figure 25(C)), female II-6, who had α-galactosidase activity below the normal range, carried the W287G mutation. Her mother, I-2, did not have the mutation in lymphocyte DNA and although her father was not analysed, he was not thought to have Fabry disease. Therefore, either female II-6 is a case of non-paternity or the mutation has arisen *de novo* in the germ-line of one of her parents or during her own development.

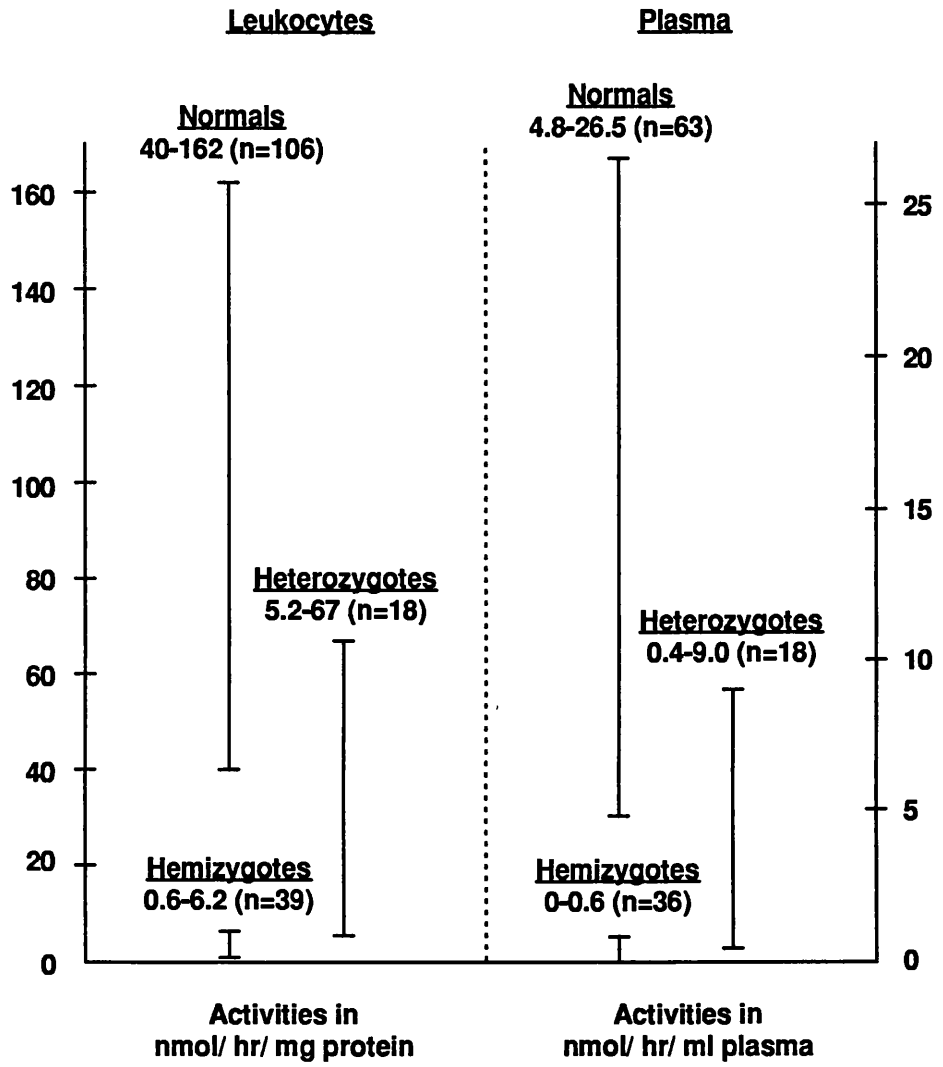
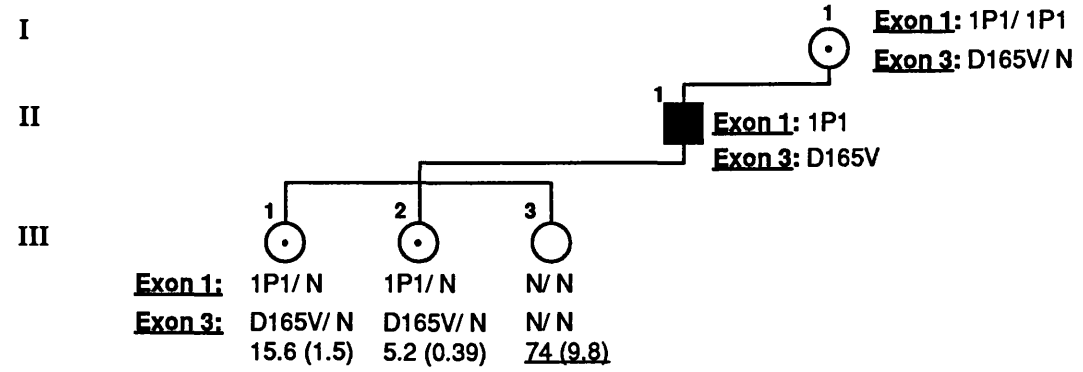


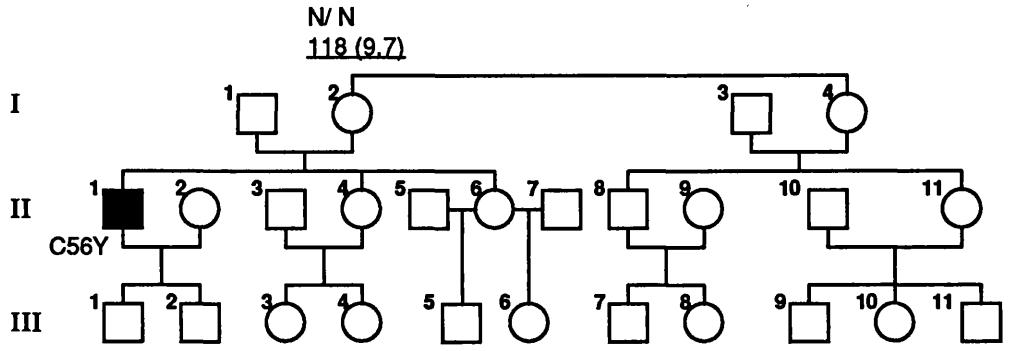
Figure 24: Alpha-galactosidase activity ranges for normals, clinically and biochemically defined hemizygotes and genetically defined heterozygotes

All α-galactosidase activity measurements were carried out by Mrs E. Young in the Enzyme Laboratory, Institute of Child Health and 'n' is the number of samples analysed. The mean normal activity for leukocytes and plasma was 72.4 and 8.6, respectively.

(A) Pedigree for family 7 and segregation of the D165V mutation and 1P1



(B) Pedigree for family 5 and segregation of C56Y mutation



(C) Pedigree for family 11 and segregation of W287G mutation

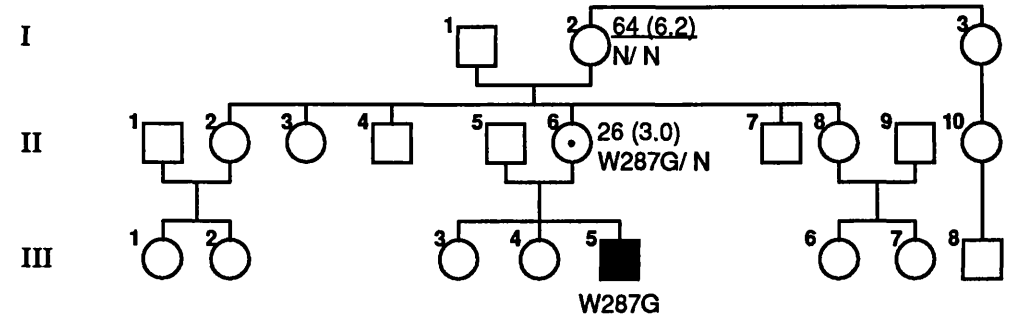


Figure 25: Pedigrees and the results of mutation detection in families 5, 7 and 11

The results of the mutation detection are shown. Enzyme activities, in nmol/hr/mg protein in leukocytes (and in nmol/hr/ml of plasma, in plasma) are shown and underlining indicates values that were within the normal range.

3.2 STUDIES OF MUTANT ALPHA-GALACTOSIDASE A

The effect of mutations in the α-galactosidase A gene on the enzymic activity was investigated by studying the residual activity in cultured cells from patients and by the transient expression in mammalian cells of mutated cDNA. This work was carried out in collaboration with Dr RJ Desnick and Dr Y Iounnou (Mount Sinai Medical Center, New York, USA), in their laboratories.

3.2.1 Enzyme activities in cultured cells from Fabry patients

The α-galactosidase A activity in fibroblasts and lymphoblastoid cell lines derived from Fabry hemizygotes was measured to see if there was any correlation between residual α-galactosidase A activity and phenotype (table 18).

All the patients, except the individual with the N215S mutation in family 8, had α-galactosidase A activities that were less than 1% of the mean normal activity. Six of the cell lines producing less than 1% of the normal activity were from hemizygotes that were classified as having the typical Fabry phenotype, in families 18, 10, 28, 3, 12 and 27. The phenotypes associated with Q327K and R342X, in families 13 and 22 were unknown (2.1.1). Lymphoblastoid cells with the N215S mutation had 6.7% of the mean, normal activity (32.6nmol/mg/ml). Although the phenotype of this patient in family 8 was unknown, four other unrelated hemizygotes with the same mutation were mildly affected (Eng *et al.*, 1993). Therefore, the highest enzyme activity was found in cells from a patient that was suspected of having a mild phenotype.

3.2.2 Transient expression of mutant cDNA in mammalian cells (COS-1)

The purpose of the expression experiments was threefold. Firstly, to verify that the mutations decreased the specific activity of the enzyme and were therefore probably disease-causing. The second aim was to investigate whether there was a relationship between the phenotype of the patient and the amount of activity produced by the known mutation for those patients from whom cell lines were not available. The third aim was to compare this vector and transfection method with the transient expression system used by others (Ishii *et al.*, 1992).

Cell type (and normal enzyme activity, in nmol/hr/mg)	Cells from Fabry hemizygote in family No.	Associated phenotype	Mutation	α-galactosidase A specific activity in cells from hemizygotes (nmol/hr/mg)		
				Range of enzyme activities	Mean	% of normal mean activity
Cultured	18	typical	W226X	0-1.20	0.66	0.7
Fibroblasts	10	typical	V269A	0-1.15	0.75	0.8
(normal range	13	unknown	Q327K	0.13-0.64	0.40	0.4
& mean =	22	unknown	R342X	0.16-1.18	0.60	0.6
75.1-143 & 98.7)	28	typical	IVS3-1,t splice mutation	0-0.04	0.03	0.03
Lymphoblastoid	3	typical	R49L	0.06-0.42	0.28	0.8
Cells	8	unknown	N215S	2.03-2.28	2.20	6.7
(normal range & mean =	12	typical	V316E	0.03-0.30	0.20	0.6
21.5-50 & 32.6)	27	typical	1087del1	0-0.16	0.08	0.2

Table 18: Alpha-galactosidase A activity in cultured cells from Fabry hemizygotes

Cell lysates were assayed for protein and α-galactosidase A activity (2.3.1). Protein assays were in duplicate. The α-galactosidase A activity was assayed in 2 sets of duplicates, except for the W226X mutant, which was from 4 sets of readings in duplicate. The normal activity data were provided by CM Eng and DF Bishop, Mount Sinai Medical Center, New York. The normal activity ranges in fibroblasts and lymphoblastoid cells were obtained from 5 and 150 samples, respectively.

The cDNAs containing the G35R, R49L, V269A, V316E, Q327K and G361R mutations identified in this study were transiently expressed to investigate their effect on the activity. One previously published mutation, R112H (Eng *et al.*, 1994b) was found in a patient with a variant phenotype but whose enzyme activity was unknown. Two other previously published mutations found in mildly affected hemizygotes, Q279E and R301Q (Sakuraba *et al.*, 1990; Nagao *et al.*, 1991; Ishii *et al.*, 1992), were also expressed in this study for comparison with the results obtained by transient expression in a previous study (Eng *et al.*, 1994b), which used a different vector and transfection method for expression.

3.2.2.1 Site-directed mutagenesis and subcloning into expression vector pMT2

The pGB6 vector (figure 7), containing full length α-galactosidase cDNA was used as the template for mutagenesis. The megaprimer method of site-directed mutagenesis was performed (figure 26, section 2.3.3). Two separate steps involving PCR amplification were used to introduce a missense mutation into the cDNA and the resulting PCR product was gel-purified (steps 1-4). *Vent* polymerase was used in the amplifications as it has a higher fidelity of polymerisation than the more commonly used *Taq* polymerase. The mutant PCR product was cut with *Eco*R I and subcloned into *Eco*R I sites in pGEM-3Z, step 5 (figure 26). A single clone was picked and a region, which contained the mutation and was between two unique restriction enzyme sites, was completely sequenced. This was carried out to confirm the presence of the correct mutation and to verify that no other mutations had been introduced by *Vent* polymerase. The location of the mutations in relation to unique restriction sites in the cDNA and the size of the regions sequenced are indicated below:

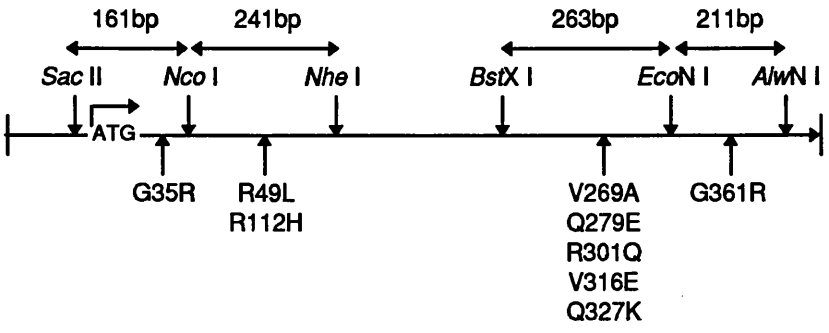
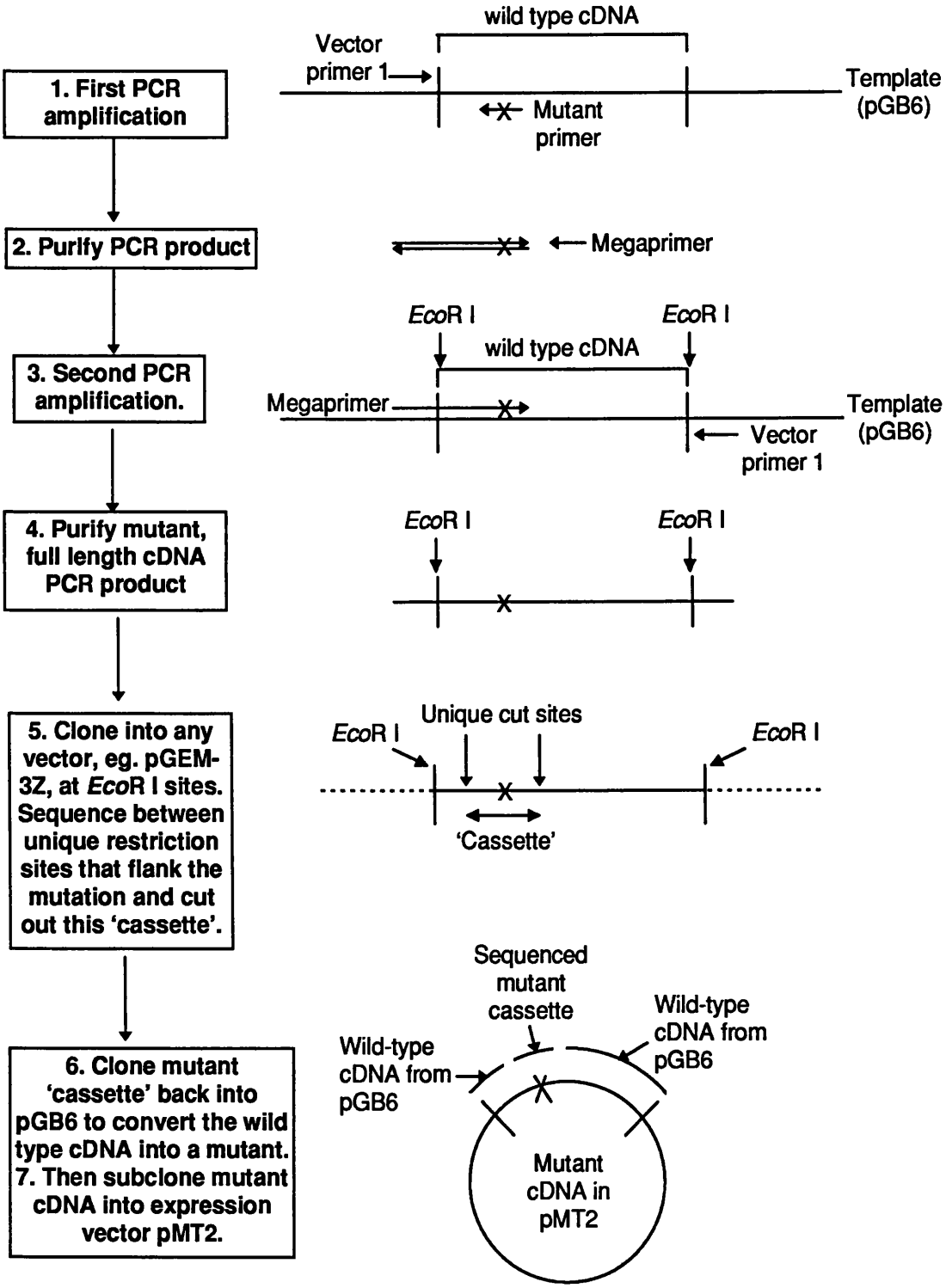
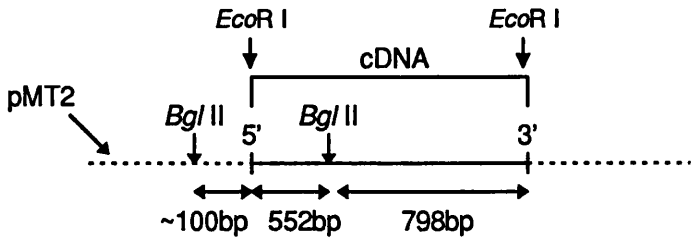


Figure 26: Strategy for the megaprimer method of site-directed mutagenesis and cloning



Sequencing of the region flanked by *Nco* I and *Nhe* I sites that contained the R112H mutation is shown in figure 27. All other constructs were sequenced in this manner. The sequenced region that contained each mutation was cut at the unique restriction enzyme sites and inserted back into the wild-type cDNA in pGB6 (figure 26, step 6). The presence of the mutation in pGB6 was verified by sequencing and then the whole mutant cDNA was cut out at *Eco*R I sites. The cDNA was subcloned into *Eco*R I sites of pMT2, a mammalian expression vector (Kaufman *et al.*, 1989, section 2.3.5.1). The presence of the mutation was again verified by sequencing. In addition, the orientation of the insert and the presence of one insert only, was checked by digestion with *Bgl* II. It cuts the vector once, at about 100bp upstream of the 5' end of the cDNA and the cDNA once, 552bp downstream from the 5' cDNA end, as indicated below:



If the insert is in the correct orientation a single fragment of approximately 650bp is released from pMT2 (a 5.1kb vector).

The 9 mutated cDNA constructs and the normal cDNA gave fragments of approximately 650bp when digested with *Bgl* II, indicating that all of them had been inserted into pMT2 in the correct orientation.

3.2.2.2 Transfection of the mutant α-galactosidase A cDNA, without cotransfection of β-galactosidase

In each experiment, COS-1 cells from a single pool were subdivided into equal portions and each portion was transfected by electroporation with either a mutant α-galactosidase A construct or the wild-type cDNA. The control consisted of a mock transfection lacking an α-galactosidase A construct, to measure the endogenous α-galactosidase activity in the COS-1 cells. The α-galactosidase activity was measured in 10ml of medium, or 1ml of cell lysate (2.3.1.3) and the results were related to the total cell protein, measured in 1ml of cell lysate (2.3.1.2). The relative activities were calculated as indicated below:

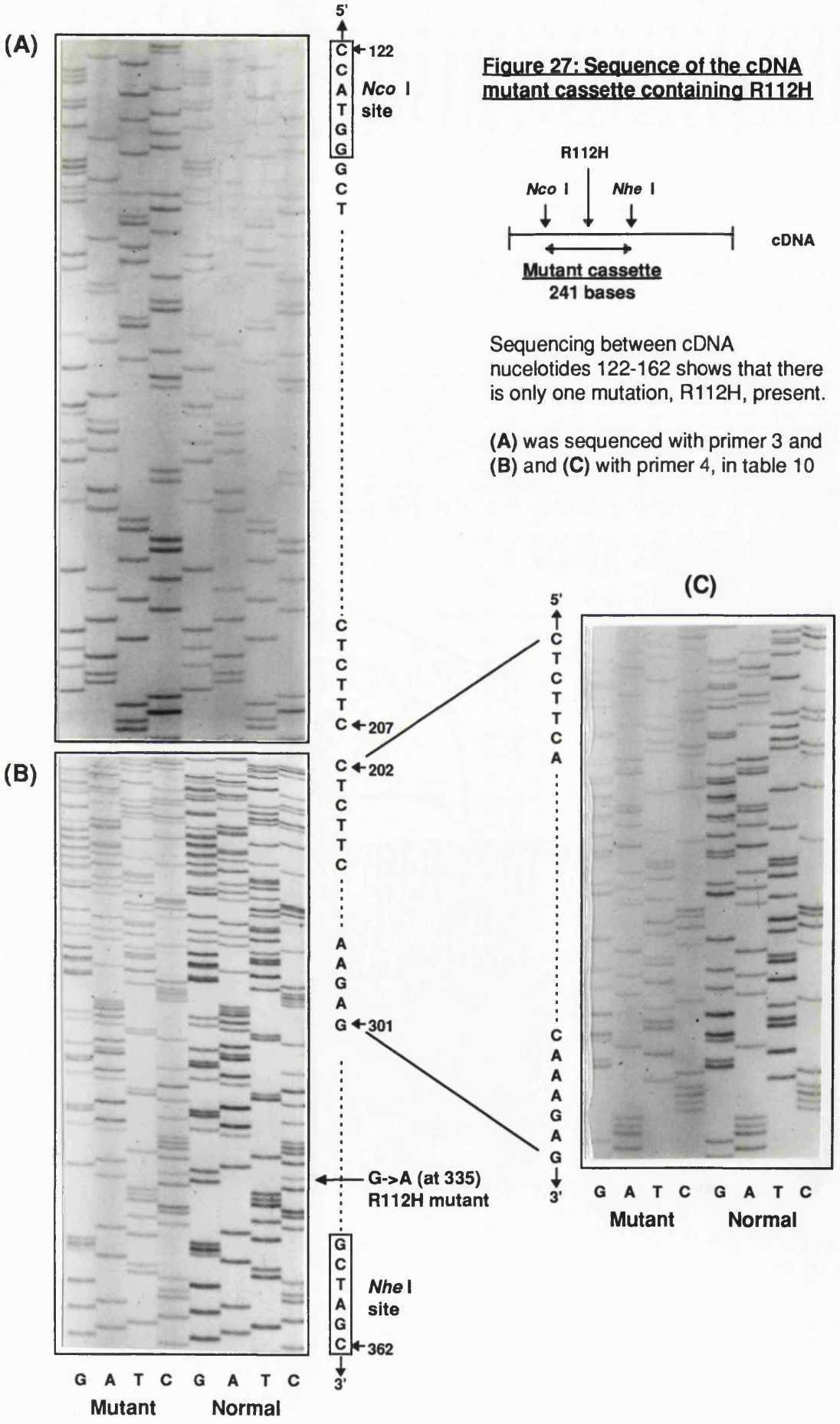


Figure 27: Sequence of the cDNA mutant cassette containing R112H

Sequencing between cDNA nucelotides 122-162 shows that there is only one mutation, R112H, present.

(A) was sequenced with primer 3 and **(B)** and **(C)** with primer 4, in table 10

$$\text{Activity (nmol/hr/mg)} = \frac{\alpha\text{-galactosidase A activity (nmol/hr/ml)} \times Z}{\text{total cell protein (mg)}}$$

where 'Z' was 1ml and 10ml for expressed intracellular and extracellular activity, respectively. The elevation of the expressed activity over endogenous COS-1 cell activity was calculated by:

$$\frac{\text{Mutant cDNA activity (nmol/hr/mg)}}{\text{Control (no construct) activity (nmol/hr/mg)}}$$

The percentage increase in activity, compared with the normal, above the endogenous activity, was calculated by:

$$\frac{\text{Mutant cDNA activity - control activity (nmol/hr/mg)}}{\text{Normal cDNA activity - control activity (nmol/hr/mg)}} \times 100$$

The results from 2 transfection experiments are shown in table 19. In the first, all 9 constructs were transfected and their relative activities compared (table 19(A) and figure 28). All caused a decrease in α-galactosidase A activity, in accordance with their presumed disease-causing status. Five constructs, R49L, R112H, V269A, V316E and Q327K produced intracellular activity that was less than twice the activity of the endogenous activity control and their activity in the medium was less than 6 fold higher than the control. They were considered to be mutations producing negligible activity. Four constructs, G35R, Q279E, R301Q and G361R, produced low residual activity above the control. The activity was more than 2 fold higher in the cells and two, Q279E and R301Q produced activity that was more than 6 fold higher in the medium than the control (table 19(A)). They were categorised as mutants producing residual enzyme activities and were ranked in their order of activity, with 1 as the highest value (table 20).

(A) Transfection 1

α-gal. Construct	Transfection 1, at 47hr					
	α-gal activity (nmoi/hr/mg)		Mutant α-gai activity/ control		% of normal activity	
	Ceiis	Medium	Cells	Medium	Cells	Medium
G35R	432	38	9.6	3.5	11.9	0.6
Q279E	189	121	4.2	11.4	4.4	2.6
R301Q	127	36	2.8	3.4	2.5	0.6
G361R	114	407	2.5	38.4	2.1	9.3
R49L	49	54	1.1	5.1	0.1	1.0
R112H	88	47	1.9	4.5	1.3	0.9
V269A	50	36	1.1	3.4	0.1	0.6
V316E	39	0	0.9	0	-0.2	-0.3
Q327K	37	24	0.8	2.2	-0.2	0.3
Normal	3286	4239	72.7	399.9	100	100
None (control)	45	11	1	1	-	-

(B) Transfection 2

α-gai. Construct	Transfection 2, at 48hr					
	α-gal activity (nmol/hr/mg)		Mutant α-gal activity/ control		% of normal activity	
	Cells	Medium	Ceils	Medium	Cells	Medium
G35R	3012	327	18.5	96.0	19.7	2.7
Q279E	2117	569	12.9	167.3	13.5	4.7
R301Q	1027	44	6.3	12.9	6.0	0.3
G361R	417	1047	2.6	307.9	1.8	8.7
Normal	14618	12011	89.4	3532.5	100	100
None (control)	163	3	1	1	-	-

Table 19: Comparison of α-galactosidase activities related to total cell protein

10μg of each α-galactosidase A construct was transfected into COS-1 cells. Alpha-galactosidase A activity was measured in the medium and cell lysate (2.3.1.3). The α-galactosidase B inhibitor, GalNAc, was not included in the assay for transfection 1 **(A)**, but was included for transfection 2, **(B)**. The activities are compared relative to the control (no construct) endogenous activity. They are also compared as a % of the normal, above the control activity.

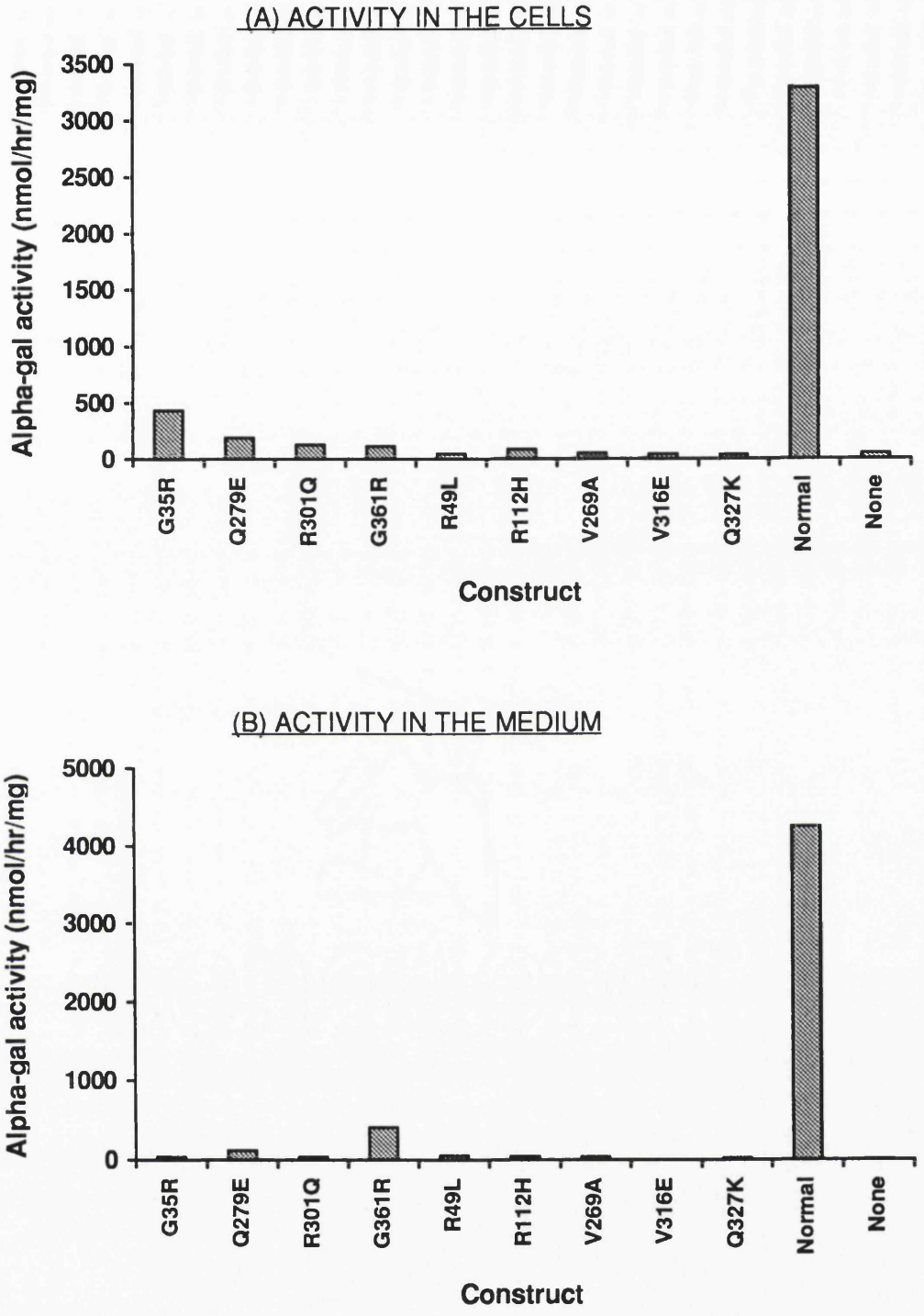


Figure 28: Comparison of expression of normal α-galactosidase A with that produced by mutant constructs, in transfection experiment 1

Histogram of the results from table 19 showing α-galactosidase activities (nmol/hr/mg total protein) produced by the 9 mutant constructs. **(A)** is in the cells and **(B)** is in the medium.

Table 20: Ranked orders of activity produced by mutants with residual activity

α-gal construct	Ranked activity in cells	
	Cells	Medium
G35R	2	4
Q279E	3	3
R301Q	4	5
G361R	5	2
Normal	1	1

It had been observed previously that the efficiency of expression is decreased if many constructs are transfected in one experiment (Dr Y. Iounnou, personal communication). This may be because the numbers of flasks required and the manipulation times are increased, to the detriment of the cells. Therefore, in all other transfection experiments only 5-6 constructs were expressed, to improve the efficiency of transfection.

The four constructs that gave residual activity were expressed in a repeat experiment, transfection 2, table 19(B). In this experiment, the α-galactosidase A inhibitor GalNAc was included in the enzyme assay to give better discrimination between the endogenous activities and the expressed activity. Comparison of the activity produced by the wild-type construct in transfections 1 and 2, in which the total α-galactosidase activity, above the control, was 7469 nmol/mg/hr and 26463 nmol/mg/hr, respectively, indicated that the efficiency of expression or transfection in experiment 2 was greater than in 1. This increase may have been due, in part, to the increased total protein, in the range of 2.2-3.3 mg and 5.4-6.1 mg, in transfections 1 and 2, indicating the presence of a larger number of cells in transfection 2. The absolute values for the specific α-galactosidase activity produced by each construct in these 2 separate experiments could not be compared directly because of the variation in the level of expression between experiments. The activities relative to the control and the percentage of the wild-type activity were also not directly comparable. However, the ranked orders of activities in transfections 1 and 2 were identical, indicating that they could be compared from one experiment to another.

In conclusion, all nine mutant cDNA constructs resulted in decreased expressed activity. Although the absolute α-galactosidase activity values relative to total cell protein varied markedly from one experiment to another, the ranked activities of the constructs were the same.

3.2.2.3 Transfection of the mutant α-galactosidase A cDNA, with cotransfection of β-galactosidase

The α-galactosidase A constructs expressed in section 3.2.2.2 were cotransfected with the plasmid vector pCMVβ (Clontech, figure 9) to investigate whether the coexpressed *E. coli* β-galactosidase would give a better reflection of the transfection efficiency. *E. coli* β-galactosidase, expressed from the cotransfected pCMVβ, is transported to the cytosol, producing intracellular activity. It has an optimal activity at pH 7 and can therefore be distinguished from the endogenous COS-1 cell, lysosomal β-galactosidase, which is optimally active at pH 4.5.

Three cotransfection experiments were carried out, 3, 4 and 5. In transfection 3, the five α-galactosidase A constructs that produced no detectable activity, R49L, R112H, V269A, V316E and Q327K, were cotransfected with 3μg of pCMVβ. In transfections 4 and 5, the four α-galactosidase A constructs that produced detectable activity were cotransfected with 3μg and 1μg of pCMVβ, respectively. The α-galactosidase activity was measured in the presence of the inhibitor, GalNAc (3.2.2.2). *E. coli* β-galactosidase activity was measured as described, 2.3.1.4. The enzyme activities, relative to the total cell protein were calculated as before, in section 3.2.2.2, (table 21(A)). The α-galactosidase A activity was related to the coexpressed β-galactosidase, in the following manner:

$$\text{α-gal/β-gal x100 activity ratio} = \frac{\text{α-galactosidase activity (nmol/hr/ml) x Z x 100}}{\text{Total β-galactosidase activity (nmol/hr)}}$$

where '*Z*' was 1ml and 10ml for expressed intracellular and extracellular α-galactosidase A activity, respectively (table 21(B)).

The α-gal/β-gal ratios were used to calculate the elevation in the expressed α-galactosidase activity, over endogenous COS-1 cell activity and the % of the normal, above endogenous activity, in an analagous manner to those in section 3.2.2.2 (table 22):

$$\text{Increase over control} = \frac{\text{Mutant cDNA activity (α-gal/β-gal ratio)}}{\text{Control (no construct) activity (α-gal/β-gal ratio)}}$$

$$\text{\% of normal} = \frac{\text{Mutant cDNA activity - control activity (α-gal/β-gal ratio)}}{\text{Normal cDNA activity - control activity (α-gal/β-gal ratio)}} \times 100$$

(A) Enzyme activities related to total cell protein

α-gal. Construct	Enzyme activity (nmol/hr/mg total protein)								
	Transfection 3 (3μg β-gal/ 54hr)			Transfection 4 (3μg β-gal/ 72hr)			Transfection 5 (1μg β-gal/ 57hr)		
	β-gal	α-gal		β-gal	α-gal		β-gal	α-gal	
	Cells	Cells	Med-ium	Cells	Cells	Med-ium	Cells	Cells	Med-ium
G35R				431	1883	422	134	989	293
Q279E				357	871	1103	122	789	499
R301Q				351	468	40	129	375	39
G361R				367	413	1517	143	129	602
R49L	727	53	38						
R112H	934	108	55						
V269A	1199	106	81						
V316E	1142	68	31						
Q327K	1076	73	30						
Normal	2214	3055	9075	304	2737	8593	107	3175	12413
None (control)	753	67	24	449	275	7	146	82	15

(B) Alpha-galactosidase activities related to co-expressed β-galactosidase

α-gal. Construct	α-gal/β-gal x100 ratios					
	Transfection 3 (3μg β-gal/ 54hr)		Transfection 4 (3μg β-gal/ 72hr)		Transfection 5 (1μg β-gal/ 57hr)	
	Cells	Medium	Cells	Medium	Cells	Medium
G35R			437	98	740	219
Q279E			244	309	648	410
R301Q			133	11	290	30
G361R			112	413	90	421
R49L	8	5				
R112H	12	6				
V269A	9	7				
V316E	6	3				
Q327K	7	3				
Normal	138	410	900	2825	2979	11646
None (control)	13	3	61	2	57	10

Table 21: Cotransfection of α-galactosidase A and β-galactosidase

10mg of each α-galactosidase A construct transfected into COS-1 cells. The α-galactosidase activity in the cell lysate and medium was determined. In **(A)**, protein was measured for calculation of the specific activities. In **(B)** activities were normalised to the *E. coli* β-galactosidase activity, expressed from the cotransfected pCMVβ.

(A) Transfection of 5 constructs that produced no detectable activity

α-gal. Construct	Transfection 3 (3μg β-gal/ 54hr)			
	Mutant α-gal activity/ control		% of normal activity	
	Cells	Medium	Cells	Medium
R49L	0.6	1.6	-4.0	0.5
R112H	0.9	2.0	0.8	0.7
V269A	0.7	2.3	-3.2	1.0
V316E	0.6	1	-5.6	0
Q327K	0.5	1	-4.8	0
Normal	10.6	136.7	100	100
None (control)	1	1	-	-

(B) Transfection of 4 constructs that produced detectable activity

α-gal. Construct	Mutant α-gal activity/ control				% of normal activity			
	Transfection 4 (3μg β-gal/ 72hr)		Transfection 5 (1μg β-gal/ 57hr)		Transfection 4 (3μg β-gal/ 72hr)		Transfection 5 (1μg β-gal/ 57hr)	
	Cells	Medium	Cells	Medium	Cells	Medium	Cells	Medium
G35R	7.2	49	13.0	21.9	44.8	3.4	23.4	1.8
Q279E	4.0	154.5	11.4	41	21.8	10.9	20.2	3.4
R301Q	2.2	5.5	5.1	3	8.6	0.3	8.0	0.2
G361R	1.8	206.5	1.6	38.5	6.1	14.6	1.1	3.5
Normal	14.8	1412.5	52.3	1164.6	100	100	100	100
None (control)	1	1	1	1	-	-	-	-

Table 22: α-gal/β-gal activity ratios relative to the wild-type and control

The α-gal/β-gal activity ratios from three electroporations in table 21(B) are compared relative to the control (no construct) endogenous activity. They are also compared as a % of the normal, above the control activity.

The constancy of the β-galactosidase expressed in each experiment can be observed in table 19(A). In transfection 3, using 3μg of pCMVβ, the β-galactosidase activities ranged from 727-2214 nmol/hr/mg, a 3 fold difference in the maximum and minimum values. This indicated that the efficiency of transfection varied markedly even in one experiment, despite using the same pool of cells and similar total cell protein, which ranged from 4.4-5.3mg. In experiment 4, again using 3μg of pCMVβ, the expressed activities were at least 1.6 fold lower than in 3 but the β-galactosidase activities were more constant, ranging from 304-449 nmol/hr/mg and with total cell protein in the range of 5.5-6.9mg. The difference in β-galactosidase activity expressed in experiments 3 and 4 did not reflect the expression of wild-type α-galactosidase A, which was similar in both and resulted in total activities of 12039 nmol/hr/mg and 11048 nmol/hr/mg, above endogenous activity, respectively. Thus, measurement of coexpressed β-galactosidase was not a good indicator of α-galactosidase A expression in different experiments. A high level of β-galactosidase activity resulted from transfection with 3μg of pCMVβ. Since it was not known whether cotransfection decreased the efficiency of transfection and expression of α-galactosidase, a decreased amount, 1μg, of pCMVβ was used in transfection 5. The β-galactosidase activity was in the range of 107-146 nmol/hr/mg, a maximum difference of 1.5 fold, with total cell protein of 5.8-6.2mg.

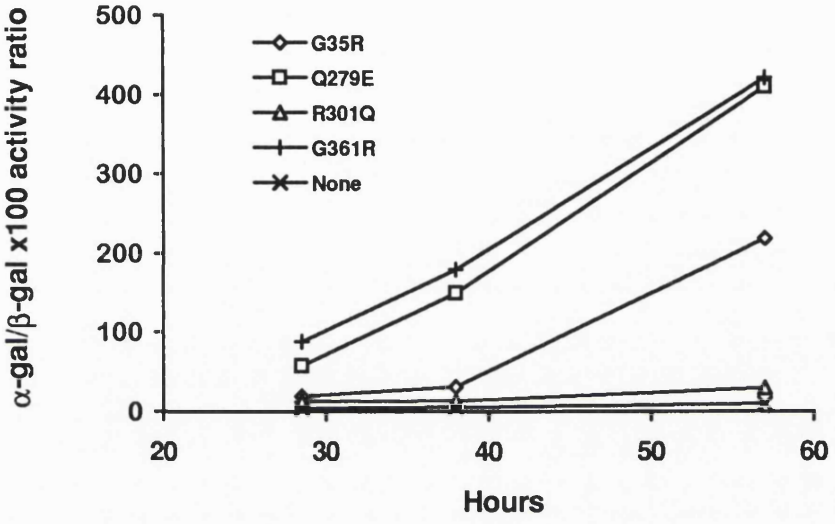
In transfection 3, the mutant constructs, R49L, R112H, V269A, V316E and Q327K, produced no detectable activity. Despite the variability of the β-galactosidase activities in this experiment, calculation of the α-galactosidase A activity, relative to coexpressed β-galactosidase, supported the conclusion that these mutations produced no activity. As a percentage of the normal, above endogenous activity, the α-gal/β-gal ratios of the mutant constructs were not more than 1% (table 22). Coexpression of β-galactosidase provided proof that the lack of activity produced by these α-galactosidase A constructs was not due to failure in the transfection. The α-galactosidase A activities produced in transfection 3, relative to total cell protein (tables 21(A)), can be compared with those in transfection 1 (table 19(A)). The intracellular activities produced by the wild-type construct, above endogenous activity, are similar, 3241 nmol/hr/mg and 2988 nmol/hr/mg in 1 and 3, respectively but the secreted activity in experiment 3 was 2 fold higher. In both, no detectable α-galactosidase activity was produced by the mutant constructs, providing further evidence that these mutations abolish activity.

In transfections 4 and 5, the four mutant constructs, G35R, Q279E, R301Q and G361R, which had produced α-galactosidase A activity on transfection alone (table 19), also produced residual activity when cotransfected with β-galactosidase, which was related to total cell protein (table 21(A)) or coexpressed β-galactosidase (tables 21(B) and 22). For the wild-type α-galactosidase A construct, cotransfection with 3 and 1 μg of pCMVβ, produced total α-galactosidase activity, above the activity from COS-1 cells, of 11048 nmol/hr/mg and 15491 nmol/hr/mg, respectively, a 1.4 fold increase using the smaller amount of pCMVβ (in experiment 5, table 21(A)). In contrast, the four mutant constructs expressed lower activities in transfection 5, using 1μg of pCMVβ. Therefore, it was not clear whether decreasing the amount of cotransfected pCMVβ resulted in an increased or decreased efficiency of α-galactosidase A expression. The α-galactosidase A activities were related to β-galactosidase, (table 21(B)) and the percentage of the normal α-galactosidase activity, above that from COS-1 cells was calculated (table 22). For Q279E, the percentages of the intracellular activity compared with the wild-type construct, in experiments 4 and 5 were similar, 21.8% and 20.2%, respectively. For others, such as G35R, there was a marked difference, 44.8% and 23.4%, in the two transfections. The specific α-galactosidase activities in experiments 4 and 5, related to total cell protein or as a percentage of the normal were not reproducible but the ranked order of the residual activities was the same in these and previous experiments, 1 and 2 (as in table 20).

In conclusion, the variation in expression of the α-galactosidase A constructs and the β-galactosidase, precludes direct comparison of specific activities from one experiment to another. However, the relative activities of the different constructs within an experiment were consistent and clearly indicated which mutations gave rise to null or residual activity alleles.

As the lack of reproducibility between different experiments could be due, in part, to a non-linear increase in expressed activity with time, the expressed activity was measured as a function of time (table 23 and figure 29).

(A) α-galactosidase A activity secreted into the medium from COS-1 cells transfected with residual activity mutants



(B) α-galactosidase A activity secreted into the medium from COS-1 cells transfected with the wild-type α-gal A cDNA construct

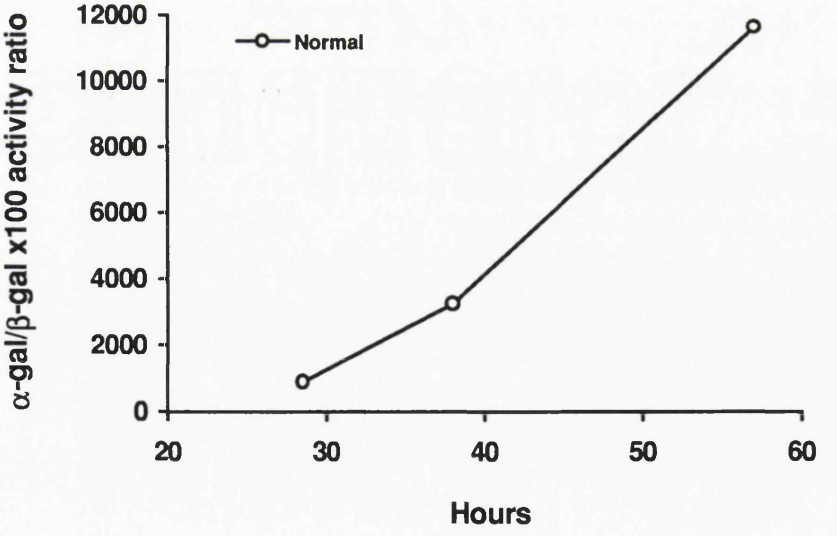


Figure 29: Increase in α-galactosidase A activity with time in the medium from COS cells transfected with residual enzyme activity mutants

Alpha-galactosidase A activity (α-gal/β-gal ratio x100 values, in table 23) from COS-1 cells electroporated with mutant cDNA constructs **(A)** and with wild-type cDNA **(B)**.

Table 23: Alpha-galactosidase in the medium relative to intracellular β-galactosidase after cell lysis at 57hr

Alpha-galactosidase A activities in the medium at 28.5, 38 and 57hr are related to coexpressed, intracellular β-galactosidase activity at 57hr.

α-gal A Construct in pMT2	α-gal/β-gal x100 activity ratio, in the medium at:		
	28.5hr	38hr	57hr
G35R	19	31	219
Q279E	58	150	410
R301Q	13	14	30
G361R	88	180	421
Normal	892	3262	11646
None (control)	4	6	10

The experiments confirmed that the ranked activities could be reliably compared and that all mutations decreased the activity of α-galactosidase A, supporting the hypothesis that they were disease-causing mutations.

3.2.2.4 Correlation between expressed activities and the patient phenotypes

The relative expressed α-galactosidase A activities produced by the constructs were compared with the clinical phenotypes observed in patients with those mutations (table 24).

Three of the five mutations producing no detectable intracellular activity during transient expression, R49L, V269A and V316E had been found in patients known to have typically severe phenotypes. The phenotype associated with a fourth, with no detectable activity, Q327K, was unknown (section 2.1.1). The α-galactosidase A activity in cultured fibroblasts or lymphoblastoid cells from patients with these four mutations was less than 1% of that found in controls (table 24). Therefore the *in vitro* expression studies reflect the situation *in vivo*. The fifth mutation, R112H, producing no detectable activity, was found in a patient with mild proteinuria as the only symptom but whose sibling was severely affected (Eng *et al.*, 1994b). No further details regarding the phenotype associated with R112H and no enzyme activity data were available for comparison with the results obtained in this study.

Mutation	α-galactosidase A activity				Patient phenotype
	Transient transfection in COS-1 cells		In patient cells		
	% of normal, range of values (tables 19 & 22)		sample type	% of normal	
	cells	medium			
R49L	-4.0 -> 0.1	0.5 -> 1.0	Lymphoblastoid cells	0.8	typical
R112H	0.8 -> 1.3	0.7 -> 0.9	-	-	variant
V269A	-3.2 -> 0.1	1.7 -> 0.6	Fibroblasts	0.8	typical
V316E	-5.6-> -0.2	-0.3 -> 0	Lymphoblastoid cells	0.6	typical
Q327K	-4.8-> 0.2	0 -> 0.3	Fibroblasts	0.4	unknown
G35R	11.9 -> 44.8	0.6 -> 3.4	Fibroblasts Leukocytes	15 (a) 5 (a)	Severe variant
Q279E	4.4 -> 21.8	2.6 ->10.9	Lymphoblastoid cells Plasma	3.8 (b) 3.0 (b)	Mild variant
R301Q	2.5 -> 8.6	0.2 -> 0.3	Lymphoblastoid cells Plasma	3.6 (b) 1.6 (b)	Mild variant
G361R	2.1 -> 6.1	3.5 -> 14.6	-	-	Severe, typical

Table 24: Comparison of the phenotypes with the expressed activities

Data were obtained from transfection experiments in this study, in which the α-galactosidase activities were related to the total cell protein (table 19) or β-galactosidase activity (table 22) and are presented as a percentage of the normal activity, above the endogenous COS-1 cell activity. Alpha-galactosidase A activity data in material from Fabry patients were obtained in this study (table 18), except for (a), information communicated by Dr H. Christomanou and (b), Ishii, *et al.*, 1992.

Four mutations, G35R, Q279E, R301Q and G361R, produced residual enzyme activity in the expression studies. Two of these, Q279E and R301Q, were found in patients with late onset and mild symptoms (cardiomyopathy only), who were diagnosed at ages 63 and 52, respectively (Nagao *et al.*, 1991; Ishii *et al.*, 1992). Comparison of the relative activities *in vivo* with the transient expression data (table 24) shows that the proportion of residual intracellular α-galactosidase A activity in the expression studies is similar or higher than those in the cultured cells and plasma. The ranked activities produced by transient expression of these two mutations were the same as their *in vivo* ranked activities.

The G35R mutation was found in a patient with severe neurological involvement at age 26 but no other symptoms. The finding of transiently expressed residual activity of 11.9-44.8% (table 24) is consistent with residual activities of 5% and 15% of the normal in leukocytes and cultured fibroblasts, respectively, from the hemizygote with G35R (H. Christomanou, personal communication). Again, the proportion of residual activity produced by overexpression is similar or higher than *in vivo*. A severe phenotype in this patient, in contrast to the milder phenotype in patients with the Q279E and R301Q mutations, which give lower expressed activities, may be due to mislocalisation of the expressed protein or decreased activity towards the natural substrate. It is also possible that there is a second unidentified mutation that lowers the activity further *in vivo*.

Expression of the G361R mutation produced activity in the medium but almost zero intracellular activity (table 24). It was associated with a severe, typical phenotype. No data on the enzyme *in vivo* was available for comparison.

In conclusion, all of the mutations cause a decrease in the expressed α-galactosidase activity, relative to the wild-type. The ranked order of residual intracellular activity expressed by the mutations and the phenotypes is consistent with higher residual activity producing a milder phenotype, with the exception of G35R.

3.2.2.5 Comparison of the activity produced by this expression system and by others

The activities produced by Q279E, R301Q and the normal, in the pMT2 vector used in this study, were compared with the activities produced in a previous expression study, in which the mutants were expressed in COS-1 cells, using the pBactE

vector and transfection by the calcium phosphate/glycerol shock technique (Ishii *et al.*, 1992). The pMT2 vector was chosen for the present studies because it was believed to produce a high level of expression in COS cells, although it had not been used previously for expression of α-galactosidase A.

Transfection experiments and expression of normal α-galactosidase A cDNA constructs in this study resulted in intracellular activity, above the control (no α-galactosidase A construct) in the range of 2462-14455 nmol/hr/mg total cell protein, in cells lysed at 47-72hr after electroporation (tables 19 and 22). In comparison, using the pBactE expression system, the expressed activity from the wild-type cDNA, in two experiments, was 2075 nmol/mg/hr and 419 nmol/mg/hr at 6 days after transfection (Ishii *et al.*, 1992), using a similar enzyme assay method. Therefore, the pBactE vector and transfection with calcium phosphate are less efficient for expression of α-galactosidase A than the pMT2 expression vector and transfection by electroporation.

Two mutated α-galactosidase A cDNAs have also been expressed in the two systems:

Table 25: Comparison of the intracellular activities produced by expression of cDNA with Q279E and R301Q in this study and a previous study

Mutant	Intracellular activities above control (nmol/hr/mg protein)		
	Using pMT2 and electroporation, this study (transfections 1, 2, 4 & 5, tables 19, 21)		Using pBactE vector and Ca-phosphate/glycerol transfection (Ishii, *et al*, 1992)
	Range	Mean	Mean (estimated from a histogram)
Q279E	144-1953	850	106
R301Q	82-864	358	75

In accordance with higher activities for expression of the wild-type α-galactosidase A, the mean activities produced by expression of Q279E and R301Q mutant constructs in this study were higher than those from pBactE, with transfection by the calcium phosphate/glycerol method technique (Ishii *et al.*, 1992). The Q279E mutant construct produced more activity than R301Q in both expression systems.

The normal α-galactosidase A cDNA has also been expressed in COS-1 cells, using the vector p91023(b) and polybrene transfection (Bishop *et el.*, 1991). The intracellular α-galactosidase activity increased to 500 nmol/hr/mg, above the control, a 4 fold increase after 72 hours, as estimated from published data. In the present study, intracellular activity was in the range of 2462-14455 nmol/hr/mg, 11-89 fold higher than in the control after 447-72 hours (tables 19 and 22).

Thus, transient expression of the α-galactosidase A cDNA in COS-1 cells, using pMT2 and electroporation is more efficient than systems used previously (Bishop *et al.*, 1991 and Ishii *et al.*, 1992).

3.2.2.6 Detection of the transiently expressed mutant proteins by immunoprecipitation and SDS-PAGE

The intracellular and extracellular fractions from COS-1 cells transfected with α-galactosidase A mutants were analysed for protein cross-reacting with anti-(α-galactosidase A) serum. The cells were metabolically labelled with ^{35}S-methionine to facilitate detection of immunoprecipitated protein. The experiments were designed to see if any enzymically inactive or structurally altered protein was produced by the mutant cDNA constructs.

All 9 mutant α-galactosidase A constructs and the wild type were transfected into COS-1 cells by electroporation, without cotransfection of the β-galactosidase construct, since this would compete with the α-galactosidase construct for the ^{35}S-methionine label. The mutant constructs that gave no enzyme activity, R49L, R112H, V269A, V316E and Q327K, were expressed in one experiment and the G35R, Q279E, R301Q and G361R mutants, which produce activity, were expressed in a second experiment. The enzyme activity in the medium was assayed for those constructs that were known to produce enzyme activity, to verify the success of the transfection (table 26). At 50-60 hours after transfection the cells were metabolically labelled with ^{35}S-L-methionine for 6 hours before immunoprecipitation with anti-(α-galactosidase A) serum and SDS-PAGE analysis.

(A) Transfection with mutant cDNAs that did not produce residual activity

α-gal A Construct in pMT2	α-gal activity (nmol/hr/ml) in the medium at:		
	23hr	32hr	57hr
R49L	6	5	12
R112H	8	12	25
V269A	10	19	58
V316E	6	6	10
Q327K	5	6	7
Normal	190	980	5439
None (control)	5	7	9

(B) Transfection with mutant cDNAs that produced residual activity

α-gal A Construct in pMT2	α-gal activity (nmol/hr/ml) in the medium at:		
	18hr	32hr	55hr
G35R	10	32	446
Q279E	24	147	606
R301Q	8	17	25
G361R	45	232	449
Normal	124	970	3862
None (control)	8	14	15

Table 26: Alpha-galactosidase A activity in the medium from cells transfected for immunoprecipitation of α-galactosidase A

10μg of each α-galactosidase A construct was electroporated into COS-1 cells. At various time points, medium was withdrawn and the α-galactosidase A activity was measured, with GalNAc in the assay (2.3.1.3). Cells were washed at 57hr and 55hr (tables **(A)** and **(B)** respectively) and grown in methionine-free medium with [^{35}S]-L-methionine for 6hr before harvesting and immunoprecipitation.

Immunoprecipitation and SDS-PAGE analysis of the normal α-galactosidase A (figure 30) showed that the major intracellular form of the protein had a relative molecular size of approximately 50kD and that two additional, faint bands, one slightly smaller and the other larger than the major band were also precipitated. The secreted protein had a slightly higher molecular weight than the major intracellular form and gave only one band.

Immunoprecipitation and SDS-PAGE analysis of the intracellular α-galactosidase A (Figure 31, (A, B)) showed that the endogenous COS-1 cell protein was detected (control, lane 'C'). It appeared to be slightly smaller in size than the protein produced by the construct with the normal human cDNA sequence (lane 'N'). Alpha-galactosidase from COS-1 cells may form the faint band that migrated slightly faster than the major band of immunoprecipitated material from each transfected cDNA construct. The intensities of the major band produced by each mutant, the normal and control, in each experiment, were compared by eye and a darker band was assumed to indicate a higher amount of immunoprecipitated α-galactosidase A. The amount of the major band produced by the wild-type (lane N) and each mutant (lanes 1-9) was higher than in the control. This indicated that all of the transfected mutant constructs produced intracellular protein that cross-reacted with α-galactosidase A antibody, even though R49L, R112H, V269A, V316E, Q327K and G361R produced very little enzyme activity (table 24). No alterations in the size of the detectable mutant proteins were observed, suggesting that truncated or incorrectly processed proteins were not produced.

In contrast to intracellular proteins, secreted cross-reacting proteins (figure 31 (C, D)) were not always detected and the amount of endogenous COS-1 cell protein that was secreted was only just detectable. No immunoprecipitated protein was detected in medium from cells transfected with the R49L and V316E mutants. This could be because they were not secreted or they were secreted but could not bind to antibody. A lack of secretion could occur if the protein is not transported correctly after translation or if the amount of protein produced is too small. All of the residual activity mutants produced cross-reacting protein in the medium, except for R301Q, for which the cross-reacting material was not significantly higher than the endogenous COS-1 cell protein. The low enzyme activity in the medium produced by R301Q correlated with its lack of CRIM. No alterations in the size of the detectable mutant proteins were seen.

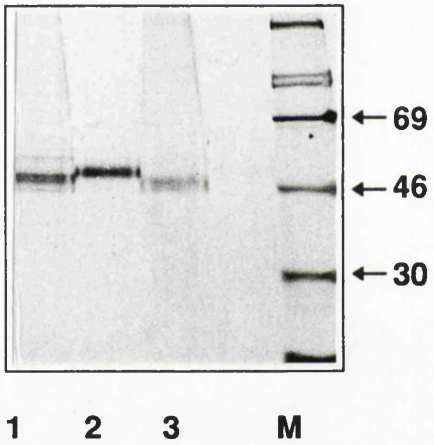


Figure 30: Comparison of the secreted and intracellular forms of α-galactosidase A by SDS-PAGE

Immunoprecipitated intracellular protein (lanes 1, 3) and secreted protein (lane 2) from COS cells electroporated with normal α-galactosidase A in pMT2. Approximately equal quantities were analysed by SDS-polyacrylamide (10%) gel electrophoresis (2.3.6.2.3) for size comparison. M is the set of [^{14}C]-methylated rainbow marker proteins (Amersham), with sizes indicated in kiloDaltons.

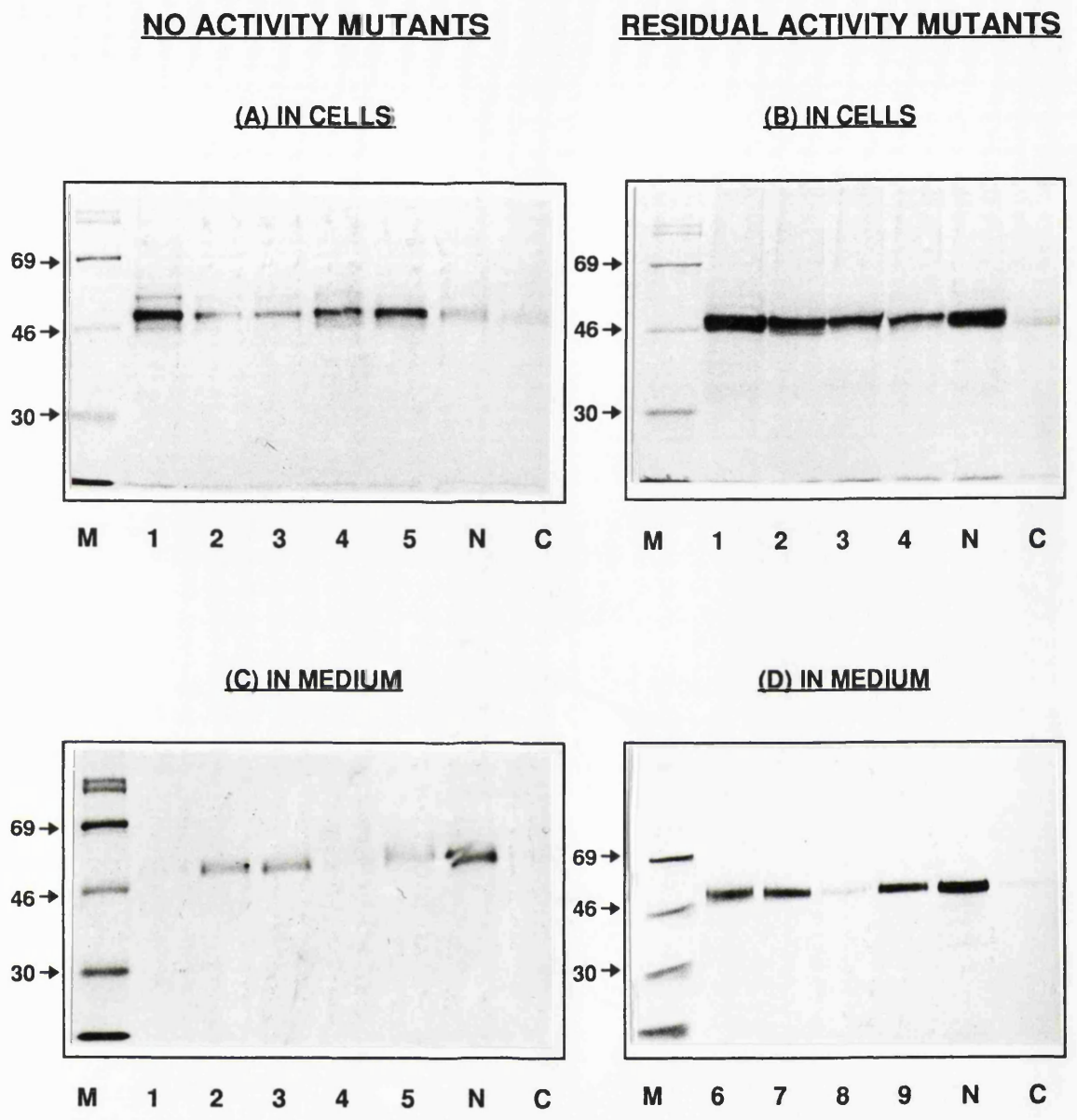


Figure 31: Immunoprecipitation and SDS-PAGE analysis of protein produced by transient expression

Transfected COS cells were grown for 55-57hrs before growth for 6 hours in methionine-free medium with ^{35}S-L-methionine. IgG-purified polyclonal antiserum was used to immunoprecipitate the α-galactosidase A protein from the medium and the cells and the protein was analysed on a 10% polyacrylamide-SDS gel (2.3.6.3).

M is the [^{14}C] methylated rainbow marker (Amersham) and protein sizes are indicated in kiloDaltons, N is electroporated with the wild-type cDNA construct and C is the no construct control. 1, 2, 3, 4, 5, 6, 7, 8 and 9 were electroporated with cDNA containing the R49L, R112H, V269A, V316E, Q327K, G35R, Q279E, R301Q and G361R mutations, respectively.

The results show that cross-reacting material was detected intracellularly and extracellularly in all the mutants that produced residual α-galactosidase activity. Catabolically inactive enzymic protein of the correct size was detected in some of the mutants that produced no detectable α-galactosidase activity, suggesting that these mutations affect the catalytic properties or stability of the enzyme rather than its synthesis.

4. DISCUSSION

4.1 THE EFFICIENCY OF SCREENING FOR MUTATIONS BY SSCP ANALYSIS

The usefulness of SSCP analysis for mutation screening was investigated in this study. SSCP analysis detects mutations that cause a shift in the migration of single-stranded DNA through a non-denaturing polyacrylamide gel and with optimal conditions it can give a detection efficiency of greater than 90% (Orita *et al.*, 1989; Hayashi, 1991). The factors that affect the sensitivity and/ or the types of mutation detected by SSCP analysis have been investigated by others and include varying the voltage for electrophoresis, the gel temperature, the buffer conditions, the polyacrylamide and glycerol gel content and the length of DNA analysed (Spinardi *et al.*, 1991; Fan *et al.*, 1993; Sheffield *et al.*, 1993). The use of a stacking gel, prior to the band resolving gel, to aid denaturation and to increase the proportion of single-stranded product has also been suggested to improve sensitivity (Yap *et al.*, 1993). These studies have shown that electrophoresis at room temperature, using a low voltage and in a non-denaturing gel which contains 5% glycerol and 6% acrylamide gives a high detection efficiency and so these conditions were used as the standard method in this study. Use of a stacking gel to improve band resolution was not considered because it complicates the process of gel preparation.

Another consideration was the size of DNA to be analysed. In one previous study the optimal DNA length was shown to be 155 bases, giving a 97% detection efficiency by analysis of 29 mutations (Sheffield *et al.*, 1993). Analysis of the same mutations in a smaller fragment, 135bp gave a detection efficiency of 90%, while increasing the size to 175, 212, 420 and 600bp decreased the detection to 76%, 70%, 58% and 3%, respectively. However, Fan *et al.* (1993) found a greater than 90% efficiency for mutation detection in a 354bp fragment and the same mutations were detected with equal efficiency in a 176bp fragment, showing no effect of sequence length. Also, some mutations in the 354bp fragment produced a larger change in mobility than in the 176bp fragment, indicating that a longer fragment did not always decrease the ability to distinguish and detect mutations. The α-galactosidase A gene contains seven exons ranging from 92-291 bases. Analysis of each exon individually, by PCR amplification, with about 100-150 bases of flanking

intron sequence, gives product sizes of 184, 274, 286, 350, 357, 400 and 441 bases for exons 4, 2, 5, 1, 6, 3 and 7, respectively. This is an acceptable size range for reliable mutation detection (Fan *et al.*, 1993).

In this study, a single set of conditions for SSCP analysis was used to screen for sequence changes in the protein-coding region and splice sites of the α-galactosidase A gene. Screening in normal individuals identified 6 different polymorphic sequence changes. Thirty unrelated Fabry families were analysed and non-polymorphic, variant band patterns were detected in the majority, 26, families. Sequencing showed that 24 were different mutations and two occurred twice. SSCP analysis did not detect sequence changes in 4 families. In at least 2 cases we have shown that the mutation did not cause a conformational change detectable under the conditions used but in the remaining 2, as no mutations were found by sequencing the whole region, they may be in another part of the gene.

The 4 families in which no mutation was found were analysed by sequencing the entire protein-coding region and splice sites to determine whether the mutations were in the region screened by SSCP analysis. Two new mutations, R49S and D92H, were found by sequencing. The location of R49S and D92H, which were undetected by the standard SSCP conditions, did not follow any obvious pattern and could not have been predicted, in agreement with results from systematic mutation studies in other genes (Fan *et al.*, 1993; Sheffield *et al.*, 1993). Other sequence changes in exon 1, such as 1P1 and 1P2, which were separated by only one base, were clearly distinguished by SSCP analysis. In contrast, the R49L allele, also in exon 1, was detected while R49S, in an adjacent nucleotide, was not. Perhaps the presence of the 1P3 polymorphism, which cosegregates with the R49L mutation, contributed to the altered conformation in exon 1, allowing it to be distinguished from both the 1P3 and the normal patterns. Studies by others, (Sheffield *et al.*, 1993; Hayashi *et al.*, 1991), indicated that mutation detection in long DNA fragments is less efficient than in shorter ones. In this study, R49S in exon 1 and D92H in exon 2, were located in DNA fragments of 350bp and 274bp, respectively, suggesting that size is not the only factor determining the success of SSCP analysis. However, shortening the 274bp PCR product containing D92H, by restriction digestion with *Hinf* I to form a 168bp fragment, permitted its detection. In two other families, no mutations were found by complete sequencing. In two other families, no mutations were found after complete sequencing. The failure to detect a mutation in these two families is apparently because the protein-coding region and splice sites did not contain any mutations.

One further consideration is that although mutations were detected by SSCP analysis in 26 families, it is not known whether more than one mutation was present in each patient. In one study, a Fabry patient was shown to have 2 sequence changes, one of these was a mutation but the other did not alter an amino acid and expression studies showed that it did not lower the α-galactosidase A activity (Koide *et al.*, 1990). No other Fabry patients have been shown to have more than one mutation and in at least 57 cases, complete sequencing of the cDNA or the protein-coding regions and splice sites has demonstrated that there is only one mutation (Koide *et al.*, 1990; Sakaruba *et al.*, 1990, 1992; Ishii *et al.*, 1990, 1991; Ploos van Amstel *et al.*, 1994; von Scheidt et al, 1991; Eng *et al.*, 1993, 1994b; Fukuhara *et al.*, 1990; Yokoi *et al.*, 1991. Thus, it is likely that SSCP analysis detected the only mutation in these families.

The ability of SSCP analysis to detect mutations in the majority of cases indicates that this is a good strategy for mutation screening in the alpha-galactosidase A gene. The simplicity of the technique and the ease with which altered band patterns are observed are a significant advantage over direct sequencing. However, in the few cases in which sequence changes are not detected, alternative strategies, such as complete sequencing, are required. If mutations are not found by these methods then they may be present in other regions of the gene. SSCP analysis is extremely useful for detecting known mutations in heterozygotes, due to the simplicity of the method and interpretation.

4.2 POLYMORPHISMS IN THE ALPHA-GALACTOSIDASE A GENE

4.2.1 The types and distribution of polymorphisms

Six polymorphisms have been found in this study, all of which are in regions of the gene that do not code for protein. Five are base substitutions and it is notable that three of these reside in the 60bp 5' untranslated region of exon 1, at -10, -12 and -30 of the cDNA sequence. The three polymorphic SSCP analysis variants that were found to contain these sequence changes have been found at a cumulative frequency of 10%. Such a high level of polymorphism in the 5' untranslated mRNA sequence has not been reported for any other genes, except for the FMR-1 gene, in which the variable length of a tandem repeat is highly polymorphic (Fu *et al.*, 1991). In comparison, polymorphic variation is thought to occur in non-coding regions such as introns with a frequency of about 1 in 100 bases and since the 3' untranslated region of many genes is often several hundred bases long, it is not surprising that this non-coding region is also frequently polymorphic (Levitt, 1991; Poduslo *et al.*, 1991; Avramopoulos *et al.*, 1993). The 5' untranslated region is shorter, usually less than 100bp (Lewin, 1994) and includes the consensus Kozak sequence, which has been shown to be important for ribosome initiation of translation (Kozak, 1986, 1987). Two polymorphisms at -10 and -12 (1P2 and 1P1 respectively) are close to the Kozak sequence, as indicated below:

Kozak consensus sequence:	**GCCACCAUGG**
α-galactosidase A sequence:	-12 -GTCACC GUGACAAUGC- +4
Polymorphisms:	A I

Bold type indicates the bases in the consensus sequence that are conserved in the α-galactosidase A mRNA. The polymorphic bases and AUG initiation site for translation are underlined. The identification of polymorphisms in this region is an indication that alterations in this sequence do not effect the production of α-galactosidase A sufficiently to cause disease. Their effect, if any, on the efficiency of translation was not investigated. It was not known whether the three polymorphisms in exon 1 cosegregated as they were detected by SSCP analysis, not sequencing. Since there were only three different band patterns, the absence of additional patterns in exon 1 suggested that 1P1, 1P2 and 1P3 did not cosegregate.

The sixth polymorphism is a 5bp deletion in which the common sequence 5'-cccccagcc-3' is mutated at undefined break-points to give 5'-cccc-3'. This mutation

may have been produced by slipped mispairing during replication of the repeated cytosine nucleotides and this mechanism of mutagenesis would predict the possibility of several sizes of deletion/ insertion at this sequence. However, sequence analysis of the 3P1 polymorphism in 4 normal, unrelated males and one Fabry hemizygote showed the same deletion and is evidence that it represents a single mutation event, rather than several independent mutations.

This brings the total number of known polymorphisms in the gene to eight. The exact sequence changes for the two mutations found by others (Desnick *et al.*, 1987; Eng *et al.*, 1994a), a *Sac* I detectable polymorphism in intron 4 and an *Nco* I polymorphism at about 10kb downstream from the last exon, have not been published. No polymorphisms have been found in the protein-coding region of the gene.

4.2.2 Detection of polymorphisms for the identification of heterozygotes

The inheritance of Fabry disease can be followed by analysis of polymorphisms which segregate with the defect in the α-galactosidase A gene. The six intragenic polymorphisms identified in this study may be useful for carrier detection. In the 60 males that were completely analysed and in which the phase of the polymorphisms was known, polymorphic patterns were observed in 17 PCR products but only 10% of chromosomes contained one or more polymorphisms. Six males were polymorphic, 2 with the 1P2-3P1-5P1-7P1 haplotype, 2 with 3P1-5P1-7P1, one with 1P1-7P1 and the other with 1P3. This was an indication that the polymorphisms were in disequilibrium. If they had segregated independently then the frequency of males expected to have one or more variant patterns would be (17/60) x 100 = 28% (table 11, 3.1.1.2). Therefore, a strategy in which all 6 polymorphisms were analysed would not be more informative for carrier detection than detection of a few polymorphisms as they often cosegregated.

In the 19 females that were analysed completely, at least 7 different haplotypes were observed in 9 females having one or more rare alleles. Although the phase of the polymorphisms was unknown, their presence was consistent with the haplotypes observed in males. For example, the three females that were heterozygous for 3P1, 5P1 and 7P1 are likely to have inherited one non-polymorphic chromosome and one with all three polymorphisms, a haplotype observed in 2 males.

A reliable estimate of the number of females that are heterozygous for one or more polymorphisms cannot be made from the data in this study. In 60 males, the frequency of each haplotype is low and so larger sample numbers are required to use this information to estimate the number of informative females. Observation of the numbers of polymorphic females, in a sample of 19, is also an unreliable indicator of the frequency of females that carry polymorphisms, due to the small sample size. In this study, analysis of 19 females showed that 47% were heterozygous for one or more polymorphisms. This level of polymorphism was much higher than the 10% observed in 60 hemizygotes and although unrelated females were randomly selected, the small group may be unrepresentative of those in the general population.

The cosegregation of polymorphisms should be considered when designing strategies for identifying carriers, by polymorphism detection. Since the 3P1, 5P1 and 7P1 variants were frequently found as one haplotype (4/6 chromosomes in males, table 11, 3.1.1.2), it is only useful to screen for one of these polymorphisms. Analysis of the 3P1 variant and simultaneous analysis of the three exon 1 polymorphisms would identify all of the polymorphic chromosomes found in the males studied and their detection would be useful for observing patterns of inheritance. However, this strategy would not have detected one female, with polymorphisms 5P1 and 7P1. The development of an SSCP detection method using ethidium bromide staining and a small, simple gel system will facilitate screening for these four polymorphisms in Fabry patients, although it has not been tested in Fabry families. Analysis by radiolabelled SSCP detection in 30 Fabry families found polymorphisms that segregated with the disease phenotype in 4 cases, 13%, a similar degree of polymorphism compared with analysis in the 60 non-Fabry males (11%).

It is concluded that carrier detection by polymorphism analysis could be usefully carried out by using a simple SSCP analysis method to detect polymorphisms in exons 1 and 3. However, the majority of Fabry hemizygotes would be non-polymorphic and so other genetic methods would be required for carrier detection in their families.

4.3 MUTATIONS IN THE ALPHA-GALACTOSIDASE A GENE

In this study, putative disease-causing mutations have been identified in 28 families and 26 of these were different mutations. This brings the total number of Fabry families with mutations reported to 106. Mutations in 93 families are different (table 4, figure 32 and table 27), 7 of which are recurrent in a further 13 families (table 28). Of the 93 different mutations, 64 are point mutations which alter codon recognition and 5 are splice-site mutations, which include four single base substitutions and a 2bp deletion. Six are large rearrangements and 18 are small (<100bp) rearrangements, including 13 deletions, 3 insertions and 2 complex mutations (table 27). The mutations that were found in more than one unrelated family are shown in table 28. No mutations have been reported in the transcription promoter region of the gene. This information is useful for identifying regions of the gene that are more prone to mutation, allowing the development of efficient screening strategies for analysis in new families to improve carrier detection and patient counselling. Investigation of the biochemical and clinical effects of these mutations may provide insight into the aetiology of the disease and identify those individual most likely to benefit from enzyme therapy in the future.

4.3.1 Analysis of families in which no mutation was found

In families 29 and 30, SSCP analysis did not detect any sequence alterations in the splice sites and exons of the α-galactosidase a gene and all exons were amplified successfully. Mutations were not identified by sequencing the same region. It is possible that the mutations were in the region sequenced but were misinterpreted as a gel artifact, since not all parts of the sequence were clearly readable. Therefore, an alternative method such as chemical mismatch cleavage, which detects close to 100% of mutations, could be used as a second screening procedure to check that the mutations were not in the protein-coding regions or splice sites.

Further analysis would be used to identify mutations in other regions of the gene. Reverse-transcription of RNA and sequencing may identify mutations in the introns that alter splicing. Mutations may also reside in the transcription promoter and enhancer sequences and could be identified by sequencing this region of the gene.

Schematic Diagram of the α-galactosidase A gene

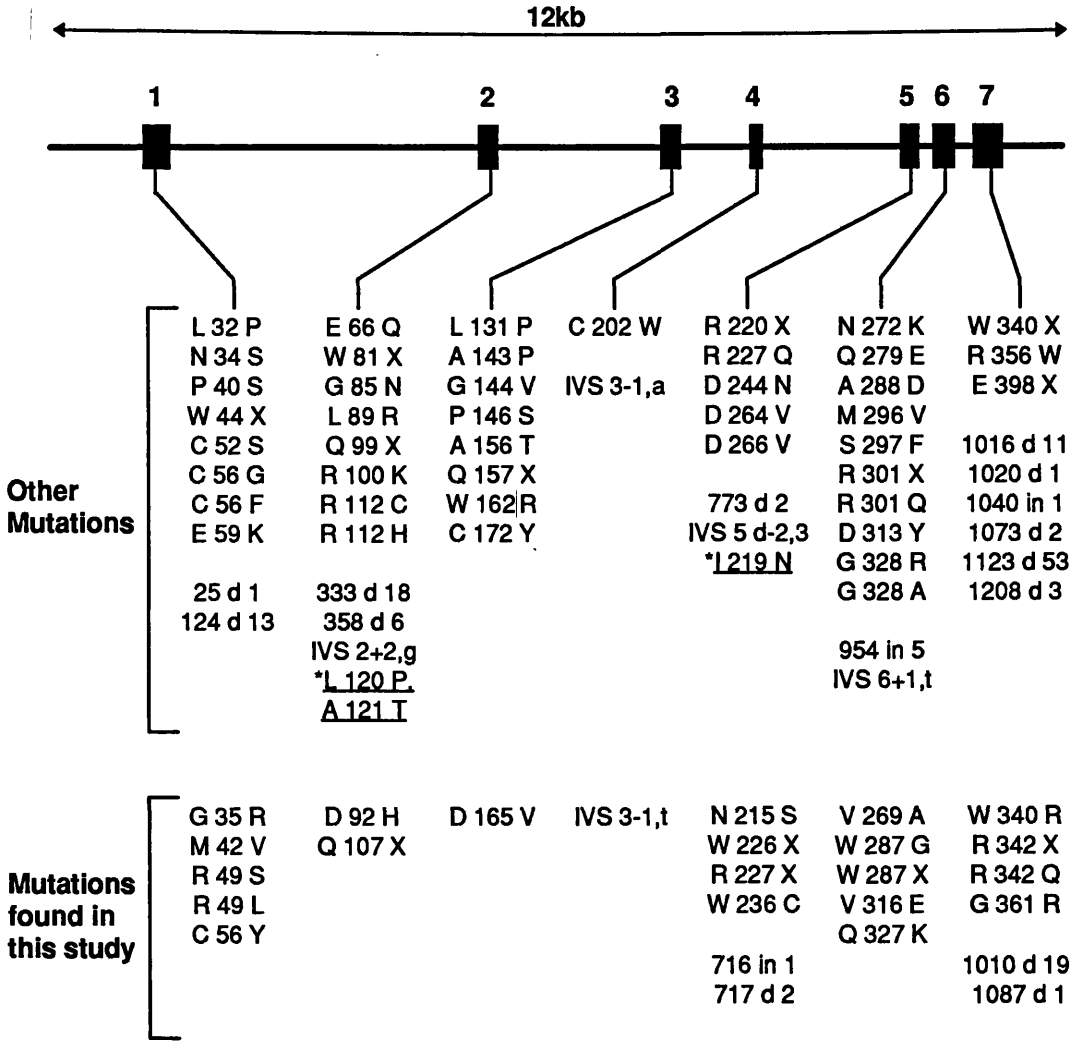


Figure 32: Locations of point mutations, splice site mutations and small rearrangements in the α-galactosidase A gene

Allele designations are as in table 4. 'd' and 'in' indicate deletions and insertions, respectively. Splice site mutations are indicated by IVS and complex mutations are indicated by * and underlining. Exons are indicated above the schematic diagram of the gene.

93 different putative disease-causing mutations		Numbers of mutations	Total
Point mutations in the protein-coding region	Transitions (non-CpG) Transitions at a CpG Transversions (non-CpG) Transversions at a CpG	30 10 21 3	64
Point mutations at splice sites	Transitions (non-CpG) Transversions	3 1	4
Small Rearrangements (1-53bp) in the protein-coding region	Deletions Duplications Complex	13 3 2	18
Small rearrangements at splice sites	Deletions	1	1
Large rearrangements (402-8112bp)	Deletions Duplications Complex	4 1 1	6

Table 27: The numbers and types of different putative disease-causing mutations found in the α-galactosidase A gene

Mutation type	Recurrent Mutations	Unrelated families with the mutation (in 106 families)	
		Number	Nationality (ethnic group)
Point mutations at CpG sites	R227X R227Q R301Q R342Q	3 2 2 4	2 British (1 Asian), 1 non-British 2 German Japanese, Danish 3 Dutch, 1 British
Point mutations at non-CpG sites	N34S N215S	3 4	2 Danish, Polish British, German, Italian, Czechoslovakian
Small deletions	717del2	2	Belgian, British

Table 28: Recurrent alleles and their frequency in 106 unrelated families

4.3.2 Evidence that the family-specific mutations cause Fabry disease

Several lines of evidence support the hypothesis that disease-causing mutations have been identified in this study. Analysis of the protein-coding region and splice sites did not find more than one mutation in each patient and all of these mutations were predicted to alter the amino-acid sequence of the protein. Further evidence was provided by *in vitro* expression studies, for 6 of the missense mutations.

4.3.2.1 Evidence from genetic analysis

The mutations detected in Fabry patients were believed to be disease-causing for the following reasons. Firstly, all but two mutations, R49S and D92H, were not found in DNA from at least 100 normal chromosomes. R49S and D92H were not detected by the standard method for SSCP analysis and normal chromosomes were not screened for these mutations. Secondly, no other SSCP changes were seen in the remaining 6 exons, other than polymorphisms, and finally, the exon containing the unique change was fully sequenced and did not contain any other mutations. However, not all regions of the gene were analysed and because we and others (Orita *et al.*, 1989; Hayashi, 1991, Spinardi *et al.*, 1991; Fan *et al.*, 1993; Sheffield *et al.*, 1993) have shown that SSCP is not 100% efficient, it is possible that an undetected, second mutation is responsible for the Fabry phenotype.

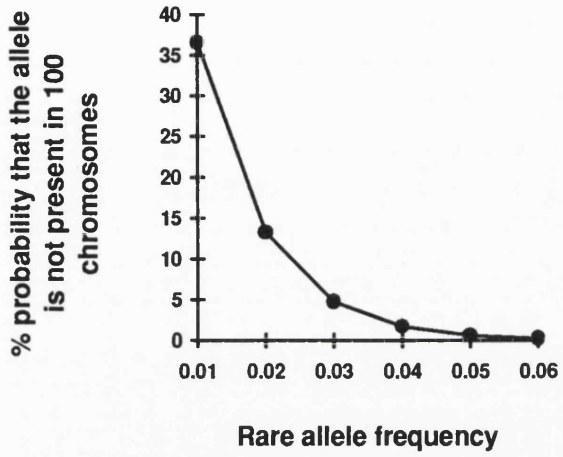
The sequence change may be a rare polymorphic allele that is present in the normal population but is not in the 100 chromosomes analysed in this study. If a two allele system is considered, then the probability that an allele is not present in a sample population is given by:

$$\text{Probability} = (1 - \text{allele frequency})^n,$$

$$\text{or,} \quad \log(\text{probability}) = n \log(1 - \text{allele frequency})$$

where 'n' is the number of chromosomes analysed. Therefore, if a rare polymorphism has a frequency of 1 allele in 100 in a sample size of n=100, there is a 37% chance that the allele will not be detected. By the same calculation for the same sample size, an allele with a frequency of 0.03 or greater will have a probability of less than 5% that it is not present. This relationship is illustrated in

the graph below, which shows the probability that an allele of a certain frequency is not present in a sample of 100 chromosomes:



Further support for the disease-causing status of these mutations comes from the fact that SSCP analysis detected a rare sequence change in 26/30 (87%) of the Fabry patients and that no other variants were detected in the protein-coding region. If it is assumed that 87% of all mutations are detected then the probability that the mutation is a rare polymorphism (with a frequency of 0.01) and that a mutation was not detected by SSCP analysis is calculated by:

$$\text{probability} = [(1-0.01)^{100}] \times (1-0.87) = 0.05$$

Therefore, there is a 95% probability that each mutation detected by SSCP analysis is not present in the normal population with a frequency of 1/100 and that it is the only mutation in the protein-coding or splice sites of the α-galactosidase A gene. For 8 mutations, M42V, R49S, C56Y, D92H, W287G, R342Q, 717del2 and 1087del1, sequencing proved conclusively that these mutations were the only ones present in the protein-coding region and splice sites.

Further evidence for the disease-causing effect of mutations comes from finding recurrent or *de novo* mutations. The N215S, R227X, R342Q and 717del2 mutations have been found in more than one unrelated family, showing segregation with the disease phenotype and absence in the normal population. The C56Y mutation was present in a clinically typical Fabry hemizygote but was

absent in lymphocyte DNA from his mother. Since no other affected family members were known, this *de novo* mutation was believed to cause Fabry disease.

Other evidence comes from the fact that analysis of the protein-coding region and splice sites by us and others (table 4, pg 56) has identified more than one sequence change in only 1/106 Fabry patients, indicating that in the majority of cases, only one mutation is present in this region. In this patient, expression studies showed that the disease-causing mutation was P40S, while a silent base change at codon 8, did not reduce the α-galactosidase activity (Koide *et al.*, 1990). Complete sequencing of either the cDNA or the protein-coding regions and splice sites of the gene has provided evidence that in at least 57 patients there were no other mutations in this region of the α-galactosidase A gene.

4.3.2.2 Biochemical evidence for the disease-causing nature of mutations

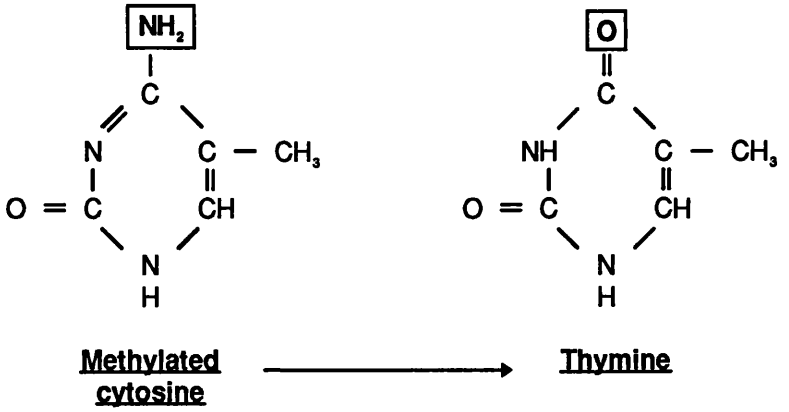
All of the mutations detected in Fabry patients were predicted to alter the protein sequence indicating that they are likely to cause disease (tables 13 and 14, section 3.1.2.1). In addition, the G35R, R49L, V269A, V316E, Q327K and G361R mutations found in this study were transiently expressed in COS-1 cells. All of them caused a decrease in α-galactosidase A activity, compared with the wild-type cDNA, supporting the hypothesis that they were disease-causing. R49L, V269A, V316E and Q327K produced no detectable activity within the cells or secreted into the medium. G361R did not produce intracellular activity but did secrete catalytically active protein into the medium. These results indicated that there would be no activity in the lysosomes and would therefore cause disease. However, the G35R mutation did produce low, intracellular activity and in four transfection experiments it was in the range of 12-45% of the normal (tables 19 and 22(B), pgs 161 and 166). The presence of residual activity was in agreement with finding 15% and 5% of the normal activity in cultured fibroblasts and leukocytes, respectively, from the hemizygote in family 1 (Dr H. Christomanou, personal communication).

In conclusion, the above lines of evidence support the hypothesis that all 26 mutations identified in this study are causative for Fabry disease.

4.3.3 Point mutations in the α-galactosidase A gene

4.3.3.1 The distribution and frequency of point mutations

From observation of the location of the 64 point mutations in the protein-coding region of the gene it can be concluded that their distribution is not random, because of the higher incidence of mutations at the CpG dinucleotide sites (table 27). Thirteen mutations, 21%, are at CpG sites, which account for only 3% (38/1290 x100) of the total number of bases in this region. The higher mutability at these sites is presumed to be due to methylation of cytosine and the ability of this base to undergo spontaneous deamination to thymine, resulting in a C to T transition (Coulondre *et al.*, 1978; Cooper *et al.*, 1988), as indicated in the diagram below:



This mechanism is supported by the observation that 10/13 of the point mutations at CpG sites are transitions of this type. Further evidence for the increased mutability at CpG sites comes from the finding of recurrent mutations (table 28) at some of these sites. The R227X allele was found in 3/106 unrelated families (2 are reported by Eng *et al.*, 1993 and 2 were found in this study, Davies, JP *et al.*, 1993 but one of the families was analysed and reported by both groups) and the R227Q allele was found in 2 families. Screening of 148 families (Eng *et al.*, 1993) specifically for these two alleles (including some of the 106 families which have been fully analysed) found the frequency of the R227X and R227Q alleles to be 5/148 and 3/148, giving a combined frequency of 5.4% of all mutations causing Fabry disease. Codon 342 is also mutated in several families and 4/106 have the R342Q mutation while 1/106 has the R342X mutation. Codon 301 is mutated in 3 families, one has the R301X mutation and two others have the R301Q mutation.

The observation that no mutations were found at 5 CpG sites at the 5' end of the α-galactosidase A cDNA sequence lends support to the hypothesis that the higher mutability at CpG sites is due to the presence of methylated cytosine bases. The 5' end of the region analysed forms part of the CpG-rich island, which is usually unmethylated in constitutively expressed genes and is therefore less mutatable. A reduced mutability of CpG dinucleotides at CpG-rich regions has been observed in other genes, notably the α-globin gene in which the CpG dinucleotides are not a hotspot for mutation, presumably due to a lack of methylation (Yoshida, 1983; Antonarakis *et al.*, 1985).

CpG-rich islands are often found at the 5' end of transcriptionally active genes and they have a G/C content above 50% and the frequency of CG sites is equal to the number of GC sites (Bird, 1986). If it is assumed that the arrangement of bases is random and that DNA has a content of 25%C:25%G, then the chance of finding a CpG site is 1 site in 16bp (1/4 x 1/4). The 5' region of the α-galactosidase A gene sequence from bases 500-1250 has 1 CpG every 18 bases, on average and this is indicative of an unmethylated CpG island. This region includes the protein-coding sequence from nucleotides 1-70 (the first 24 amino acids of the signal peptide) and contains five CpG sites. Therefore, it is likely that only 14 of the 19 CpG sites in the protein-coding region are methylated and are potential mutation 'hotspots'. The 1220bp region from cDNA nucleotides +71-1290, outside the CpG-rich island, contains 28 bases, 2%, in CpG dinucleotides. The increase in the number of putative disease-causing C to T or G to A transition mutations at CpG sites, compared with those at non-CpG sequences in this region is calculated by:

$$\text{Mutation frequency} = \frac{\text{Number of transitions (table 27)}}{\text{number of base pairs in region}}$$

$$\text{mutation frequency at CpGs / mutation frequency at non-CpGs} = (10/28\text{bp}) / (30/1192\text{bp})$$

= 14 fold elevation of transitions at CpGs

This increased incidence of transition mutations at CpG sites is lower than that observed in the factor IX in which the ratio of transitions at CpG to non-CpG sites is 24:1. However, the above calculation does not take into account the additional, recurrent mutations at CpG sites.

The increased mutability at the CpG sites accounts partly for the higher numbers of transition mutations compared with transversions in the protein-coding region. If mutations at CpG sites are not included then the difference in the numbers of transitions and transversions is less marked with 59% (30/51) and 41% (21/51), respectively. This agrees with data from mutation analysis in the factor IX gene in which the numbers of transitions at non-CpG sites is elevated compared with transversions but is less marked than at the CpG sites (Yoshida, 1983; Bottema *et al.*, 1991).

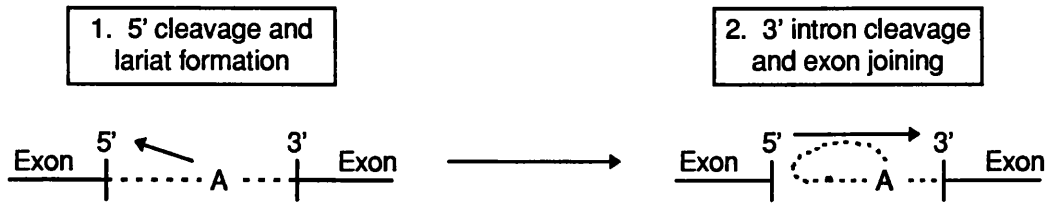
Interestingly, two of the recurrent point mutations were not at CpG sites, N34S and N215S. The reason for the presence of the N34S allele in three families (Eng *et al.*, 1994a; Madsen *et al.*, 1995) is not obvious as it is not at a CpG site and does not cause a mild phenotype. However, the presence of the N215S mutant in 4 families may be related to the mild phenotype with which it is associated, allowing it to persist in the population for longer than mutations giving a severe phenotype. These families may be distantly related or they may represent independent mutational events that have not been eliminated from the population.

It is noteworthy that no point mutations of any type have been found in the first 90bp and that only one is found in the last approximately 200 bases of protein-coding sequence. Perhaps there is a decreased susceptibility of the DNA sequence to mutation, other than the known influence of the CpG dinucleotides in the first 70bp. Alternatively, it may that the phenotype is unrecognised, has no symptoms or it lethal in utero. It is possible that the 31 amino acid signal peptide, encoded by the 5' sequence, is more tolerant of mutations, since it is cleaved and does not participate in the catalytic action of the protein. The observation that only two mutations have been reported to occur in the signal peptide of any lysosomal enzymes, a point mutation, P5R, in the fucosidase gene of a patient with fucosidosis (H. Cragg Ph.D thesis, 1995) and a 1bp cytosine base deletion in the iduronate 2-sulphatase gene of a patient with mild Hunter disease, lends support to this idea. With regard to the lack of point mutations in the last 200 bases of the cDNA, encoded by exon 7, it may be significant that this region has the least homology with N-acetylgalactosaminidase, 15.8%, while the remaining protein-coding region has 56.4% identity (Wang, AM *et al.*, 1990, 1991). Exon 7 may be more tolerant to mutations. In support of this is the identification of a 3bp deletion in exon 7, which causes a mild phenotype in one Fabry patient (Eng et al., 1994b).

4.3.3.2 The effects of splice site mutations

In this study, one splice-site mutation was found, an ag->at substitution at the 3' end of intron 3, IVS3-1,t. It was found in a patient with a severe phenotype (table 29). The mutation was predicted to cause mis-splicing of exon 4 because this sequence was part of a highly conserved consensus sequence (Shapiro *et al.*, 1987). Evidence supporting this prediction comes from knowledge of the splicing reaction and the DNA sequences that are involved.

Eukaryotic genes have a protein-coding sequence which is split (Breathnach *et al.*, 1981) into exons, and these are separated by non-coding intron sequences. Since both the introns and exons are transcribed to produce hnRNA, the production of mature RNA for translation into protein requires the removal of the introns. Normally, splicing involves a two-step transesterification reaction, which is dependent on the presence of three consensus intronic sequences. Introns are joined and exons are removed by cleavage of the 5' intron junction at a conserved sequence, formation of a lariat structure, cleavage at the 3' intron conserved sequence and joining of two adjacent exons (Nilsen, 1994; Sharp, 1994), as indicated below:



The two intron sequences at the 5' and 3' splice site junctions contain the almost invariant GT and AG dinucleotide sequences but a few exceptions exist (Jackson, 1991). The invariant residues are indicated in bold type in the following consensus sequences:

```
Position:        -3   ->    +6                -14     ->      +1
sequence:   EXON--CAG | gtaagt ----intron---- (c or t)10 xcag | G--
EXON
```

The vertical lines indicate the intron-exon boundary, between the +1 and -1 base positions, 'x' indicates any base. Mutations in many genes, at the conserved dinucleotides have been reported to cause disease (Krawczak *et al.*, 1992) and they result in either exon skipping or use of a cryptic splice site in the intron or the exon.

Mutation type	Allele	Phenotype (family No.)	Age	Onset age and clinical symptoms									
				1	2	3	4	5	6	7	8	9	Other
Missense	**G35R**	**variant (1)**	**26**	**-**	**-**	**N**	**-**	**-**	**N**	**N**	**N**	**Y**	**-**
	M42V	typical (2)	30	7	Y	Y	-	N	-	N	-	N	-
	R49S	typical (3)	d44	2	Y	Y	-	-	Y	-	Y	Y	-
	R49L	typical (4)	33	-	Y	Y	Y	-	-	N	N	-	-
	C56Y	typical (5)	49	15	Y	Y	-	Y	Y	-	Y	N	-
	D92H	typical (6)	37	5	Y		-	Y	N	-	N	N	-
	D165V	unknown (7)	No hemizygote information										
	N215S	**mild?? (8)**	**No hemizygote information**										
	W236C	typical (9)	No hemizygote information										
	V269A	typical (10)	39	-	-	Y	-	-	-	-	Y	Y	-
	W287G	typical (11)	15	8	Y	Y	-	Y	N	-	N	N	-
	V316E	typical (12)	27	8	Y	Y	Y	-	N	N	-	-	-
	Q327K	unknown (13)	No hemizygote information										
	W340R	typical (14)	No hemizygote information										
	R342Q	typical (15)	No hemizygote information										
	G361R	typical (16)	-	8	Y	Y	-	Y	N	-	N	N	-
Nonsense	Q107X	typical (17)	31	-	-	-	-	-	-	-	Y	-	-
	W226X	typical (18)	24	18	Y	N	Y	-	N	Y	N	-	Skeletal
	R227X	typical (19)	29	5	Y	Y	-	Y	N	-	N	N	-
	R227X	typical (20)	38	9	Y	Y	Y	-	-	Y	N	-	
	W287X	typical (21)	54	9	Y	Y	-	-	Y	N	Y	N	Pedal oedema
	R342X	unknown (22)	No hemizygote information										
Small	716ins1	typical (23)	11	-	Y	Y	N	-	-	-	-	N	-
deletions/	717del2	typical (24)	38	6	Y	Y	N	N	Y	Y	Y	N	oedema
insertions	717del2	unknown (25)	No hemizygote information										
	1010del19	typical (26)	33	22	Y	Y	Y	Y	N	N	Y	N	Pedal oedema
	1087del1	typical (27)	37	5	Y	Y	Y	Y	N	Y	N	Y	-
Splice site mutation	IVS3-1, t	typical (28)	d45	32	Y	Y	Y	-	Y	Y	Y	N	

Table 29: The clinical phenotypes of hemizygotes in this study

'Y' and 'N' indicate the presence and absence of symptoms and '-' is unknown information. 'd' is the age of death, and phenotypic variants are in bold type. (A) and (B) are hemizygous relatives.

The column numbers refer to: **1.** onset age; **2.** acroparesthesia and pain crises; **3.** angiokeratoma; **4.** eye involvement; **5.** hypohidrosis; **6.** proteinuria; **7.** cardiac involvement; **8.** kidney involvement; and **8.** neurological involvement, excluding pain.

Although no RNA analyses or studies of the protein size were conducted to determine the effect of the mutation, it is hypothesised that IVS3-1,t will result in skipping of exon 4. This is supported by the identification of another mutation, IVS3-1,a, at the same base, an ag->aa alteration, which was analysed by RT-PCR and shown to produce a mRNA that lacked exon 4 (Yokoi *et al.*, 1991). In addition, the absence of an alternative 3' intron splice-site consensus sequence close to or within exon 4, suggests that exon-skipping would probably occur. Since the exon 3/4 and 4/5 boundaries are not in frame, a frameshift would result in exon 5 to create a stop signal at codon 219. Consequently, the protein may be truncated. The protein may also be decreased in amount, either due to a destabilisation of the abnormal mRNA, (Urlaub *et al.*, 1989; Peltz *et al.*, 1993) or because the protein is misfolded and subjected to proteolysis (Kassenbrock *et al.*, 1988; Rothman, 1989; Klausner *et al.*, 1990). Therefore, although there is no direct evidence for the effect of this mutation, the prediction of a severe effect was in accordance with the severe clinical phenotype and finding almost zero α-galactosidase A activity in cultured fibroblasts from this patient (table 18, pg 155).

In other studies, four splice site mutations were found in Fabry patients with severe phenotypes, in accordance with their predicted, severe effects. However, not all mutations at the invariant splicing dinucleotides cause abnormal splicing. For example, a point mutation in the 'invariant' GT (5') splice sites of intron 2 in the adenine deaminase gene produced some RNA that was normally spliced and the clinical phenotypes of two affected siblings were variant, with one mildly and the other severely affected (Arredondo-Vega *et al.*, 1987).

4.3.3.3 The effects of nonsense mutations

Clinical information was obtained for 5 out of the 6 hemizygotes in families 17-22, in which a nonsense mutation was found (table 29 and section 2.1.1). Four different mutations, Q107X, W226X, R227X and W287X were present in patients with the typical symptoms and R227X was found in two unrelated patients. The phenotype of the sixth hemizygote, with R342X, was unknown. Four of these families were of interest.

In family 18, the hemizygote with W226X had the typical symptoms and also had the less common symptom of retarded growth and late onset of puberty, with

skeletal dysplasia. A related hemizygote had no skeletal dysplasia but did have a diffuse lack of calcium.

The R227X mutation was detected in two unrelated families, one of which, family 19, was Asian and the other was British family 20, which had been extensively studied and could be traced back over 6 generations. In family 20 (Hamers *et al.*, 1979; Johnston *et al.*, 1966) the symptoms were typical but there was evidence of phenotypic variation in 11 hemizygous relatives and 13 heterozygotes and the severity did not correlate with the urinary excretion of glycosphingolipids. At the two extremes of severity, one 32 year old hemizygote, V-13 (figure 6, section 2.1.1), was severely affected with pains, angiokeratoma, slowed growth, flexed fingers and oedema while another, III-38 was less severely affected, with almost no angiokeratoma. He had suffered from two cerebrovascular thromboses at ages 48 and 55 but recovered and was still able to work at age 59. The blood groups of all but three hemizygotes were known to be A_1, A_2 or O and so the phenotypic variability was not related to increased substrate load caused by the presence of the blood group B antigens.

A third mutation, W287X was of interest because although the range of symptoms was typical, they were slightly less severe than in other patients with nonsense mutations. The hemizygote with this mutation was not diagnosed biochemically until age 54 and was still fit enough to work at this time.

The specific effects of these nonsense mutations have not been investigated experimentally, except in family 20 in which two hemizygotes were known to be CRIM negative to α-galactosidase A antiserum (Hamers *et al.*, 1979). Evidence for the predicted effects of these nonsense mutations comes from investigations in other genes. They may have one or more of the following consequences: (a) they produce a truncated protein, as predicted by the genetic code; (b) they cause skipping of the mutant exon (Fukada, I *et al.*, 1992; Naylor *et al.*, 1992; Dietz *et al.*, 1993); (c) they cause a reduction in the amount of mRNA. It has been shown that in the dihydrofolate reductase gene stop mutations reduce the mRNA to 20% of the normal amount (Urlaub *et al.*, 1989). Patients with Sandhoff disease, caused by inheritance of stop mutations on one or both chromosomes, produced mRNA in the range of 0.8-30% of the normal amount (Zhang, Z-X *et al.*, 1994); or (d) the codon is read through, producing a normally sized protein, as seen in the aspartylglucosaminidase gene (Peltola *et al.*, 1994) and the fucosidase gene (Yang *et al.*, 1992, 1993). Exon skipping and a reduction in the amount of mRNA are

thought to involve a specific exon-scanning mechanism to remove mRNA containing a premature stop codon (Peltz *et al.*, 1993; Dietz *et al.*, 1994).

The R342X mutation cannot cause exon skipping because it is in the last exon and may not cause a reduction in the mRNA. Evidence for this comes from analysis of the triosephosphate isomerase gene in which stop mutations at the end of the penultimate or in the last exon do not cause a decrease in the amount of mRNA, while more proximal mutations do (Cheng *et al.*, 1990). However, exon skipping and a reduction in mRNA may be caused by the Q107X, W226X, R227X and W287X mutations, the former three of which would produce frameshifts in the sequences of exons 2, 5 and 5 respectively, as their exon boundaries are not in-frame. Skipping of exon 6 in the W287X would not produce a frameshift but would produce a truncated protein. These consequences are consistent with observed classic phenotypes and the finding of almost zero enzyme activity, less than 1% in cultured fibroblasts from patients with W226X and R342X (table 18, pg 155). No RNA analyses or studies of the protein size were carried out to determine the precise effects of these five mutations.

4.3.3.4 The effects of missense point mutations

Of the 16 missense mutations found in this study , 9 were found in hemizygotes with the classic symptoms, G35R was associated with a variant phenotype and for 6, no information was available from hemizygotes (table 29 and 2.1.1). For these missense mutations it is predicted from the genetic code that the size of the protein is not altered unless they affect the mRNA splicing. None of the mutations here is predicted to create a consensus splice site (Shapiro *et al.*, 1987). However, altered splicing has been shown to occur for missense and silent mutations that do not create a consensus splice site. For example, several missense and silent mutations have caused abnormal splicing in the hypoxanthine-guanine phosphoribosyltransferase gene (Steingrimsdottir *et al.*, 1992), all of which are located close to the correct splice site. Most notable of these is the silent base substitution, located 13bp upstream of the natural 5' intron splice site, causing approximately 90% of the transcript to lack the mutant exon. No mRNA analysis was performed to determine the effect of the missense mutations in this study.

4.3.3.4.1 Missense mutations that give a typical phenotype

Thirteen of the missense mutations were classified as giving a severe, typical phenotype and a further one, D165V was thought to be typical but there were no clinical details (table 29 and 2.1.1). For two of these mutations, W236C and W340R, clinical information was only available from symptomatic heterozygotes. The only clinical information regarding another symptomatic heterozygote who had the Q327K mutation was that she had angiokeratoma. The R342Q mutation was found in a family for which limited clinical information was available but the phenotypic characteristics in symptomatic females were indicative of a typical phenotype. Thus, all of the missense mutations were found in patients that were classified as typical, except for G35R and N215S.

The phenotypes of hemizygotes with D92H, in family 6 were of interest because although the ranges of symptoms were typical of Fabry patients, they were described as slightly milder and clinically variant, by two clinicians. Two affected male cousins, II-4 and II-7 were identical ages and one had hypohidrosis but no renal manifestations or pedal oedema while the other did not have hypohidrosis but had a low creatinine clearance and pedal oedema. Another, II-6, continued to work at age 39, with no obvious symptoms except for pains and no further information on this patient was available. In addition, all females were asymptomatic. Further clinical information is required to determine whether the phenotype is truly atypical in this family. It is interesting to note that the R112H mutation, which is close in the linear sequence to D92H, is found in one 56 year old patient with proteinuria as the only symptom, while a relative was severely affected and died of kidney failure (Eng *et al.*, 1994b).

Biochemical studies have been carried out to determine the effects of some of these mutations. The α-galactosidase A activities in lymphoblastoid cells from patients with R49L and V316E were negligible, as were the activities produced by V269A and Q327K in cultured fibroblasts (table 18, pg 155). This is consistent with the severe phenotype observed in patients with these mutations. Expression of cDNAs containing these same four mutations in COS-1 cells also gave negligible activity and is evidence that these mutations are sufficient to cause the Fabry phenotype. A fifth mutant that was expressed, G361R, was found to produce low residual activity in the medium, although the intracellular activity was negligible, correlating with the severe phenotype.

Material cross-reacting with anti-(human α-galactosidase A) serum was found intracellularly in COS-1 cells in which cDNA constructs containing the mutations R49L, V269A, V316E, Q327K and G361R had been transiently expressed. The immunoprecipitated protein had the same size as the wild-type protein, suggesting that the mutations did not alter splicing of the RNA. However, the possibility that abnormally-sized protein was produced but could not bind to the anti serum, could not be ruled out. In one of the two expression experiments, in which the normal and R49L, V269A, V316E, Q327K and R112H mutant constructs were expressed and subjected to immunoprecipitation and SDS-PAGE analysis, faint bands slightly larger and smaller than the major protein band were observed intracellularly. The larger of these may be a precursor of the mature α-galactosidase A enzyme and the smaller band may be partially degraded protein or the endogenous COS-1 cell enzyme, which appears to be slightly smaller than the over-expressed α-galactosidase A. As the additional bands were present in the immunoprecipitated, expressed wild-type protein, they were not a consequence of the presence of the mutations. Evidence for different, intracellular, post-translational processing pathways comes from observation that the secreted protein was larger in size than the major, intracellular form. A larger secreted form of the wild-type α-galactosidase A has been detected in other cell types types (Lemansky *et al.*, 1987; Iounnou *et al.*, 1992). The processing intermediates could have been studied further by 'pulse-chase' experiments in which the cells are grown in the presence of label for a short time before removing the label and immunoprecipitating the protein after several different periods of time have elapsed. This type of experiment may identify differences in processing of the mutant proteins, compared with the normal.

The effects of the other mutations have not been investigated and the influence of these mutations on the protein structure, function, transport and processing is unknown. Alpha-galactosidase A protein has been crystallised (Murali *et al.*, 1994) but the location of the individual amino acid side chains within this structure has not been determined. The folded structures of other lysosomal glycosidases are unknown and so possible regions of conserved protein topology that may be affected by mutations are unknown. Some mutations may affect the protein by substituting a small group with an obviously more bulky amino acid side chain, as with C56Y, D92H and G361R. Others may alter the bonds between adjacent side chains. The importance of hydrophobic interactions is indicated by the effect of the V269A mutation in which substitution of a -CH $(CH_3)_2$ side chain for a CH_3 is sufficient to completely abolish the enzyme activity. Covalent interactions, such as

the formation of disulphide bridges between cysteine residues are also important for protein stability (Price *et al.*, 1988). So, the removal of a cysteine by the C56Y mutation and the insertion of one by W236C could disrupt the conformation of the protein and alter its stability and susceptibility to proteolysis (Kassenbrock *et al.*, 1988; Rothman, 1989; Klausner *et al.*, 1990). The relative importance of these effects requires more knowledge of the protein structure. The following mutations were at amino acids that were conserved in the α-N-acetylgalactosaminidase protein sequence (Wang, AM *et al.*, 1990): M42V, R49L, R49S, C56Y, D92H, D165V, W236C, W287G, Q327K and R342Q. This is an indication that changes at these amino acid positions might affect the active conformation of the enzyme. Four other mutations, V269A, V316E, W340R and G361R were at non-conserved amino acid residues.

4.3.3.4.2 Missense mutations that give a variant phenotype

Two of the mutations found in this study, G35R and N215S were present in patients classified as phenotypic variants (tables 29 and 30). The G35R mutation was found in a 26 year old patient with reversible hemiparesis and recurrent fevers, indicating neurological involvement but with no other clinical characteristics of Fabry disease. Both of his parents, at the ages of 48 and 54 years, were clinically normal but no material from the parents was available for genetic analysis. No reliable clinical information could be obtained for the hemizygote with the N215S mutation in this study. However, in three other unrelated families, N215S was shown to be associated with a variable, mild phenotype. One of these, a 42 year old Italian hemizygote, had proteinuria and rheumatoid arthritis only and there were no cardiac or other manifestations (Bishop *et al.*, 1981b). In contrast, a 63 year old Czechoslovakian (Elleder *et al.*, 1990a) and a 63 year old German (Eng *et al.*, 1993), with N215S, both had cardiac involvement only.

Studies of the activity of the G35R mutant in cultured fibroblasts and leukocytes (Dr. H. Christomanou, personal communication) and of the N215S mutant in a lymphoblastoid cells in this study (table 18, pg 155) and by others (Bishop *et al.*, 1981b; Eng *et al.*, 1993), have shown residual, intracellular enzyme activity. In this study, the transiently expressed G35R mutation produced intracellular activity, which cross-reacted with the α-galactosidase A antiserum (3.2.2.6). The effects of these mutations on the protein structure are unknown.

Mutant Allele	Phenotype	Biochemical analysis of the protein					References
		in vivo studies	*In vitro* studies during transient expression in COS celis				
		Residual enzyme activity	Residual enzyme activity, using		CRIM +ve*		
			pBactE vector	pMT2* vector	I	S	
G35R*	severe variant	Y	-	*Y (I)	*Y	*Y	1
N215S*	mild variant	Y*	-	-	-	-	2,3
Q279E	mild variant	Y	Y (I)	*Y (IS)	*Y	*Y	4,5
R301Q	mild variant	Y	Y (I)	*Y (I)	*Y	*Y/N	4,6,7
R112H	mild-severe variant	-	-	*N	*Y	*Y	8
P146S	Mild	Y	-	-	-	-	9
M296V	mild	Y	-	-	-	-	10

Table 30: Comparison of the genotype, clinical phenotype and protein studies in atypical hemizygotes

* indicates results from this study; I is intracellular and S is secreted protein. Y is yes and N is negligible activity or CRIM, while Y/N indicates that there is a small amount of CRIM and '-' indicates that no result was obtained. The references are as follows. 1 is a personal communication from Dr H. Christomanou. 2-10 are Eng *et al.*, 1993; Eng *et al.*, 1994a; Ishii *et al.*, 1992; Ishii *et al.*, 1993; Sakuraba *et al.*, 1990; Madsen *et al.*, 1995; Eng *et al.*, 1994b; Ploos van Amstel *et al.*, 1994; and von Scheidt *et al.*, 1991, respectively.

However, the N215S is predicted to remove an asparagine that is normally glycosylated (Desnick *et al.*, 1994) and the mild effect suggests that the lack of glycosylation at position 215 does not completely abolish the catalytic activity. In one study (Desnick *et al.*, 1994), transient expression of α-galactosidase A cDNA containing an N215Q mutation at this site, produced intracellular and secreted enzyme activity that was 47% and 3.4% of the normal, respectively, in COS-1 cells. Transient expression of the same amino acid substitution at the two other glycosylation sites, N139 and N192, produced intracellular activity that was 90% and 95% of the normal, respectively. This indicates the importance of glycosylation at asparagine 215. The presumed position of the mutant amino acid is at the protein surface and its mutation may cause less disruption of the enzyme structure than those affecting the protein core (Matthews, 1993), allowing activity and milder clinical symptoms. In another study (Bishop *et al.*, 1981b), the residual α-galactosidase A activity in fibroblasts from a patient with N215S was more thermolabile at pH 4.6 and 50°C and more unstable at pH 7.4 and 37°C, than the normal enzyme. The instability at pH 7.4 was in accordance with finding almost zero α-galactosidase A activity in the plasma and urine, while the lack of storage material observed in a liver biopsy could be explained by the presence of intracellular enzymic activity.

Studies by others have identified five additional missense mutations that result in variant phenotypes (section 1.3.2.1, table 3 and table 30). Three of these, Q279E (Nagao *et al.*, 1991; Ishii *et al.*, 1992), R301Q (Sakuraba *et al.*, 1990; Nagao *et al.*, 1991; Ishii *et al.*, 1992; Madsen *et al.*, 1995) and R112H (Eng *et al.*, 1994b), have been transiently expressed in COS-1 cells in this study. The Q279E and R301Q produced residual intracellular enzyme activity. This is consistent with finding intracellular activity *in vivo* (Ishii *et al.*, 1992). Previous analysis of the purified Q279E mutant protein, produced by transient expression in insect cells and COS-1 cells, has shown it to be a stability mutant (Ishii *et al.*, 1993). The fourth mutation, R112H did not produce any detectable activity on expression. Information regarding the activity in the patient cells was not available for comparison.

Comparison of the intracellular enzyme activities expressed in COS-1 cells with the phenotype, does not always give the expected results. A higher concentration of intracellular activity in one mutant, compared with another may not indicate a higher capacity to degrade the lysosomal substrate and a consequent milder phenotype. This is because the activity may be located in the incorrect intracellular compartment or the artificial substrate used in the enzyme assay does

not give a true reflection of the activity towards the natural substrates *in vivo*. Alternatively, the amount and catalytic activity of mutant enzyme produced *in vivo* may depend on the cell type. This may explain why, in this study, *in vitro* expression of the G35R mutant produces a higher, intracellular activity than the Q279E and R301Q mutants but it gives a more severe clinical phenotype. The presence of activity towards the natural substrate but not the artificial one used for assays in this study, may explain why the R112H mutation gives a mild phenotype but no detectable activity by transient expression. Alternatively, other genetic or environmental factors may moderate the effect of the R112H mutation, which would also explain the intrafamilial variation that is associated with R112H (section 4.3.2.4.1).

Comparison of the clinical effects and the distribution of the mutations along the gene sequence shows no obvious pattern in the location of mutations giving a variant phenotype. Mutations associated with variants were located in exons 1, 2, 3, 5 and 6. Even two mutations that were at the same codon had different effects. The R112H mutant gave a variant phenotype while the R112C mutant gave the typical symptoms. The mutations in hemizygotes with a mild phenotype were also close to those with a severe phenotype, most notably, the M296V mutant was mild while the adjacent S297F mutant was severe. Therefore it is concluded that the effects of amino acid substitutions on the structure and function of α-galactosidase A are subtle and will require a detailed 3-dimensional picture of the enzyme or enzyme-substrate complex for understanding. All of the mutations found in patients with a variant phenotype are at amino acids in α-galactosidase A that are conserved in the α-N-acetylgalactosaminidase protein sequence (Wang, AM *et al.*, 1990).

Comparison of the types of amino acid substitutions that give a variant phenotype, P146S, G35R, R112H, N215S, Q279E, M296V, R301Q or residual enzyme activity and a severe phenotype, G361R, with those found in this study that produce a typical clinical effect and negligible enzyme activity, is shown in figure 33. Six out of eight mutations associated with a variant phenotype and/or residual activity, were at polar residues and were replaced with alternative polar residues. This supports the hypothesis that hydrophobic residues in the protein core are important for protein stability and that hydrophilic amino acids on the protein surface are more tolerant to mutation (Matthews, 1993).

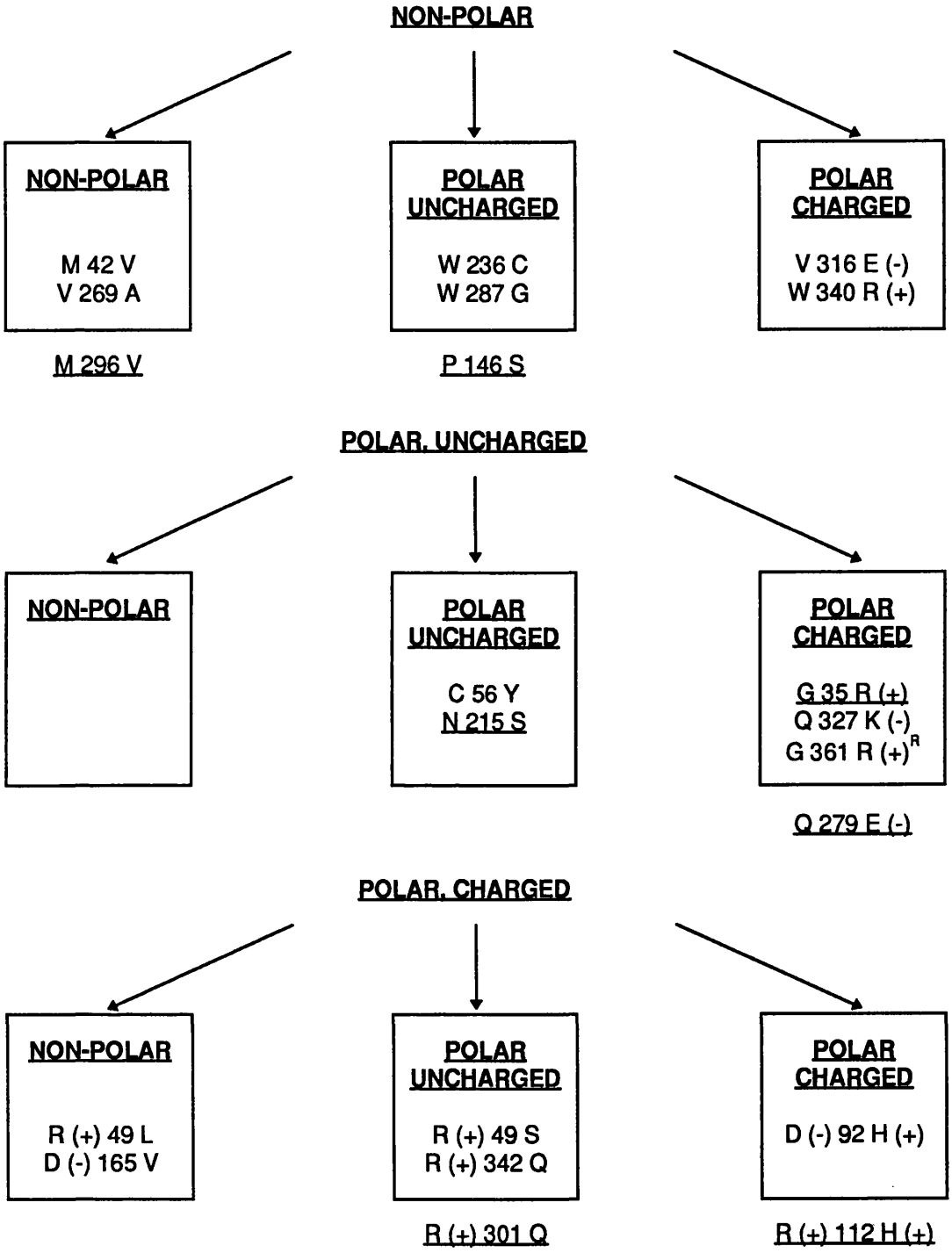


Figure 33: Diagram indicating the types of amino acid substitutions found in this study and all others that were found to give an atypical phenotype

The alterations in hydrophobicity caused by amino acid substitutions are indicated by the arrows. Side chains which are basic and acidic are indicated by a (+) and (-) respectively. Underlining indicates a mild or variant phenotype and the ^{R} indicates a mutant that was found in a patient with typical symptoms but with residual activity in *in vitro* expression. Mutations in the boxes were found in this study and those outside the boxes were found by others.

Further insight into the influence of mutations on the phenotype may be gained from purification and kinetic study of the mutant proteins. Studies of the *in vivo* processing and trafficking of the mutants and knowledge of the folded protein structure will increase our understanding of the effects of mutations.

4.3.4 Small rearrangements in the α-galactosidase A gene

4.3.4.1 The distribution and frequency of small rearrangements

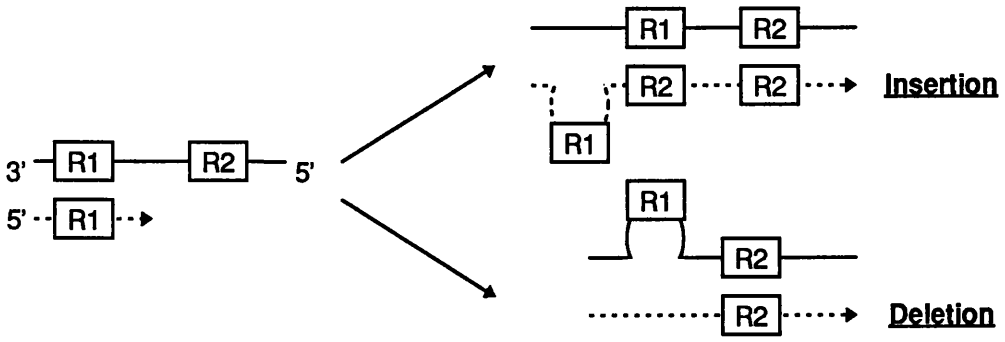
In this study, three small deletions (717del2, 1010del19 and 1087del1) and one insertion (716ins1) have been identified. The locations of these rearrangements and of other published rearrangements are indicated in figure 34. Two of these, 716ins1 and 717del2 are found in a region containing repeated nucleotides and they may have been formed by misalignment during DNA synthesis at the replication fork, as indicated in figure 35(A). The 717del2 mutation has been found in two families, of Belgian and British ancestry, that are thought to be unrelated. The presence of these two mutations in three families indicates a region that is particularly prone to replication errors. The 1087del1 mutation is flanked by two CTC direct repeats and since neither of the repeats is lost, the mechanism indicated in figure (A) does not explain this deletion. An alternative mechanism has been suggested to account for this type of deletion, figure 35(B) (Cooper *et al.*, 1991; Krawczak *et al.*, 1991). The CTC direct repeats can allow a misalignment during formation of the replication fork and excision of the resulting looped DNA may have deleted this base.

The mechanism for generating the 1010del19 mutation is unclear, since it is not flanked by repeats. However, it is in an area of DNA which includes three out of the four consensus sequences (GAG, GCC, GCG and ACG, Weaver *et al.*, 1982), indicated in figure 34, which cause DNA polymerase α to pause in replication. A second mutation, 1016del11, that is not flanked by direct repeats has also been found in this region and it has been proposed for this mutation that slowing of the polymerase may increase the possibility of an error during replication (Eng *et al.*, 1993, 1994a). Perhaps these two deletions occur during synthesis of DNA in the 3'-5' direction. The polymerase may perhaps dissociate at this sequence and if the secondary structure of the parent strand forms a loop then adjacent, newly synthesised Okazaki fragments may be brought together and ligated without synthesis of the intervening sequence.

Figure 34: Diagram indicating the location of 16 small rearrangements in the cDNA

Underlining indicates the region of duplicated or deleted DNA and the allele designations are indicated above. The boxes indicate the mutations found in this study. Consensus sequences are indicated below the sequence and in italics: α are DNA polymerase α arrest sites (Weaver, 1982) and poly A are signals for translation termination. Exon boundaries are indicated by vertical lines in the sequence. The alleles are designated according to the cDNA nucleotide number which is adjacent and 5' to the first altered sequence (reading in a 5'-3' direction), 'd' and 'i' indicate that the bases are deleted and inserted, respectively and this is followed by the number of nucleotides involved. The exact breakpoints are not always ascertainable.

(A) Slipped mispairing during DNA synthesis causes small deletions and insertions



(B) Slipped mispairing during formation of the replication fork and excision causes deletions

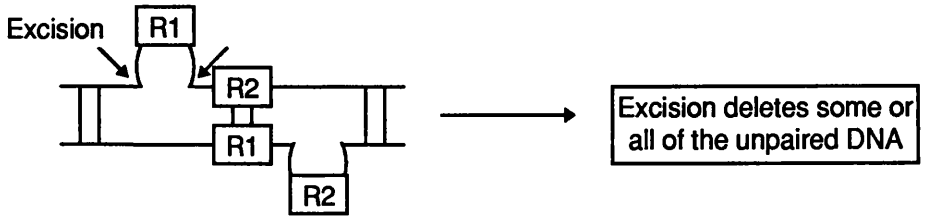


Figure 35: Slipped mispairing creates small deletions and insertions at the replication fork

'R' is a directly repeated sequence; —— is DNA at the replication fork; ····▸ is newly synthesised DNA at the replication fork.

Therefore, three different mechanisms could account for the formation of the four small rearrangements found in this study: slippage during DNA synthesis, promoted by repeated nucleotides, loop formation at direct repeats followed by excision and polymerase errors at α arrest sites. Similarly, all 16 mutations small rearrangements found in Fabry disease, shown in figure 34 can be explained by these three mechanisms. Eight involve direct repeat sequences and 4 involve single, repeated nucleotides. Two other, 25del1 and 773del2, delete part of the flanking inverse repeat sequences and may have been generated by formation of a hairpin loop structure, followed by DNA excision. The remaining 2, 1010del19, 1016del11, do not involve repeated sequences but are in a region containing DNA polymerase α arrest sequences. The presence of direct repeats, inverse repeats and palindromic DNA sequences at the sites of small rearrangements has been noted and reviewed (Cooper *et al.*, 1991; Krawczak *et al.*, 1991) and lends support to the hypothesis that slipped mispairing is involved in their genesis. In addition to the deletions and insertions indicated in figure 34 and discussed above, two 'complex' in frame mutations have also been reported (Eng *et al.*, 1994b). In both mutations, I219N and L120P-A121T, either two point mutations have occurred or a deletion and insertion have altered two bases in each mutation.

Although these rearrangements are promoted by the surrounding sequence, most of the mutations are in different regions. However, a few areas have more than one rearrangement mutation. These are in exons 2, 5 and 7. In exon 2, the 333del18, 358del6 and the complex mutation, L120P-A121T, are in close proximity. In exon 5, the 717del2 mutation is found in two families and the 716ins1 involves slippage at the same site. Finally, exon 7 has four rearrangements that are within a 30bp region, 1010del19, 1016del11, 1020del1 and 1040ins1.

4.3.4.2 The effects of small rearrangements

All four small rearrangements are predicted to create a premature stop codon in the same exons in which they occurred. Therefore the effects of the mutations are expected to be similar to those for stop mutations, such as production of a truncated protein or cause exon skipping or a reduction in the amount of mRNA. Examples of small and large deletions that create stop codons and cause exon-skipping have been reported from analysis of other genes, for example, a 2bp deletion in the retinoblastoma gene (Mori *et al.*, 1990) and a 52bp deletion in dystrophin (Matsuo *et al.*, 1991). The premature stop codons caused by 716ins1

and 717del2 could result in skipping of exon 5, a reduction in the amount of mRNA and production of a truncated protein. A frameshift would occur from exon skipping because the boundary of exon 4 to 5 is not in-frame. The nonsense codons produced by two frameshift deletions, 1010del19 and 1087del1, in exon 7 would not be expected to cause exon skipping since they are in the last exon. A reduction in the mRNA is not predicted for stop mutations in the last exon (Cheng *et al.*, 1990) and production of truncated proteins is most likely. However, no studies on the mRNA or protein in cells from patients with these mutations have been carried out to confirm these predictions.

These considerations suggest that severe phenotypes are expected from frameshift mutations. All four small frameshift mutations identified in this study, 716ins1, 717del2, 1010del19 and 1087del1, are associated with typical Fabry phenotypes and three, 716ins1, 1010del19 and 1087del1, were found in hemizygotes (in families 23, 26 and 27 respectively) with no residual α-galactosidase activity (tables 5, 18 and 29). No frameshift mutations in this or other studies have been found in Fabry hemizygotes with a mild or variant phenotype. However, a variant phenotype was found in a Fabry hemizygote with an in-frame, 3bp deletion, 1208del3, that deleted an arginine at amino acid 404 (Eng et al., 1994b). The amino acid sequence encoded by exon 7 of the α-galactosidase A gene has only 15.8% identity with that encoded by the corresponding α-N-acetylgalactosaminidase sequence (Wang, AM *et al.*, 1990) and it includes the deleted arginine 404, which is not conserved. The patient phenotype was moderately severe with cardiac symptoms primarily (Eng *et al.*, 1993). In other lysosomal storage diseases, even frame-shift deletions have been shown to cause a mild effect. For example, in two mildly affected Hunter patients, one had a 1bp deletion in exon 1 of the iduronate 2-sulphatase gene and the other had a mutant splice site in intron 7, both which were detected by analysis of the mRNA. They were predicted to cause frameshifts and premature stop codons (Hopwood *et al.*, 1993a).

It is concluded that frameshift mutations in the α-galactosidase A gene correlate with severe phenotypes, consistent with their predicted effects at the protein level but that this is not true of all lysosomal diseases.

4.3.5 Conclusions for genotype and phenotype correlation

Attempts to correlate the clinical phenotype of Fabry patients with the genotype have proved difficult for two reasons, relating to the collection of clinical data. Firstly, the disease is slowly progressive and the available clinical data have been collected at different ages by different clinicians. So direct comparison of different patients is complicated. Secondly, most mutations in Fabry disease are unique to one family. Clinical data from more than one affected male with the same mutation is often unavailable to aid the correlation of genotype with phenotype. However, despite these problems, seven obviously atypical Fabry patients have been identified and their mutations have been identified in this and other studies.

Information regarding the clinical phenotype associated with a particular mutation may be useful in patient counselling. However, it is complicated by intrafamilial and interfamilial variation. The intrafamilial variation associated with the R227X mutation in family 20 has been described (4.3.2.3) as has the variation in hemizygotes with the D92H mutation in family 6 and the phenotypes associated with R112H (4.3.2.4.1). Intrafamilial variation is not restricted to Fabry disease and has been observed in other lysosomal storage diseases, such as α-mannosidosis (Mitchell *et al.*, 1981), Sanfilippo syndrome type A and Hurler (McDowell *et al.*, 1993) and several others (Zlotogora, 1987). Interfamilial variation has also been observed in Fabry patients with the N215S mutation, in which one patient had proteinuria and two others cardiac symptoms (Eng *et al.*, 1994a). A splice site mutation in another lysosomal storage disorder, Sandhoff disease, caused different phenotypes in patients from different racial backgrounds (McInnes *et al.*, 1992). This variability shows that factors other than the disease-causing mutation can influence the phenotype.

In Fabry disease, one of the genetic factors that may influence the phenotype is inheritance of certain blood group antigens, groups B and P, which are substrates for α-galactosidase A, and whose presence may increase the substrate load on the lysosomal system. Fabry patients with the B group antigens were found to be more severely affected than those without in one study (Kint *et al.*, 1973). However, inheritance of the B antigen did not account for the clinical variability in family 20 with R227X, in which all hemizygotes had blood groups A_1, A_2 or O. Another genetic factor may be the expression of glycotransferase B^2, which is more efficient in producing the B antigens than the B^1 form (Wherret *et al.*, 1973; Yoshida, 1983; Desnick *et al.*, 1989). However, the involvement of these factors in altering the

phenotype of Fabry patients has not been investigated. The effects environmental factors, such as diet, in altering the phenotype have not been studied either.

In conclusion, genotype correlation with the phenotype is complicated by variation caused by unknown factors and this should be considered for patient counselling. Analysis of the types and locations of mutations have shown no obvious patterns for those producing variant and severe effects. Understanding their effects requires information on the protein structure.

4.3.6 Carrier identification by direct mutation analysis

4.3.6.1 Comparison of genetic data and α-galactosidase activity in heterozygotes

Carrier detection by analysis of the intragenic polymorphisms detected in this study is limited by the frequency of the alleles. However, detection of specific mutations by either SSCP analysis using the standard conditions, digestion with restriction enzymes or a modified method of SSCP analysis, has allowed conclusive identification of the carrier status of 27 females in 12 families. Of these, 18 were heterozygotes and 9 were normal. Their α-galactosidase activities are shown in tables 16 and 17, pgs 149-150. In 8/18 heterozygotes (44%), either the leukocyte and/ or the plasma enzyme activities were within the normal range and they could not have been designated as carriers. All females with enzyme activities below the normal range were confirmed as carriers.

Previous carrier ranges of α-galactosidase activity in leukocytes and plasma were derived from females known to be obligate heterozygotes (table 31). This range included females with an affected father or those with an affected descendant and at least one other affected relative in another branch of the family. However, family history can be unreliable, as obligate carriers with an apparently affected father could be the result of non-paternity. This had been suspected in female III-3, in family 7 and was formally shown in this thesis by analysis of polymorphisms (figure 17(A), pg 131) and the D165V mutation (figure 25(A), pg 153). Female III-3 had normal activity in hair roots, leukocytes and plasma (Personal communication from Dr B. Winchester and Mrs E. Young).

The previous ranges of enzyme activities for heterozygotes, based on obligate carriers, can now be modified to incorporate data from an additional 8, newly defined heterozygotes, analysed by genetic methods. Mutation analysis has confirmed the status of the 11 obligate heterozygotes that were already included in this range (table 31). Normal female III-3 from family 7 was removed from the heterozygote data group, decreasing the range of plasma activities to 0.39-6.4nmol/hr/mg but leaving the leukocyte activity range unchanged. The addition of 8 values for non-obligate carriers, increased the plasma range to 0.39-9.0 nmol/hr/mg but again, did not change the leukocyte range. The mean activities in heterozygotes, for leukocytes and plasma, are less than 1 standard deviation away from the minimum normal activities. Therefore, it is not surprising that 6/18 (33%) of enzyme assays on leukocytes and the same number on plasma from heterozygotes, were within the normal ranges (tables 16 and 17, pgs 149-150). The highest activities in plasma, 9.0, and in leukocytes, 67, of genetically defined heterozygotes were almost 3 and 1-2 standard deviations away from the mean, respectively and well within the normal ranges. The carrier status of a female with the highest leukocyte activity, 89, could not be confirmed genetically but she had one affected son and at least one other affected relative, who was not a descendant. Thus, incorporation of data from heterozygotes proven by DNA analysis has not extended the enzyme activity ranges, although the amount of accurate data and therefore the calculated means and standard deviations have been improved. These studies show conclusively that all females with α-galactosidase activities below the normal range, in leukocytes and/or plasma, are heterozygotes and that mutation detection is a feasible and useful approach to carrier detection in Fabry families.

The normal ranges of α-galactosidase activities in plasma and leukocytes (table 31) were unchanged by the results from 9 females that were genetically normal (table 16, pg 149). However, it should be noted that three of these apparently normal females, one (figure 25(B), pg 153, I-2) in family 5 and two (table 16, pg 149, I-2 and II-6) in family 11, may be somatic and/or germ-line mosaics (Hall, 1988) for the disease-causing mutation.

Tissue	Activities in	Sample numbers	α-galactosidase activity (nmol/hr/mg)		
			Range	Mean	Standard deviation
Leukocytes	Normals	106	40-162	72.4	21.2
	(A) Obligate heterozygotes	21	5.2-89	36.2	21.5
	(B) Genetically determined & obligate heterozygotes	28	5.2-89	35.9	19.5
Plasma	Normals	63	4.8-26.5	8.6	4.3
	(A) Obligate heterozygotes	17	0.39-9.8	3.5	2.4
	(B) Genetically determined & obligate heterozygotes	24	0.39-9.0	3.6	2.0

Table 31: Activity ranges in normals, obligate heterozygotes and with data incorporated from genetically determined heterozygotes

All enzyme data was obtained from Mrs E. Young in the Enzyme Laboratories at the Institute of Child Health. (A) indicates the previously determined heterozygote range. The new range, (B), incorporates data from tables 16 and 17.

4.3.6.2 New mutations and somatic mosaicism

In this study, one base substitution, C56Y, was found to be a *de novo* mutation and two others, W287G and M42V, were suspected of being new mutations. In family 5 (figure 25(B), pg 153), C56Y was not found in lymphocyte DNA from the mother, I-2, of an affected child and so the mutation must have arisen in her germ-line or early in the development of her affected son. In family 11, the mother, II-6, of an affected child carried the W287G mutation but the grandmother, I-2, did not (table 16, pg 149). Similarly, in family 2 (figure 23(B), pg 127), the mother, I-2, of the hemizygote carried the M42V mutation while the grandmother, I-1, did not. In families 11 and 2, DNAs from the grandfathers were not available for analysis but they was not thought to be clinically affected. The mutation may have arisen *de novo* in the grandmother, grandfather or mother, or alternatively, the mutations are the result of non-paternity. The occurrence of mosaicism, due to *de novo* mutations, presents a problem in calculating the risk of producing affected children and in genetic counselling.

The probability that an individual has a *de novo* mutation that will be inherited, is dependent on the mutation rate and the proportion of germ-line cells that carry the mutation. In X-linked disorders, the proportion of *de novo* mutations in males is related to rate of mutation and their elimination from the population, which is influenced by the ability of affected males to reproduce and is given by the equation:

$$(1 - f)\,\mu \,/\, 2\,\mu + v$$

where 'f' is the effective fertility of males, and 'v' and 'μ' are the mutation rates per generation in male and female gametes, respectively (Haldane, 1935; Weatherall, 1991). If the males cannot reproduce then f = 0 and the proportion of mutations arising in males is 1/3, when the mutation rates in males and females are equal, ie. $v = \mu$. The proportion of *de novo* mutations in males is decreased in disorders, such as Fabry disease, in which the males can reproduce. This information can be incorporated into calculations of the probability of mosaicism in the mother of an isolated, affected male.

In Fabry disease, *de novo* mutations have been found, by others, in only 2 families, one with a 402bp deletion which was not present in the grandparents but was found in the mother of an affected child (Bernstein *et al.*, 1989) and one male with a complex mutation, I219N, that was not detected in his mother (Eng *et al.*, 1994b). In a third family, a 1710bp deletion/inversion mutation was also suspected of being a *de novo* mutation which was present in the mother of an affected male, but not the grandmother. The L32P mutation, found in a heterozygote but not in either of her two parents, was also believed to be a *de novo* mutation, although the possibility of non-paternity was not excluded (Madsen, *et al.*, 1995). However, no systematic study of the incidence or origin of *de novo* mutations in Fabry hemizygotes has been carried out.

In individuals that are known to be germinal mosaics it is possible to calculate the risk that they will have an affected child (Jeanpierre, 1992; van der Meulen *et al.*, 1995) based on empirical data, as in Duchenne Muscular Dystrophy (Bakker *et al.*, 1989; van Essen *et al.*, 1992) and models of gametogenesis (Hartl, 1971). The recurrence risk is dependent on the total number of cell divisions during gametogenesis and the generation in which the mutation occurred, determining the proportion of germ-line cells that carry the mutation. A single mutation in the first generation will segregate to produce 50% of cells with the mutation and the children will have a 50% risk of inheritance. Since it is not possible to determine the exact proportion of mutated gametes, risk assessments are based on probability calculations. If it is assumed that each cell generation is equally likely to produce a mutation and that only one parent is a mosaic then the risk of inheriting a mutation is calculated by:

$$\sum_{i=1}^{n} [\, 1 / n\,(2^{1})\,]$$

where 'n' is the total number of cell divisions (van der Meulen *et al.*, 1995). Gametogenesis produces an estimated 5-7 million oocytes and a even greater number of spermatozoa and these must have been produced from at least 22 and 30 divisions, respectively (i.e., 2^n cells, where 'n' is the number of divisions). For this many generations, the above equation for the inheritance risk approximates to 1/n, giving a risk of 4.5% if n=22 (van der Meulen *et al.*, 1995). A more accurate risk assessment can be obtained by taking into account the number of unaffected and affected children produced by a germinal mosaic. Thus, the probability of

inheritance of a mutation in families in which there are 1, 2, 3 and 4 affected children and no unaffected children, can be calculated as 4.8%, 33%, 42% and 47%, respectively (Hartl, 1971; van der Meulen *et al.*, 1995). These calculations assume that a particular individual is a mosaic but in the Fabry families analysed in this study, the individual in which the mutation arose is not known and so risk assessment, based on the probability that a new mutation has occurred and that there is a certain degree of germinal mosaicism, is more complicated. In Fabry disease there are no reported cases of mosaicism, although there are many reports in other X-linked disorders, such as Duchenne Muscular Dystrophy (Bakker et al., 1989), X-linked agammaglobulinaemia (Hendriks et al., 1989), X-linked Severe Combined Immunodeficiency (Puck et al., 1995), Haemophilia A (Gitschier, 1988) , Haemophilia B (Taylor et al., 1991) and ornithine transcarbamylase deficiency (Maddalena et al., 1988).

Molecular genetic analyses simplifies the risk assessments by proving that there is germ-line mosaicism. If one affected and another unaffected child have inherited the same chromosome then mosaicism in proven. In family 5, the risk of Fabry disease in siblings of the mother, who may be a germinal mosaic, is decreased from the probability that they inherited a pre-existing mutation to the probability that another new mutation will arise in the α-galactosidase A gene.

4.3.6.3 Improvements of the strategy for mutation detection

Analysis of the distribution and frequency of the mutations has not revealed many common mutations that can be screened for, or small regions of the gene that are particularly prone to mutation. The recurrent mutations R227X and R227Q account for only 5.4% of mutations (Eng *et al.*, 1993). The distribution and frequency of the fourteen, presumably methylated CpG dinucleotides that are more prone to mutation, is shown in table 32, below:

Table 32: Comparison of the exon length, the number of CpG sites outside the CpG island and the frequency of all types of mutation

All 93 different mutations found in Fabry patients, from table 4, are included.

In codons or splice sites for exons	Exon length (bp)	Number of CpG sites in exon	Number of mutations	% of total mutations (93)	[No. of mutations/ No. of bp in exon] x100
1	254	2	15	16	5.9
2	175	2	14	15	8.0
3	178	2	9	10	5.1
4	92	0	3	3	3.3
5	162	2	14	15	8.6
6	198	2	17	18	9.6
7	289	4	15	16	5.2
large rearrangements	-	-	6	6	-

The fact that the locations of the CpG dinucleotides are scattered does not allow for fast detection of mutations at these sites and would only detect mutations in about 20% of families. A more efficient screening strategy could be to analyse exons in a preferred order, based on the proportions of mutations that have been found in each exon. The number of different mutations found in the exons and their flanking splice sites varies from one region to another, which is partly related to the size of the exons, although some exons have a higher proportion of mutations per number of bases than others (table 31). Exon 6 has more than three times more mutations in its codons and splice sites, per base, than exon 4. Based on the observation than some regions have more different mutations than others and coupled with the fact that the recurrent mutations at codon 215, 227 and 342 are in exons 5, 5 and 7, an efficient mutation-screening strategy would analyse exons 5, 6 and 7 first. This would detect mutations in more than 50% of families if the distribution of mutations in other sample populations is the same as those already studied. Exons 1 and 2 could then be examined, while exons 3 and 4, which contain the fewest numbers of mutations, could be analysed last.

In conclusion, the use of genetic methods to detect unknown mutations for use in carrier detection is preferable to diagnosis by enzyme analysis. Although there are no common mutations or small regions that are particularly prone to mutation, an efficient strategy for identifying mutations can be based on knowledge of their distribution.

4.3.7 The limitations of the expression system and possibilities for its modification

The transient expression studies have successfully indicated which mutants give rise to residual activity and/or produce material that cross-reacts with anti-(α-galactosidase A) serum. However, a number of problems were associated with these studies. Direct comparison of the absolute α-galactosidase activity values produced during one transfection experiment required several assumptions to be made regarding the constancy of factors that could influence the expressed activities. Another problem was the irreproducibility of absolute activity values from one experiment to another. A third problem was the inability to distinguish between the endogenous COS-1 cell enzyme and the transiently expressed enzymes. This prevented complete isolation of the mutant proteins and accurate studies of properties of the mutant enzymes. Isolation of the mutant proteins and their study may provide further information regarding the effects of mutations at the protein level.

Direct comparison of the results obtained for expression of different constructs, even in the same experiment, is dependent on several assumptions: (a) the number and state of the COS-1 cells used for transfection of each construct was similar; (b) the endogenous specific activity of α-galactosidase A in the samples of COS-1 cells was the same; (c) the expression of the mutant construct did not alter the amount of background α-galactosidase A that was produced from the COS-1 cell gene, or the proportion that was secreted; (d) the efficiency of β-galactosidase transfection and its translation was independent of that of the α-galactosidase A constructs. These assumptions allowed comparison of the relative levels of expression of the constructs that were electroporated in the same experiment when using aliquots from the same pool of cells, but could not be tested.

The first assumption is probably correct, if the cells are well mixed before being subdivided for transfection. The other assumptions are less certain. There is no way of distinguishing between the endogenous COS-1 cell activity and the expressed mutants as the polyclonal, human antiserum cross-reacts with both. Therefore there is no way of knowing whether the amount or distribution of endogenous α-galactosidase A alters during these experiments. As, α-galactosidase A is a homodimeric protein, it is possible that the endogenous enzyme will dimerise with the expressed endogenous protein. This is not a problem when measuring the activity of the normal human α-galactosidase A construct, which is high but for the mutants with low residual activity it may significantly alter the results. For

example, the G361R mutant produces intracellular activity which is only 1.6 fold higher than the endogenous activity when no construct is electroporated. This activity may represent human:COS-1 α-galactosidase A heterodimers and there may be zero activity produced by the mutant homodimer. Therefore, the sensitivity and accuracy of the system are decreased. The kinetic analysis of the activity would also be inaccurate, making characterisation of catalytically active mutant proteins difficult. In addition, it is not known whether the coexpressed *E. coli* β-galactosidase activity is a true reflection of the efficiency of the transfection and expression of the α-galactosidase A constructs and therefore whether it is a reliable control for normalisation of enzyme activities. Few studies have been conducted to determine the parameters which affect the transfection efficiency in mammalian cells and it is not clear whether cotransfection alters this (Chu *et al.*, 1987).

An ideal system would either produce no endogenous α-galactosidase A activity, or it would allow the expressed protein to be distinguished from the endogenous enzyme. This could be achieved by expression in cells that do not produce α-galactosidase A or by devising a method to selectively isolate the expressed protein. An ideal system would also permit measurement of reliable reference marker to allow direct comparison of expressed enzyme activities from different transfection experiments. The following options for modification of the expression system could be considered in future studies.

Cell which do not have endogenous α-galactosidase A activity, such as enzyme-deficient lymphoblastoid cells or fibroblasts from Fabry patients, could be used. However, they would require genetic modification to allow sufficient overexpression with existing vectors to observe the mutant enzymes with low activities. No α-galactosidase A-deficient derivatives of the mammalian cell lines used for overexpression, such as COS-1 cells or CHO cells, have been reported and so genetic manipulation would be required to inactivate the endogenously produced enzyme.

Alternatively, the expressed proteins could be formed with an attached fusion protein tag, allowing them to be selectively purified. No suitable systems have been described for expression in mammalian cells. However, several purification systems have been described for proteins expressed in bacteria. In one example, using the pRIT2T vector (Pharmacia catalogue), a fusion protein which includes a fragment of protein A is expressed in bacterial cells, allowing purification by binding IgG-Sepharose. Using the pGEX vector (Pharmacia), the fusion protein

includes a portion of glutathione S-transferase, which binds to glutathione-Sepharose. Genetic manipulation of the existing mammalian expression vectors could be used to express similar fusion proteins. However, the effects of the fused protein on normal enzyme activity would need to be evaluated to verify that the stability and catalytic activity were not changed. This is probably the easiest method for allowing the expressed protein to be distinguished and would also allow easy purification and subsequent kinetic and stability analysis of the low activity mutants. In addition, this system would allow the enzyme activity to be related to the amount of α-galactosidase A protein, rather than to the variable total cell protein and it obviates the need for cotransfection of a reference marker.

The reliability of using a coexpressed reference enzyme as a control may be improved if the control protein is expressed from the same transfected vector as the α-galactosidase A mutant. Since coexpression of genes from one vector, under the control of 2 promoters can result in preferential expression of one of the 2 genes (Emerman *et al.*, 1984), vectors have been designed to allow translation of bicistronic sequences. In one such retroviral vector (Aran *et al.*, 1994), a drug resistance gene and human glucocerebrosidase have a common promoter and have been successfully translated during stable transfection into mouse cells.

The expression system used in this study allows the presence or absence of mutant activity or protein to be detected but modifications to the system may enable more extensive and accurate information to be obtained for the characterisation of mutant proteins.

4.4 FUTURE ASPECTS

From the results of this study it is concluded that SSCP analysis is an efficient method to localise unknown mutations, prior to sequencing and for heterozygote detection in Fabry disease. This genetic analysis allows carrier females to be conclusively identified and is a major improvement on the current biochemical detection methods. Mutation analysis and genotype-phenotype correlation will also aid genetic counselling by indicating that certain mutations result in variant and sometimes milder clinical manifestations. Therefore, both aims, to improve carrier detection and to correlate genotype with phenotype have been achieved. Further information regarding the effect of the genotype on the phenotype may be gained by studying the physical and chemical properties of the defective enzymes, along with their processing and transport. This will be facilitated by purification of the mutant enzymes, perhaps involving the use of different expression systems to the one used in this study.

The development of genetic techniques for carrier detection has overcome the problem of heterozygote detection in Fabry patients. However, the major problem of therapy has not been solved. The success in treatment of Gaucher disease by enzyme replacement, (Barton *et al.*, 1991), provides a good model for developing a therapeutic method for Fabry disease (Barton *et al.*, 1991). Ideally, the α-galactosidase A enzyme used in treatment will be stable, minimising the frequency of administration required to maintain a therapeutic threshold of activity. The ability and efficiency of the enzyme to reach the target cells is also an important consideration. Thus, studies aimed at determining the folded structure of the enzyme and the functional domains will allow identification of the essential amino acid sequences required for activity and targeting and may allow improvement of the native enzyme for therapy. Mutation analysis and investigation of the effects at the protein level will facilitate these studies. Information regarding the specific mechanisms for uptake of exogenously applied enzymes in the affected cell type is also required. Receptors that bind carbohydrate for receptor-mediated endocytosis of glycoproteins have been identified in lymphocytes, fibroblasts, glial cells, macrophages, myoblasts and hepatocytes and the distribution of different receptors varies with the cell type (Bou-Gharios *et al.*, 1993). Since the vascular endothelial cells are the primary site for accumulation and pathologic consequences in Fabry disease, a study of receptor-mediated uptake mechanisms present in these cells for internalising α-galactosidase A, will be of value. Methods of enzyme production are a third important consideration for enzyme therapy. Production of recombinant

enzyme that is secreted for easy collection and purification would be ideal. Recent clinical trials in type 1 Gaucher patients, using a recombinant form of glucocerebrosidase, have been successful (Grabowski *et al.*, 1995). This enzyme, cerezyme, is overexpressed and secreted from CHO cells, providing a theoretically infinite source of enzyme and reducing the possibility of infection by enzyme from human sources. Preliminary studies have been carried out for α-galactosidase A production in CHO cells (Iounnou *et al.*, 1992) and insect cells (Coppola *et al.*, 1994). One disadvantage with enzyme therapy is the requirement for repeated administration of enzyme. This could be overcome by transplantation of cells to provide enzyme continuously. Bone marrow transplants (BMT) have been carried out in many patients with metabolic disorders, although not in Fabry patients. BMT provides multiple enzyme-expressing cell types but is often limited by the availability of immunologically similar tissue (Hobbs, 1990).

Gene therapy has been successful in metabolic correction of a number of disorders caused by single enzyme deficiencies and strategies are currently under investigation for the treatment of lysosomal disorders. *Ex vivo* therapy, in which cells are withdrawn, genetically modified using retroviruses and transplanted back into the patient, circumvents the problem of tissue-matching for transplantation. This strategy has allowed effective treatment in animal models for several disorders, such as adenosine deaminase deficiency (van Beuschem *et al.*, 1990) and familial hypercholesterolaemia (Chowdhury *et al.*, 1992). Other strategies have been developed for gene therapy, including the use of lipofection and electroporation for *in vitro* manipulation and of adenoviral vectors, herpes-based vectors and direct gene transfer by injection, for *in vivo* therapy (Friedmann, 1992). The application of these types of method to therapy in Fabry disease has not been investigated and is limited largely by the lack of an animal model for testing. Thus, future research may be focused on creating a model for this disease to permit progress in research for gene therapy.

Treatment of Fabry patients is most likely to be achieved by enzyme therapy, rather than gene therapy, due to lack of an animal model for testing. However, one important consideration for all of these strategies is whether the patients produce protein that is immunologically similar to wild-type α-galactosidase A. Patients with no similar protein may be immunologically intolerant to the enzyme produced during therapy. Therefore it is important to investigate the nature of the genetic defect and its effect on the protein. This study has provided information that nine mutant proteins containing a missense mutation do cross-react with anti-(wild-

type α-galactosidase A) serum, indicating that they are immunologically similar to the wild-type enzyme. Patients with these mutations may be candidates for clinical trials of enzyme therapy in the future. Thus, this study has provided information that may facilitate future therapeutic strategies.

APPENDIX

Direct sequencing has identified 26 putative disease-causing mutations in the α-galactosidase A gene (figure 36).

In figure 36(A), mutations were detected by SSCP analysis, followed by sequencing. The Sequenase II method (2.2.7.2) was used to sequence all exons, except for exons 1 and 4, which were cycle-sequenced (2.2.7.3).

In figure 36(B), mutations were detected by sequencing all 7 exons of the gene, using the Sequenase II method (2.2.7.2).

The sequencing results for the mutations shown in figures 36 (A) and (B) are given in section A and section B, respectively. They are in order of their family number and the sequencing results for families 20 and 25 were the same as those in families 19 and 24, respectively. Gene numbering was as in Kornreich *et al.* (1989) and cDNA nucleotides were numbered as in Bishop, *et al.* (1991).

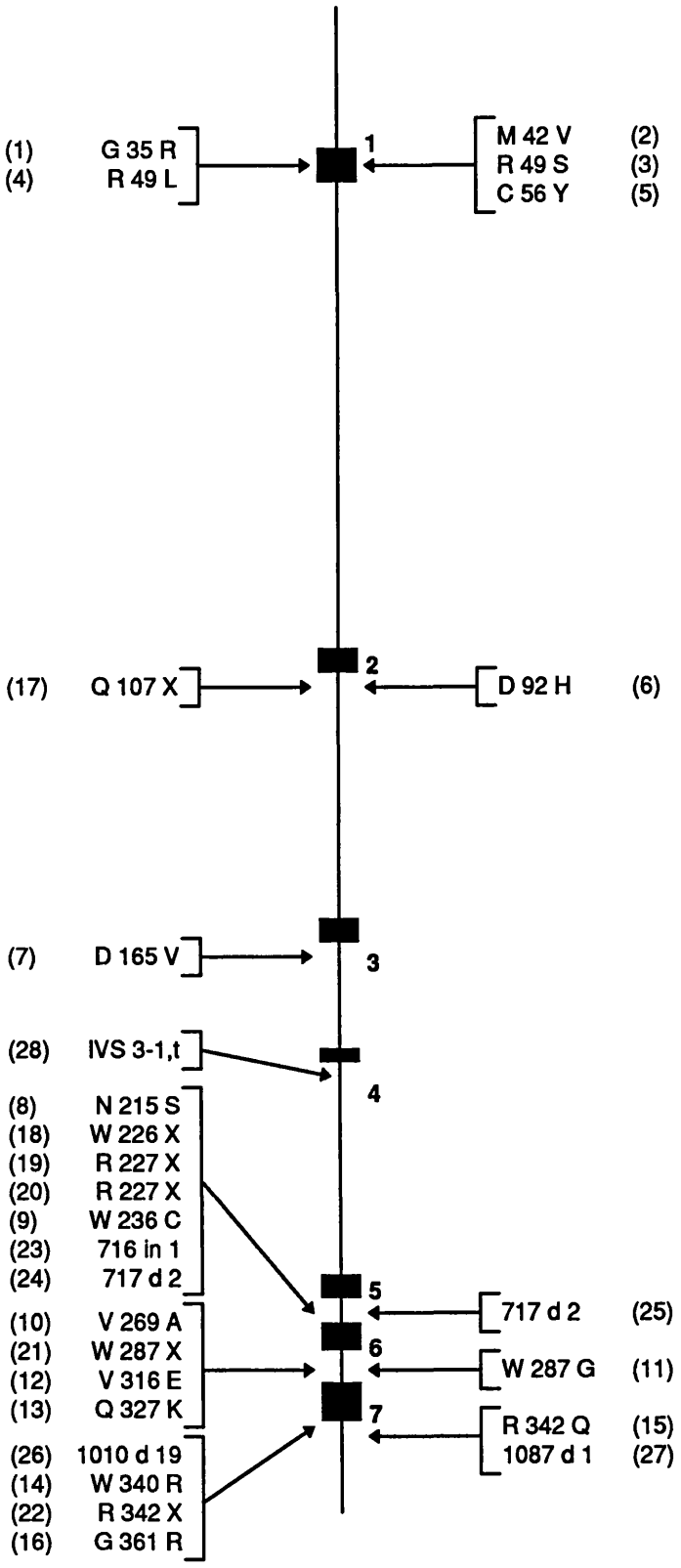


Figure 36: Location of 26 different mutations in 28 unrelated Fabry families

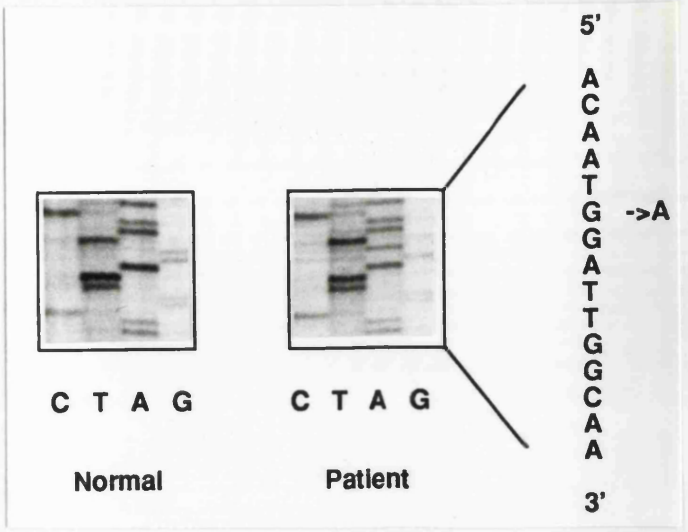
R 227X and 717d2 were found to occur twice. Family numbers are shown in brackets. Filled boxes represent the exons, which are numbered.

Appendix, Section A

Family 1: G 35 R (Hemizygote)

Glycine 35 -> Arginine
(Polar -> Polar, +ve)

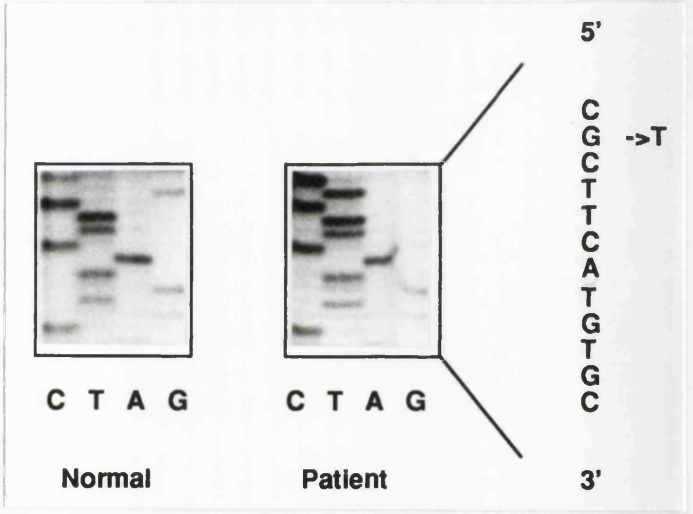
Gene No. 1282
cDNA No. 103 (Exon 1)



Family 4: R 49 L (Hemizygote)

Arginine 49 -> Leucine
(Polar, +ve -> Non-polar)

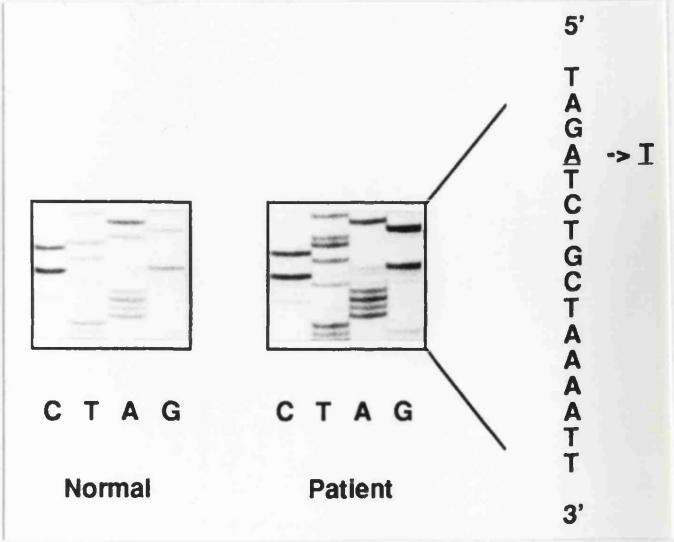
Gene No. 1325
cDNA No. 146 (Exon 1)



Family 7: D 165 V (Hemizygote)

Aspartic acid 165 -> Valine
(Polar, -ve -> Non-polar)

Gene No. 7393
cDNA No. 494 (Exon 3)

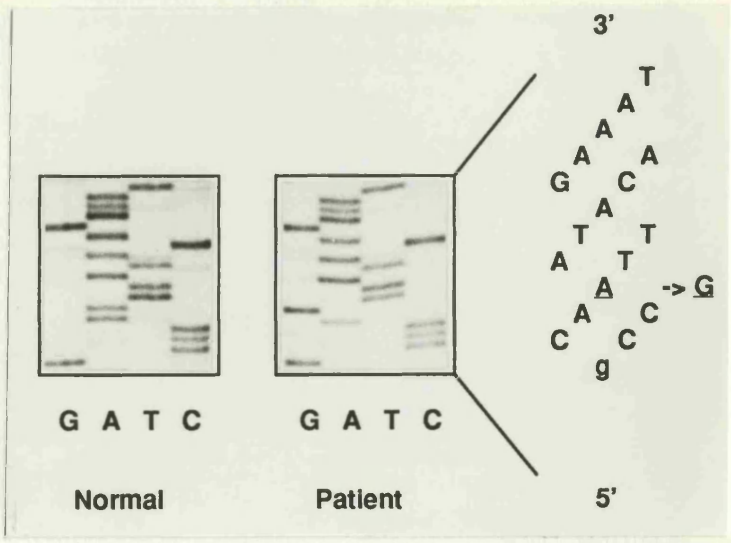


Appendix. Section A

Family 8: N 215 S (Hemizygote)

Asparagine 215 -> Serine
(Polar -> Polar)

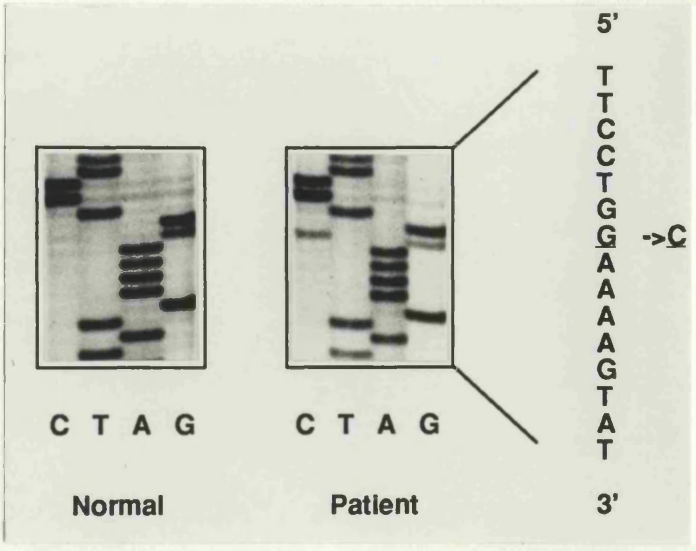
Gene No. 10135
cDNA No. 644 (Exon 5)



Family 9: W 236 C (Heterozygote)

Tryptophan 236 -> Cysteine
(Non-polar -> Polar)

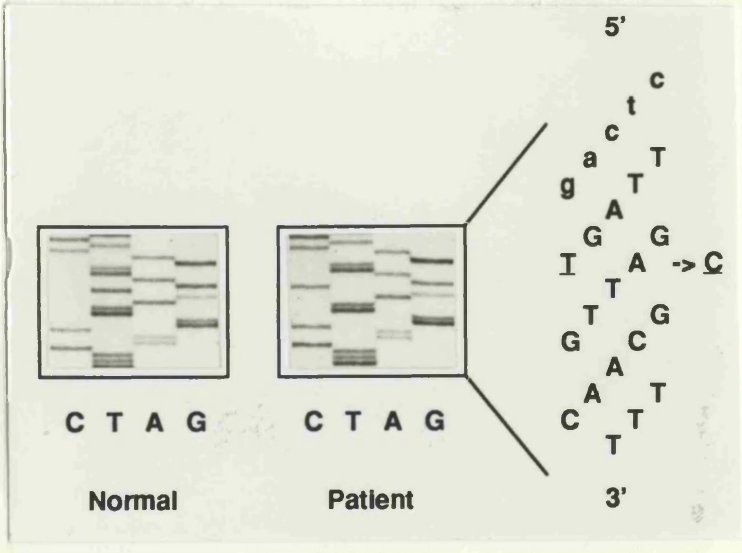
Gene No. 10199
cDNA No. 708 (Exon 5)



Family 10: V 269 A (Hemizygote)

Valine 269 -> Alanine
(Non-polar -> Non-polar)

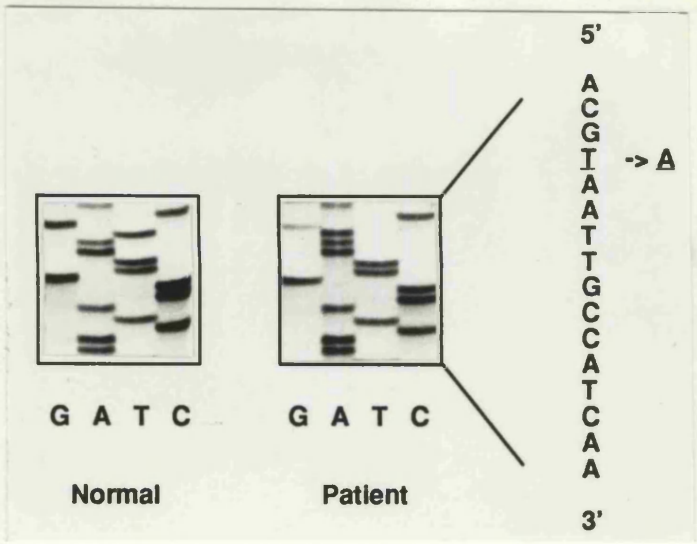
Gene No. 10514
cDNA No. 806 (Exon 6)



Family 12: V 316 E (Hemizygote)

Valine 316 -> Glutamate
(Non-polar ->Polar, -ve)

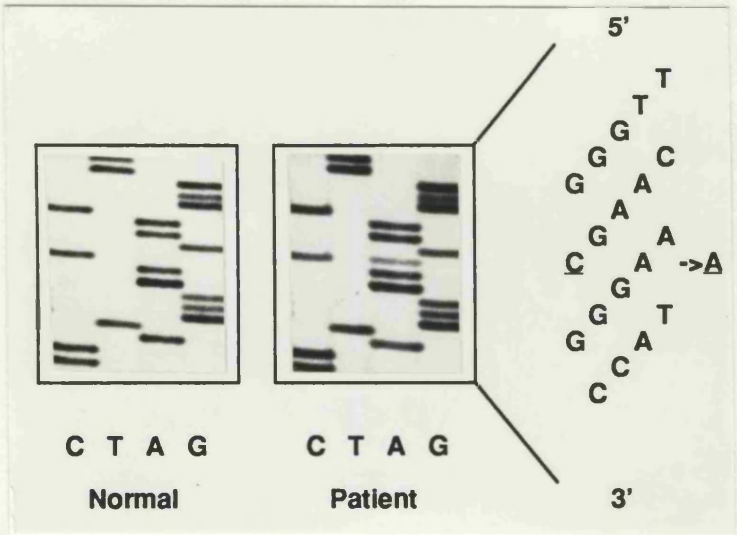
Gene No. 10655
cDNA No. 947 (Exon 6)



Family 13: Q 327 K (Heterozygote)

Glutamine 327 -> Lysine
(Polar -> Polar, +ve)

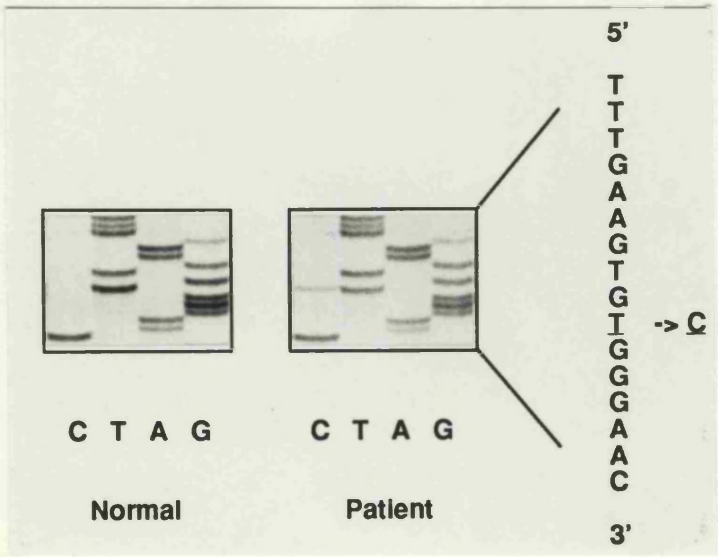
Gene No. 10687
cDNA No. 979 (Exon 6)



Family 14: W 340 R (Heterozygote)

Tryptophan 340 -> Arginine
(Non-polar -> Polar, +ve)

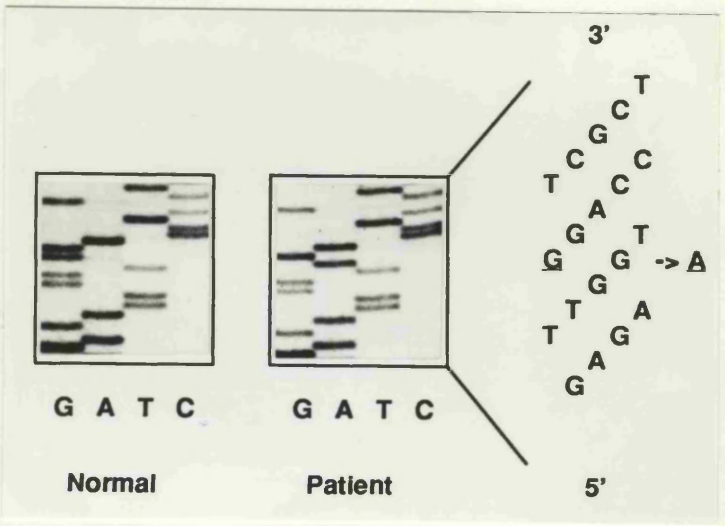
Gene No. 10996
cDNA No. 1018 (Exon 7)



Family 16: G 361 R (Hemizygote)

Glycine 361 -> Arginine
(Polar -> Polar, +ve)

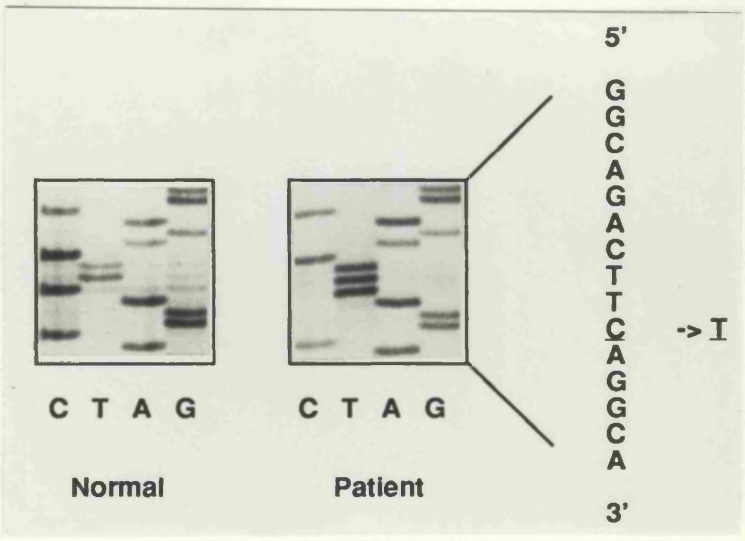
Gene No. 11059
cDNA No. 1081 (Exon 7)



Family 17: Q 107 X (Hemizygote)

Glutamine 107 -> Stop

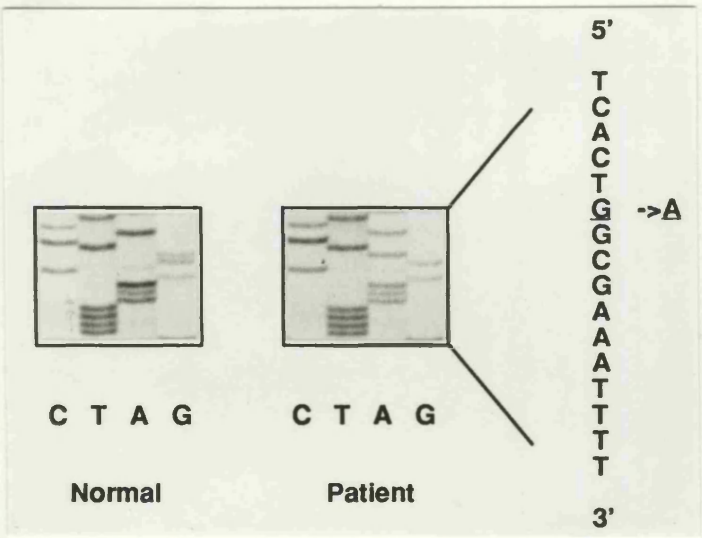
Gene No. 5218
cDNA No. 319 (Exon 2)



Family 18: W 226 X (Hemizygote)

Aspartic acid 226 -> Stop

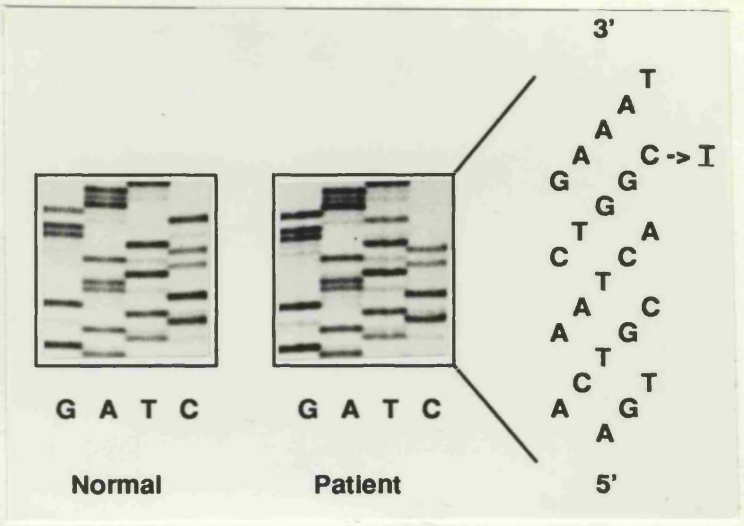
Gene No. 10168
cDNA No. 677 (Exon 5)



Family 19: R 227 X (Hemizygote)

Arginine -> Stop

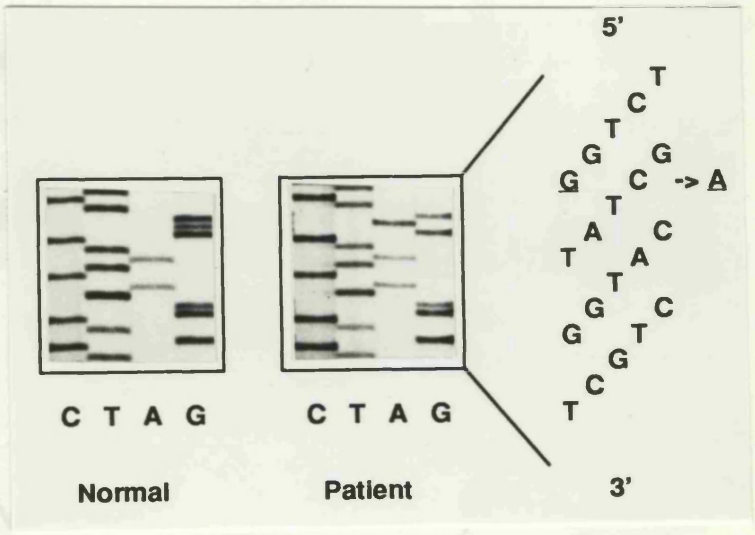
Gene No. 10170
cDNA No. 679 (Exon 5)



Family 21: W 287 X (Hemizygote)

Tryptophan 287 -> Stop

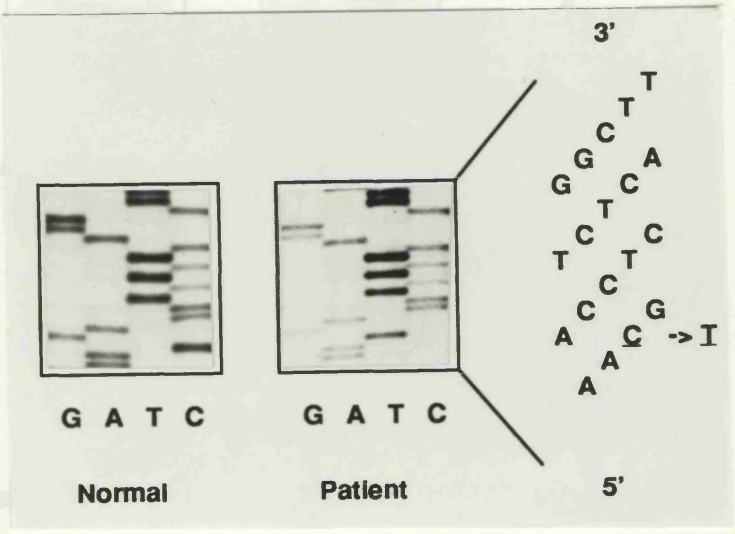
Gene No. 10569
cDNA No. 861 (Exon 6)



Family 22: R 342 X (Hemizygote)

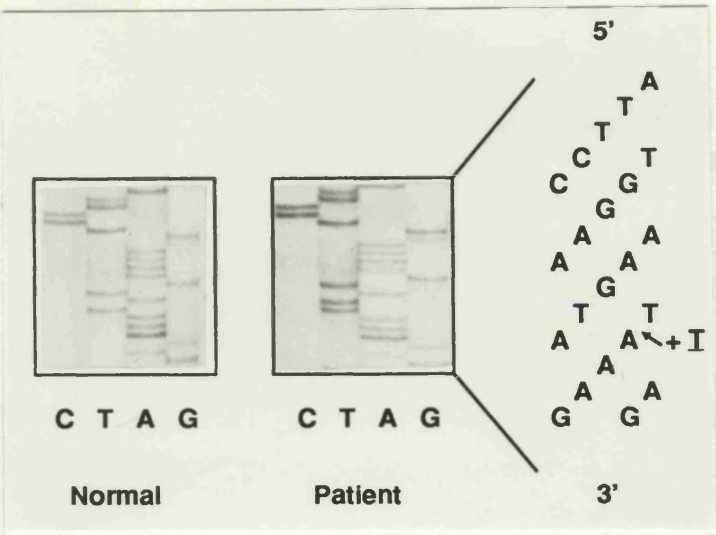
Arginine 342 -> Stop

Gene No. 11002
cDNA No. 1024 (Exon 7)



Appendix. Section A

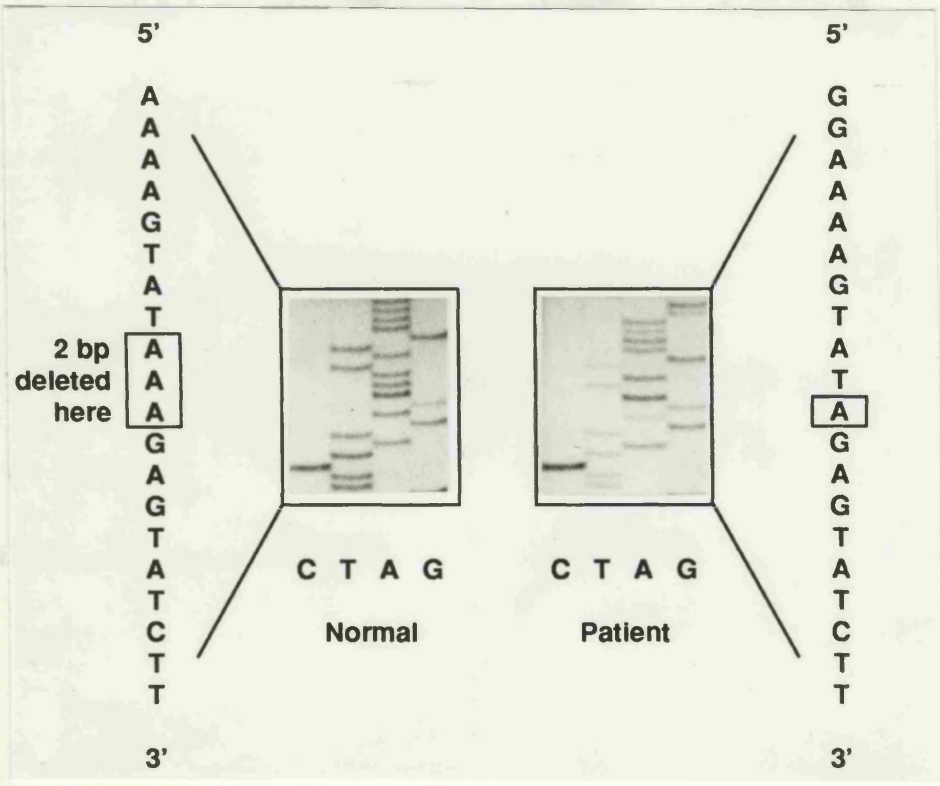
Family 23: 716 ins 1 (Hemizygote)



1 bp insertion alters amino acids 241-249: SILDWTSFN -> EYLGLDIF-stop

Insertion is between gene nucleotides 10207-10208 and cDNA Nos. 716-717 (Exon 5)

Family 24: 717 del 2 (Hemizygote)

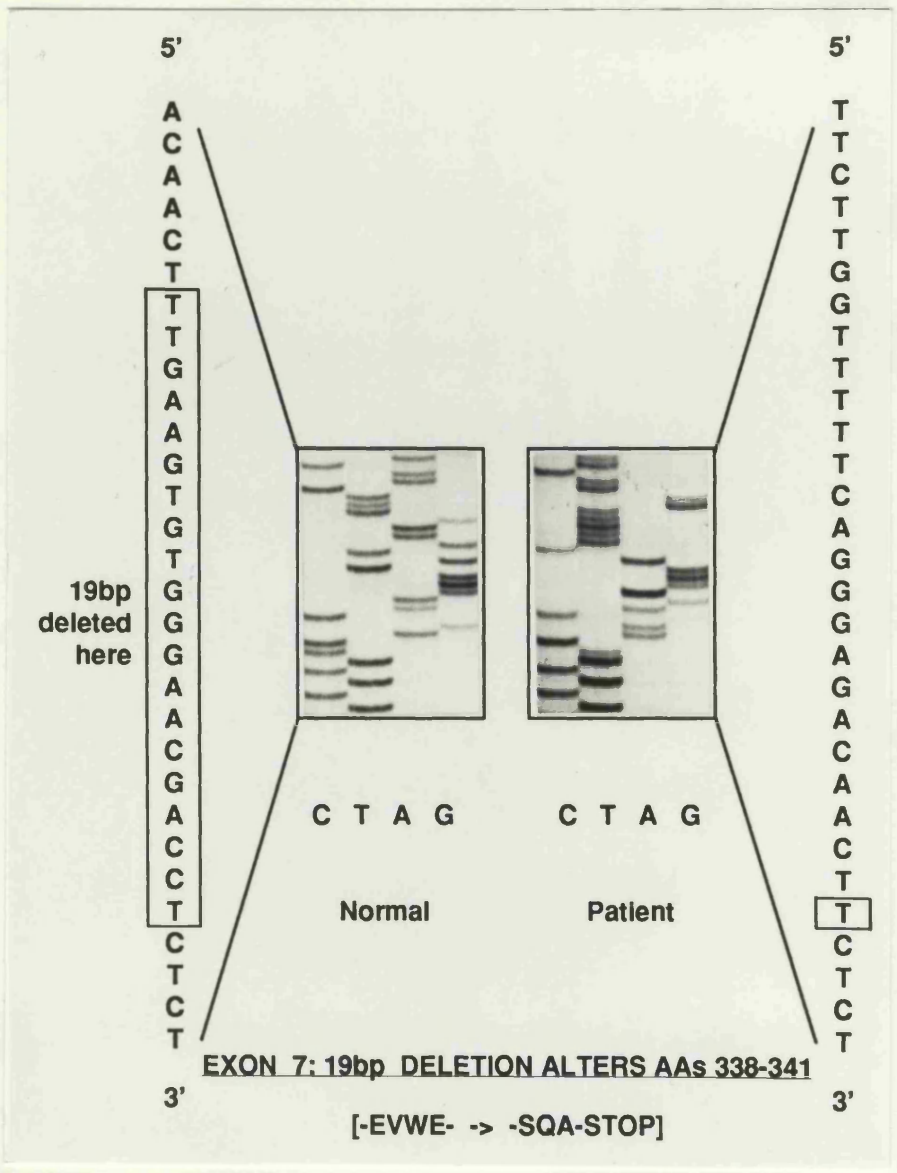


2 bp deletion alters amino acids 240-248: KSILDWTSF -> EYLGLDIF-stop

Deletion is between gene nucleotides 10208-10210 and cDNA Nos. 717-719 (Exon 5)

Appendix, Section A

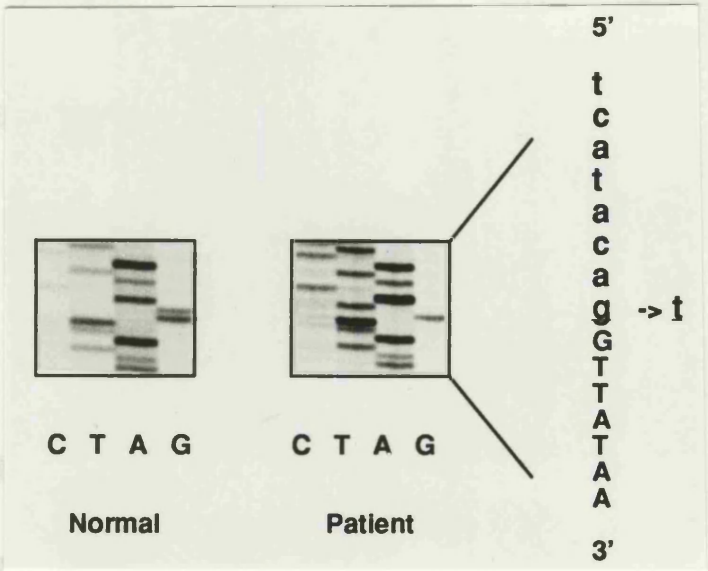
Family 26: 1010 del 19 (Hemizygote)



Deletion is between gene nucleotides 10998-11007 and cDNA Nos. 1010-1029 (Exon 7)

Appendix, Section A

Family 28: IVS3-1, t (Hemizygote)

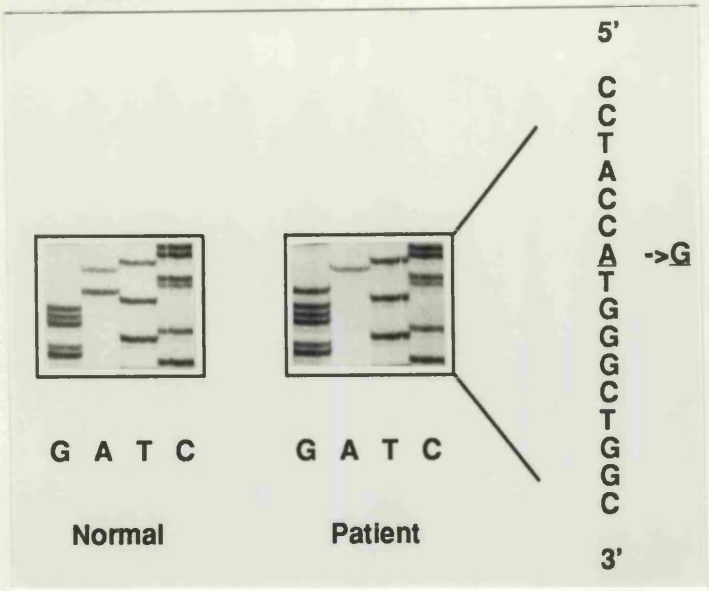


Base substitution is at gene nucleotide 8321,
ag -> at, in the 3' splice site of intron 3

Family 2: M 42 V (Hemizygote)

Methionine 42 -> Valine
(Non-polar -> Non-polar)

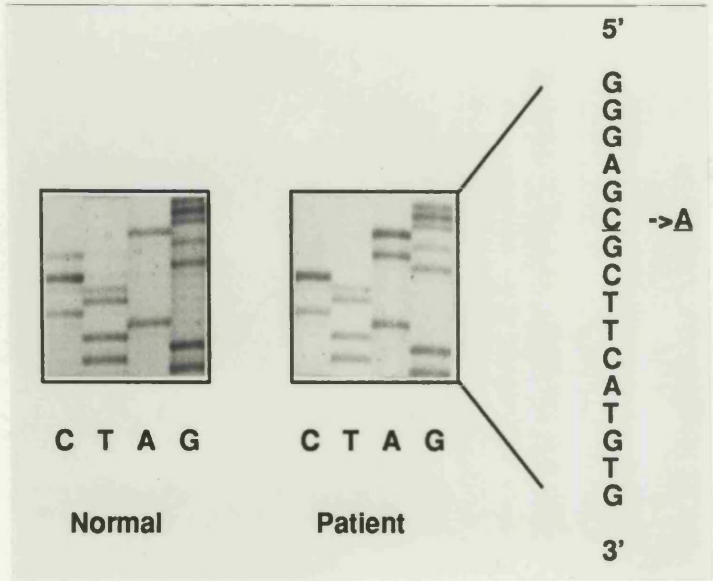
Gene No. 1303
cDNA No. 124 (Exon 1)



Family 3: R 49 S: (Hemizygote)

Arginine 49 -> Serine
(Polar, +ve -> Polar)

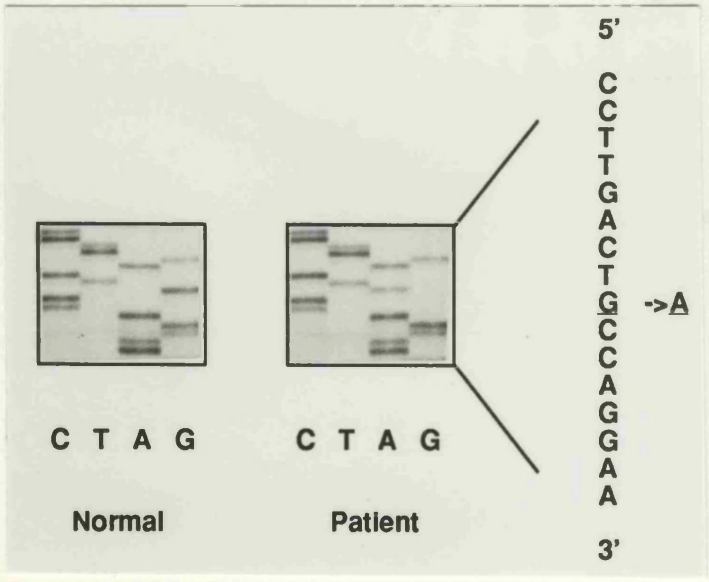
Gene No. 1324
cDNA No. 145 (Exon 1)



Family 5: C 56 Y (Hemizygote)

Cysteine 56 -> Tyrosine
(Polar -> Polar)

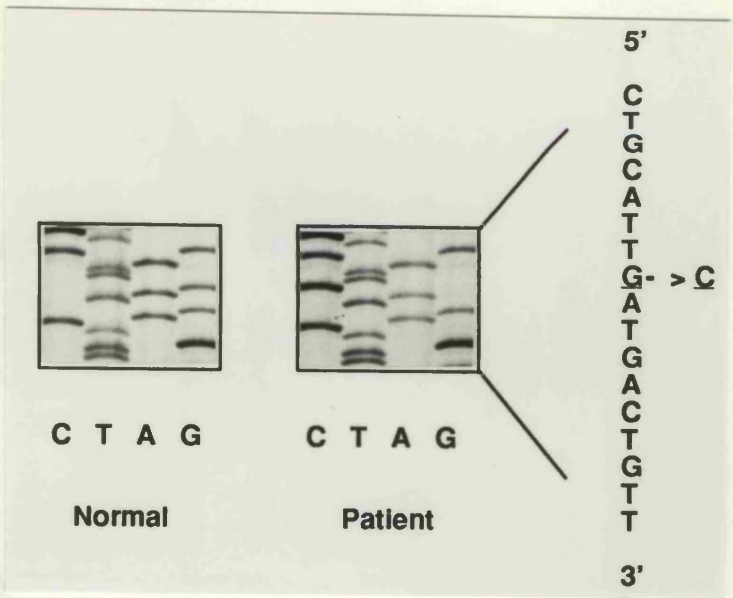
Gene No. 1346
cDNA No. 167 (Exon 1)



Family 6: D 92 H (Hemizygote)

Aspartate 92 -> Histidine
(Polar, +ve -> Polar, +ve)

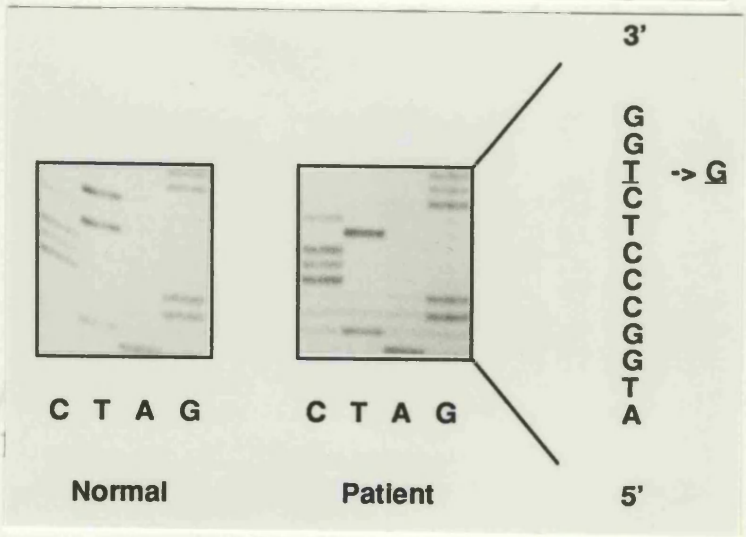
Gene No. 5173
cDNA No. 274 (Exon 2)



Family 11: W 287 G (Hemizygote)

Tryptophan 287 -> Glycine
(Non-polar -> Polar)

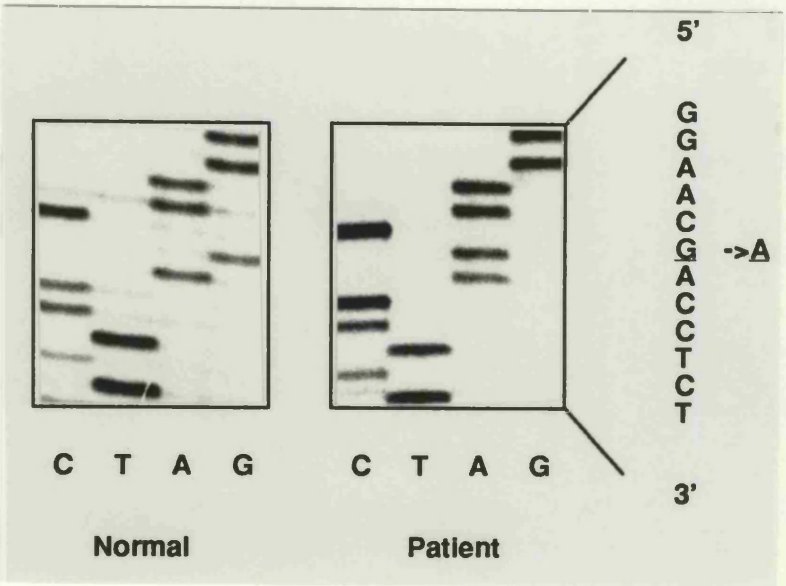
Gene No. 10567
cDNA No. 859 (Exon 6)



Family 15: R 342 Q (Hemizygote)

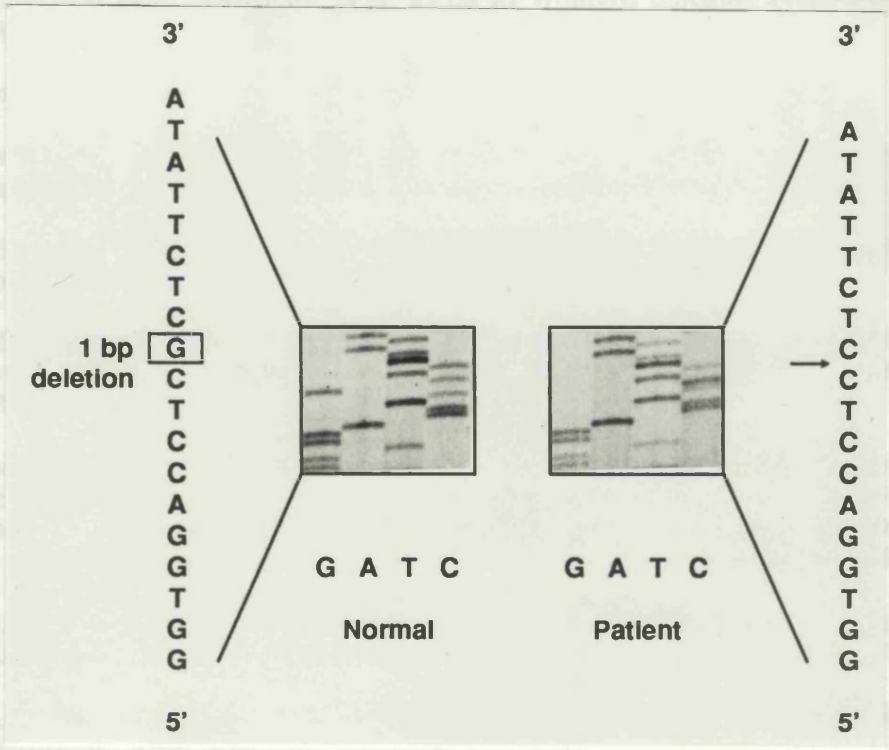
Arginine 342-> Glutamine
(Polar, +ve -> Polar)

Gene No. 11003
cDNA No. 1025 (Exon 7)



Appendix, Section B

Family 27: 1087 del 1 (Hemizygote)



1 bp deletion alters amino acids 363-389 and creates a premature stop codon at 390

Deletion is at gene nucleotide 11066 and cDNA No. 1088 (Exon 7)

REFERENCES

Aerts,JMFG; Schram,AW; Strijland,A; van Weeley,S; Jonsson,LMV; Tager,JM; Sorrell,SH; Ginns,EI; Barranger,JA; Murray,GJ (1988): Glucocerebrosidase, a lysosomal enzyme that does not undergo oligosaccharide phosphorylation. *Biochim. Biophys. Acta* 964, 303.

Allen,RC; Belmont,JB (1992): Dinucleotide repeat polymorphism at the DXS178 locus. *Hum. Mol. Gen.* 1, 216.

Anderson,RA; Byrum,RS; Coates,PM; Sando,GN (1994): Mutations at the lysosomal acid cholesterol ester hydrolase gene locus in Wolman disease. *Proc. Natn. Acad. Sci. USA* 91 (7), 2718.

Anderson,W (1898): A case of angiokeratoma. *Brit. J. Derm.* 10, 113.

Antonarakis,SE; Kazazian,HH Jr; Orkin,SH (1985): DNA polymorphism and molecular pathology of the human globin gene clusters. *Hum. Genet.* 69, 1.

Arai,K; Shimaya,A; Hiratani,N; Ohkuma,S (1993): Purification and characterization of lysosomal H+-ATPase. *J. Biol. Chem.* 268 (8), 5649.

Aran,JM; Gottesman,MM; Pastan,I (1994): Drug-selected coexpression of human glucocerebrosidase and P-glycoprotein using a bicistronic vector. *Proc. Natn. Acad. Sci. USA* 91, 3176.

Arredondo-Vega,FX; Santistiban,I; Kelly,S; Schlossman,CM; Umetsu,DT; Hershfield,MS (1987): Correct splicing despite mutation of the invariant first nucleotide of a 5' splice site: a possible basis for disparate clinical phenotypes in siblings with adenosine deaminase deficiency. *Am. J. Hum. Genet.* 54, 820.

Ashwell,G; Morell,AG (1974): The role of surface carbohydrates in the hepatic recognition and transport of circulating glycoproteins. *Adv. Enzymol.* 41, 99.

Aubert,JP; Biserte,G; Locheuux-Lefebvre,MH (1976): Carbohydrate-peptide linkage in glycoproteins. *Arch. Biochem. Biophys.* 175, 410.

Avila,JL; Convit,J; Velazquez-Avila,G (1973): Normal alpha-galactosidase activity and urinary-sediment glycosphingolipid levels in two obligate heterozygotes. *Brit. J. Derm.* 89, 149.

Avramopoulos,D; Chakravarti,A; Antonarakis,SE (1993): DNA polymorphisms in the 3' untranslated region of genes on human chromosome 21. *Genomics* 15, 98.

Bach,G; Rosenmann,E; Karni,A; Cohen,T (1982): Pseudodeficiency of alpha-galactosidase A. *Clin. Genet.* 21, 59.

Bakker,E; Veenema,H; Den Dunnen,JT; Van Broeckhoven,C; Grootscholten,PM; Bonten,EJ; Van Ommen,GJB; et al. (1989): Germinal mosaicism increases the recurrence risk for 'new' Duchenne muscular dystrophy mutations. *J. Med. Genet.* 26, 553.

Baranski,TJ; Faust,PL; Kornfeld,S (1990): Generation of a lysosomal-enzyme targeting signal in the secretory protein pepsinogen. *Cell* 63, 281.

Barik,S (1993): Site-directed mutagenesis by double polymerase chain reaction. In: Methods in Molecular Biology. Vol. 15: PCR protocols: current methods and applications. (Ed: White,BA) Humana Press Inc., Totawa, NJ, 277-286.

Barranger,JA; Rice,E (1993): An overview of Gaucher disease. *Gaucher Clinical Perspectives* 1 (1), 1.

Barrett,AJ; Heath,MF (1977): Lysosomal enzymes. In: Lysosomes, a Laboratory Handbook. 2nd ed. (Ed: Dingle,JT) North Holland, Amsterdam, 19-146.

Barton,NW; Brady,RO; Dambrosia,JM; Di Bisceglie,AM; Doppelt,SH; Hill,SC; Mankin,HJ; Murray,GJ; Parker,RI; Argoff,CE (1991): Replacement therapy for inherited enzyme deficiency-macrophage-targeted glucocerebrosidase for Gaucher's disease. *N. Eng. J. Med.* 324, 1464.

Ben-Yoseph,Y; Pack,BA; Mitchell,DA; Elwell,DG; Potior,M; Melancon,SB; Nadler,HL (1986): Characterization of the mutant N-acetylglucosaminyl-phosphotransferase in I-cell disease and pseudo-Hurler polydystrophy. Complementation analysis and kinetic studies. *Enzymology* 35, 106.

Bernstein,HS; Bishop,DF; Astrin,KH; Kornreich,R; Eng,CM; Sakuraba,H; Desnick,RJ (1989): Fabry disease: six gene rearrangements and an exonic point mutation in the alpha-galactosidase gene. *J. Clin. Invest.* 83, 1390.

Beutler,E; Kuhl,W (1973): Absence of cross-reactive antigen in Fabry disease. *N. Eng. J. Med.* 289, 694.

Beutler,E; Westwood,B; Dale,GL (1981): The effect of phlebotomy as a treatment of Fabry disease. *Biochem. Med.* 30, 363.

Bielicki,J; Freeman,C; Clements,PR; Hopwood,JJ (1990): Human liver iduronate-2-sulphatase. *Biochem. J* 271, 75.

Bird,A (1986): CpG-rich islands and the function of DNA methylation. *Nature* 321, 209.

Birnboim,HC; Doly,J (1979): A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucl. Acids Res.* 7, 1513.

Bishop,DF; Sweeley,CC (1978): Properties and comparisons with tissue alpha-galactosidases. *Biochim. Biophys. Acta* 525, 399.

Bishop,DF; Desnick,RJ (1981a): Affinity purification of alpha-galactosidase A from human spleen, placenta, and plasma with elimination of pyrogen contamination. *J. Biol. Chem.* 256 (3), 1307.

Bishop,DF; Grabowski,GA; Desnick,RJ (1981b): Fabry disease: an asymptomatic hemizygote with significant residual alpha-galactosidase A activity. *Am. J. Hum. Genet.* 33, 71A.

Bishop,DF; Kornreich,R; Desnick,RJ (1988): Structural organisation of the human alpha-galactosidase A gene: further evidence for the absence of a 3' untranslated region. *Proc. Natn. Acad. Sci. USA* 85, 3903.

Bishop,DF; Kornreich,R; Eng,CM; Ioannou,YA; Fitzmaurice,TF; Desnick,RJ (1991): Human alpha-galactosidase characterisation and eukaryotic expression of the full length cDNA and structural organisation of the gene. In: Lipid storage disorders - biological and medical aspects. (Eds: Salvayre,R; Douste-Blazy,L; Gatt,S) Plenum press, New York, 809-822.

Bohlen,P; Stein,S; Dairman,W; Udenfriend,S (1973): Fluorometric assay of proteins in the nanogram range. *Arch. Biochem. Biophys.* 155, 213.

Bottema,CDK; Bottema,MJ; Ketterling,RP; Yoon,H-S; Janco,RL; Phillips,JA; Sommer,SS (1991): Why does the human factor IX have a G + C content of 40%. *Am. J. Hum. Genet.* 49, 839.

Bou-Gharios,G; Abraham,D; Olsen,I (1993): Lysosomal storage diseases: mechanisms of enzyme replacement therapy. *Histochem. J.* 25, 593.

Brady,RO; Gal,AE; Bradley,RM; Martensson,E; Warshaw,AL; Laster,L (1967): Enzymatic defect in Fabry's disease: ceramide trihexosidase deficiency. *N. Eng. J. Med.* 276, 1163.

Brady,RO; Uhlendorf,BW; Jacobson,CB (1971): Fabry's disease: antenatal detection. *Science* 172, 174.

Braulke,T; Hille,A; Huttner,WB; Hasilik,WB; von Figura,K (1987): Sulfated oligosaccharides in human lysosomal enzymes. *Biochem. Biophys. Res. Commun.* 143, 178.

Braun,SE; Aronvich,EL; Anderson,RA; Crotty,PL; McIvor,RS; Whitley,CB (1993): Metabolic correction and cross-correction of mucopolysaccharidosis type II (Hunter syndrome by retroviral-mediated gene transfer and expression of human iduronate-2-sulfatase. *Proc. Natn. Acad. Sci. USA* 90, 11830.

Breathnach,R; Chambon,P (1981): Organisation and expression of eucaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50, 349.

Brown,MS; Kovanen,PT; Goldstein,JL (1981): Regulation of plasma cholesterol by lipoprotein receptors. *Science* 212, 628.

Burnside,J; Schneider,DL (1982): Characterization of the membrane proteins of the rat liver lysosomes-composition, enzyme activities and turnover. *Biochem. J* 204, 525.

Cariello,NF; Skopek,TR (1993): Mutational analysis using denaturing gradient gel electrophoresis and PCR. *Mutation Research* 288, 103.

Chen,HC; Tsai,JH; Lai,YH; Guh,JY (1990): Renal changes in heterozygous Fabry's disease - a family study. *Am. J. Kidney Dis.* 15, 180.

Chen,YQ; Rafi,MA; de Gata,G; Wenger,DA (1993): Cloning and expression of cDNA encoding human galactocerebrosidase, the enzyme deficient in globoid cell leukodystrophy. *Hum. Mol. Gen.* 2, 1841.

Cheng,J; Fogel-Petrovic,M; Maquat,LE (1990): Translation to near the distal end of the penultimate exon is required for normal levels of spliced triosephosphate isomerase mRNA. *Mol. Cell. Biol* 10 (10), 5215.

Chiang,H-L; Terlecky,SR; Plant,CP; Dice,JF (1989): A role for a 70kDa heat shock protein in lysosomal degradation of intracellular proteins. *Science* 246, 382.

Chou,H-F; Vadgama,J; Jonas,AJ (1992): Lysosomal transport of small molecules. *Biochem. Med. Metab. Biol.* 48, 179.

Chowdhury,JR; Grossman, M; Gupta, S; Chowdhury, NR; Baker, JR; Wilson, JM (1992): Long-term improvement of hypercholesterolemia after ex vivo therapy in LDLR-deficient rabbits. *Science* 254, 1802.

Chu,G; Hayakawa,H; Berg,P (1987): Electroporation for the efficient transfection of mammalian cells. *Nucl. Acids Res.* 15 (3), 1311.

Clarke,JTR; Knaack,J; Crawhall,JC; Wolfe,LS (1971): Ceramide trihexosidosis (Fabry's disease) without skin lesions. *N. Eng. J. Med.* 284, 233.

Cohen,SN; Chang,ACY; Hsu,L (1972): Non-chromosomal antibiotic resistance in bacteria: genetic transformation of Eschericia coli by R factor DNA. *Proc. Natn. Acad. Sci. USA* 69, 2110.

Columbi,A; Kostyal,A; Brachner,R; Gloor,F; Mazzi,R; Tholen,H (1967): Angiokeratoma corporis diffusum-Fabry's disease. *Helv. Med. Acta* 34, 67.

Cooper,DN; Youssoufian,H (1988): The CpG dinucleotide and human genetic disease. *Hum. Genet.* 78, 151.

Cooper,DN; Krawczak,M (1991): Mechanisms of insertional mutagenesis in human genes causing genetic disease. *Hum. Genet.* 87, 409.

Coppola,G; Yan,Y; Hantzopoulos,P; Segura,E; Stroh,JG; Calhoun,DH (1994): Characterization of glycosylated and catalytically active recombinant human alpha-galactosidase A using a baculovirus vector. *Gene* 144, 197.

Cotton,RGH; Rodrigues,NR; Campbell,RD (1988): Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine osmium tetroxide and its application to the study of mutations. *Proc. Natn. Acad. Sci. USA* 85, 4397.

Coulondre,C; Miller,JH; Farabaugh,PJ; Gilbert,W (1978): Molecular basis of base substitution hotspots in Escherichia coli. *Nature* 274, 775.

Cuozzo,JW; Sahagian,GG (1994): Lysine is a common determinant for mannose phosphorylation of lysosomal proteins. *J. Biol. Chem.* 269, 14490.

Dahms,N; Lobel,P; Kornfeld,S (1989): Mannose 6-phosphate receptors and lysosomal enzyme targeting. *J. Biol. Chem.* 264, 12115.

Davies,JP; Winchester,BG; Malcolm,S (1993): Mutation analysis in patients with the typical form of Anderson-Fabry disease. *Hum. Mol. Gen.* 7, 1051.

Dawson,G; Sweeley,CC (1970): In vivo studies on glycosphingolipid metabolism in porcine blood. *J. Biol. Chem.* 245, 410.

Dean,KJ; Sweeley,CC (1979): Studies on human liver alpha-galactosidases. *J. Biol. Chem.* 254, 9994.

de Duve,C; Pressman,BC; Gianetto,R; Wattiaux,R; Appelmans,F (1955): Intracellular distribution patterns of enzymes in rat liver tissue. *Biochem. J* **60**, 604.

Desnick,RJ (1983): Treatment of inherited metabolic diseases: current status and prospects. In: Genetic diseases: diagnosis and treatment. (Ed: Dietz,AA) American Association for Clinical Chemistry, Washington DC., 183.

Desnick,RJ; Allen,KY; Simmons,RL; Woods,JE; Anderson,CF; Najarian,JS; Krivit,W (1972): Correction of enzymatic deficiencies by renal transplantation: Fabry's disease. *Surgery* **72**, 203.

Desnick,RJ; Allen,KY; Desnick,SJ; Raman,MK; Bernlohr,RW; Krivit,W (1973): Fabry disease: enzymatic diagnosis of hemizygotes and heterozygotes. *J. Lab. Clin. Med.* **81**, 157.

Desnick,RJ; Dean,KJ; Grabowski,G; Bishop,DF; Sweeley,CC (1979): Enzyme therapy in Fabry's disease: differential in vivo plasma clearance and metabolic effectiveness of plasma and splenic alpha-galactosidase A isoenzymes. *Proc. Natn. Acad. Sci. USA* **76**, 5326.

Desnick,RJ; Grabowski,GA (1981): Advances in the treatment of inherited metabolic diseases. *Adv. Hum. Genet.* **11**, 281.

Desnick,RJ; Bernstein,HS; Astrin,KH; Bishop,DF (1987): Fabry disease: molecular diagnosis of hemizygotes and heterozygotes. *Enzyme* **38**, 54.

Desnick,RJ; Bishop,DF (1989): Fabry disease: alpha-galactosidase A deficiency; Schindler disease: alpha-N-acetylgalactosaminidase deficiency. Chap. 70. In: The metabolic basis of inherited disease. 6th ed. (Eds: Scriver,CR; Beaudet,AL; Sly,WS; Valle,D) McGraw Hill, New York, 1751-1796.

Desnick,RJ; Iounnou,YA; Eng,CM (1994): Alpha-galactosidase A deficiency: Fabry disease. In: The Metabolic Basis of Inherited Diseases. 7th ed. (Eds: Scriver,CR; Beaudet,AL; Sly,WS; Valle,D) McGraw Hill, New York.

de Weers,M; Mensick,RGJ; Kenter,M; Schuurman,RKB (1992): Three dinucleotide repeat polymorphisms at the DXS178 locus. *Hum. Mol. Gen.* **1**, 653.

Dice,JF (1992): Selective degradation of cytosolic proteins by lysosomes. *Ann. N.Y. Acad. Sci.* **674**, 58.

Dietz,HC; Valle,D; Francomano,CA; Kendzior,RJ; Pyeritz,RE; Cutting,GR (1993): The skipping of constitutive exons in vivo induced by nonsense mutations. *Science* **259**, 680.

Dietz,HC; Kendzior,RJ,Jr (1994): Maintenance of an open reading frame as an additional level of scrutiny during splice site selection. *Nature Genetics* **8**, 183.

Dulbecco,R; Freeman,G (1959): Plaque formation by the polyoma virus. *Virology* **8**, 396.

Elleder,M; Bradova,V; Smid,F; Budesinsky,M; Harzer,K; Kustermann-Kuhn,B; Ledvinova,J; Belohlavek; Krai,V; Dorazilova,V (1990a): Cardiocyte storage and hypertrophy as a sole manifestation of Fabry's disease. Report on a case simulating hypertrophic non-obstructive cardiomyopathy. *Virchows. Arch. [A]* **417**, 449.

Elleder,M; Ledvinova,J; Vosmik,F; Zeman,J; Stejskal,D; Lageron,A (1990b): An atypical ultrastructural pattern in Fabry's disease: a study on its nature and incidence in 7 cases. *Ultrastruct. Pathol.* 14, 267.

Emerman,M; Temin,HM (1984): Genes with promoters in retrovirus vectors can be independently suppressed by an epigenetic mechanism. *Cell* 39, 459.

Eng,CM; Resnick-Silverman,LA; Niehaus,DJ; Astrin,KH; Desnick,RJ (1993): Nature and frequency of mutations in the alpha-galactosidase A gene that cause Fabry disease. *Am. J. Hum. Genet.* 53, 1186.

Eng,CM; Desnick,RJ (1994a): Molecular basis of Fabry disease: mutations and polymorphisms in the human alpha-galactosidase A gene. *Hum. Mutat.* 3, 103.

Eng,CM; Niehaus,DJ; Enriquez,A; Burgert,TS; Ludman,MD; Desnick,RJ (1994b): Fabry disease: twenty-three mutations including sense and antisense CpG alterations and identification of a deletional hot-spot in the alpha-galactosidase A gene. *Hum. Mol. Gen.* 3, 1795.

Ezaki,J; Wolfe,LS; Higuti,T; Ishidoh,K; Kominami,E (1995): Specific delay of degradation of mitochondrial ATP synthase subunit c in Late infantile Neuronal ceroid lipofuscinosis (Batten disease). *J. Neurochem.* 64, 733.

Fabry,J (1898): Ein Beitrag Zur Kenntnis der perpura haemorrhagica nodularis (Purpura papulosa hemorrhagica Hebrae). *Arch. Dermatol. Syph.* 43, 187.

Fan,E; Levin,DB; Glickman,BW; Logan,DM (1993): Limitations in the use of SSCP analysis. *Mutation Research* 288, 85.

Farragina,T; Churg,J; Grisham,E; Strauss,L; Prado,A; Bishop,DF; Schuchman,E; Desnick,RJ (1981): Light and electron microscopic histochemistry of Fabry's disease. *Am. J. Pathol.* 103, 247.

Flomen,RH; Green,EP; Green,PM; Bentley,DR; Giannelli,F (1993): Determination of the organisation of the coding sequences within the iduronate sulphate sulphatase (IDS) gene. *Hum. Mol. Gen.* 2, 5.

Flynn,DM; Lake,BD; Boothby,CB; Young,EP (1972): Gut lesions in Fabry's disease without a rash. *Arch. Dis. Child.* 47, 26.

Fodde,R; Losekoot,M (1994): Mutation detection by denaturing gradient gel electrophoresis (DGGE). *Hum. Mutat.* 3, 83.

Forrest,SM; Dahl,HH; Howells,DW; Dianzani,I; Cotton,RG (1991): Mutation detection in phenylketonuria by using chemical cleavage mismatch: importance of using probes from both the normal and patient samples. *Am. J. Hum. Genet.* 49, 175.

Freshney,RI (Ed.) (1994): Culture of animal cells: a manual of basic technique. 3rd ed. Wiley Liss, New York. 486 pages.

Friedmann,T (1992): A brief history of gene therapy. *Nature Genetics* 2, 93.

Froissart,R; Blond,J-L; Maire,I; Guibaud,P; Hopwood,JJ; Mathieu,M; Bozon,D (1993): Hunter syndrome: gene deletions and rearrangements. *Hum. Mutat.* 2, 138.

Fu,YH; Kuhl,DPA; Pizzuti,A; Pieretti, M; Sutcliffe, JS; Richards, S; Verkerk, AJMH; Holden, JJA; Fenwick, RG; Warren, ST; Oostra, BA; Nelson, DL; Caskey, CT (1991): Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 67, 1047.

Fukada,I; Ogawa,K (1992): Alternatively-spliced p53 mRNA in the FAA-HTC1 rat hepatoma cell line without the splice site mutations. *Cell Structure and Function* 17, 427.

Fukada,M (1991): Lysosomal membrane glycoproteins. *J. Biol. Chem.* 266, 21327.

Fukushima,H; Dewet,JR; O'Brien,JS (1985): Molecular cloning of a cDNA for alpha-L-fucosidase. *Proc. Natn. Acad. Sci. USA* 82, 1262.

Furbish,FS; Steer,CJ; Krett,NL; Barranger,JA (1981): Uptake and distribution of placental glucocerebrosidase in rat hepatic cells and effects of sequential deglycosylation. *Biochim. Biophys. Acta* 673, 425.

Furst,W; Sandhoff,K (1992): Activator proteins and topology of lysosomal sphingolipid catabolism. *Biochim. Biophys. Acta* 1126, 1.

Galjart,NJ; Gillemans,N; Harris,A; van der Horst,GTJ; Verheijen,FW; Galjaard,H; d'Azzo,A (1988): Expression of cDNA encoding the human "protective protein" associated with lysosomal beta-galactosidase and neuraminidase: homology to yeast proteases. *Cell* 54, 755.

Gemignani,F; Pietrini,Y; Tagliavini,F; Lechi,A; Neri,TM; Asinari,A; Savi,M (1979): Fabry's disease with familial lymphedema of the lower limbs. *Eur. Neurol.* 18, 84.

Gitschier,J (1988): Maternal duplication associated with gene deletion in sporadic hemophilia. *Am. J. Hum. Genet.* 43, 274.

Gluzman,Y (1981): SV40-transformed Simian cells support the replication of early SV40 mutants. *Cell* 23, 175.

Goldstein,JL; Brown,MS (1977): The low density lipoprotein pathway and its relation to atheroschlerosis. *Annu. Rev. Biochem.* 46, 897.

Gottschalk,S; Waheed,A; Schmidt,B; Laidler,P; von Figura,K (1989): Sequential processing of lysosomal acid phosphatase by a cytoplasmic thiol proteinase and a lysosomal aspartyl proteinase. *EMBO J.* 8, 3215.

Grabowski,GA (1993): Genetic aspects of Gaucher disease. *Gaucher Clinical Perspectives* 1, 5.

Grabowski,GA; Barton,NW; Pastores,G; Dambrosia,JM; Banerjee,TK; McKee,MA; Parker,C; Schiffmann,R; Hill,SC; Brady,RO (1995): Enzyme therapy in type 1 Gaucher disease: comparative efficacy of mannose-terminated glucocerebrosidase from natural and recombinant sources. *Ann. Intern. Med.* 122, 33.

Grace,ME; Berg,A; He,G; Goldberg,L; Horowitz,M; Grabowski,GA (1991): Gaucher disease: heterologous expression of two alleles associated with neuronopathic phenotypes. *Am. J. Hum. Genet.* 49, 646.

Grompe,M (1993): The rapid detection of unknown mutations in nucleic acids. *Nature Genetics* 5, 111.

Grosshans,E (1986): A propos de la revue generale la maladie de Fabry. *Ann. Dermatol. Venereol.* 113, 277.

Haldane,JBS (1935): The rate of spontaneous mutation for a human gene. *J. Genet.* 31, 317.

Hall,JG (1988): Review and hypotheses: germline mosaicism: observations related to clinical genetics. *Am. J. Hum. Genet.* 43, 355.

Ham,RG (1963): An improved nutrient solution for diploid Chinese hamster and human cell lines. *Exp. Cell Res.* 29, 515.

Hamers,MN; Wise,D; Ejiofor,A; Strijland,D; Robinson,D; Tager,JM (1979): Relationship between biochemical and clinical features in an English Anderson-Fabry family. *Acta Med. Scand.* 206, 5.

Harter,C; Mellman,I (1992): Transport of the lysosomal membrane glycoprotein lgp120 (lgpA) to lysosomes does not require appearance on the plasma membrane. *J. Cell Biol.* 117, 311.

Hartl,DL (1971): Recurrence risk for germinal mosaics. *Am. J. Hum. Genet.* 23, 124.

Hashimoto,K; Lieberman,P; Lamkin,N Jr (1976): Angiokeratoma corporis diffusum (Fabry disease). *Arch. Dermatol.* 112, 1416.

Hasilik,A (1992): The early and late processing of lysosomal enzymes: proteolysis and compartmentation. *Experientia* 48, 130.

Hasilik,A; Waheed,A; von Figura,K (1981): Enzymatic phosphorylation of lysosomal enzymes in the presence of UDP-N-acetylglucosamine. Absence of the activity in I-cell fibroblasts. *Biochem. Biophys. Res. Commun.* 98, 761.

Hayashi,K (1991): PCR-SSCP: a simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods and Applications* 1, 34.

Hendriks,RW; Mensink,EJBM; Kraakman,MEM; Thomson,A; Schuurman,RKB (1989): Evidence for male X-chromosomeal mosaicism in X-linked agammaglobulinaemia. *Hum. Genet.* 83, 267.

Hers,HG (1965): Inborn lysosomal storage diseases. *Gastroenterology* 48, 625.

Hobbs,JR (1990): Displacement bone marrow transplantation for some inborn errors. *J. Inher. Metab. Dis.* 13, 572.

Hoefsloot,LH; Hoogeveen-Westerveld,M; Reuser,AJJ; Oostra,BA (1990): Characterisation of the human alpha-glucosidase gene. *Biochem. J* 272, 493.

Hoekstra,D; Kok,JW (1992): Trafficking of glycosphingolipids in eukaryotic cells; sorting and recycling of lipids. *Biochim. Biophys. Acta* 1113, 227.

Holtzmann,E (1989): Lysosomes. Plenum Press, New York. 439 pages.

Honing,S; Hunziker,W (1995): Cytoplasmic determinants involved in direct lysosomal sorting, endocytosis and basolateral targeting of rat lgp120 (lamp-I) in MDCK cells. *J. Cell Biol.* 128, 321.

Hopwood,JJ; Bunge,S; Morris,CP; Wilson,PJ; Steglich,M; Beck,M; Schwinger,E; Gal,A (1993a): Molecular basis of mucopolysaccharidosis type II: mutations in the Iduronate-2-sulphatase gene. *Hum. Mutat.* 2, 435.

Hopwood,JJ; Vellodi,A; Scott,HS; Morris,CP; Litjens,T; Clements,PR; Brooks,DA; Cooper,A; Wraith,JE (1993b): Long-term clinical progress in bone-marrow transplanted mucopolysaccharidosis type I patients with a defined genotype. *J. Inher. Metab. Dis.* 16, 1024.

Horowitz,M; Wilder,S; Horowitz,Z; Reiner,O; Gelbart,T; Beutler,E (1989): The human glucocerebrosidase gene and pseudogene: structure and evolution. *Genomics* 4, 87.

Hozumi,I; Nishizawa,M; Ariga,T; Miyatake,T (1990): Biochemical and clinical analysis of accumulated glycolipids in symptomatic heterozygotes of angiokeratoma corporis diffusum (Fabry's disease) in comparison with hemizygotes. *J. Lipid Res.* 31, 335.

Hubbes,M; Callahan,J; Gravel,R; Mahuran,D (1989): The amino-terminal sequences in the pro-alpha and beta polypeptides of human lysosomal beta-hexosaminidase A and B are retained in the mature isozymes. *FEBS Lett.* 249, 316.

Hultman,T; Stahl,S; Hornes,E; Uhlen,M (1989): Direct solid-phase sequencing of genomic and plasmid DNA using magnetic beads as a solid support. *Nucl. Acids Res.* 17, 4937.

Huttner,WB; Baeuerle,PA (1988): Protein sulfation on tyrosines. *Modern Cell Biol.* 6, 97.

Idriss,J-M; Jonas,AJ (1991): Vitamin B12 transport by rat liver lysosomal membrane vesicles. *J. Biol. Chem.* 266, 9438.

Iounnou,YA; Bishop,DF; Desnick,RJ (1992): Overexpression of human alpha-galactosidase A results in its intracellular aggregation, crystallization in lysosomes, and selective secretion. *J. Cell Biol.* 119, 1137.

Ishii,S; Sakuraba,H; Shimmoto,M; Minamikawa-Tachino,R; Suzuki,T; Suzuki,Y (1991): Fabry disease: detection of 13bp deletion in alpha-galactosidase A gene and its application to gene diagnosis of heterozygotes. *Ann. Neurol.* 29, 560.

Ishii,S; Sakaruba,H; Suzuki,Y (1992): Point mutations in the upstream region of the alpha-galactosidase A gene exon 6 in an atypical variant of Fabry disease. *Hum. Genet.* 89, 29.

Ishii,S; Kase,R; Sakaruba,H; Suzuki,Y (1993): Characterization of a mutant alpha-galactosidase gene product for the late-onset cardiac form of Fabry disease. *Biochem. Biophys. Res. Commun.* 197, 1585.

Jackson,IJ (1991): A reappraisal of non-consensus mRNA splice sites. *Nucl. Acids Res.* 19, 3795.

Jeanpierre,M (1992): Germinal mosaicism and risk calculation in X-linked diseases. *Am. J. Hum. Genet.* 50, 960.

Johnson,DL; del Monte,MA; Cotlier,E; Desnick,RJ (1975): Fabry disease: diagnosis of hemizygotes and heterozygotes by alpha-galactosidase activity in tears. *Clin. Chim. Acta* 63, 81.

Johnston,AW; Warland,BJ; Weller,SDV (1966): Genetic aspects of angiokeratoma corporis diffusum. *Ann. Hum. Genet.* 30, 25.

Jongkind,JF; Verkerk,A; Niermeijer,MF (1983): Detection of Fabry's disease heterozygotes by enzyme analysis in single fibroblasts after cell sorting. *Clin. Genet.* 23, 261.

Kaplan,A; Achord,DT; Sly,WS (1977): Phosphohexosyl components of a lysosomal enzyme are recognised by pinocytosis receptors on human fibroblasts. *Proc. Natn. Acad. Sci. USA* 74, 2026.

Kassenbrock,CK; Garcia,PD; Walter,P; Kelly,RB (1988): Heavy chain binding protein recognizes aberrant polypeptide translocated in vitro. *Nature* 333, 90.

Kaufman,RJ; Davies,MV; Pathak,V; Hershey,JWB (1989): The phosphorylation efficiency of eukaryotic initiation factor II alters translational efficiency of specific messenger-RNAs. *Mol. Cell. Biol* 9 (3), 946.

Keen,J; Lester,D; Inglehearn,C; Curtis,A; Bhattacharya,S (1991): Rapid detection of single base mismatches as heteroduplexes on hydrolink gels. *Trends Genet.* 7, 5.

Kint,JA (1970): Fabry's disease: alpha-galactosidase deficiency. *Science* 167, 1268.

Kint,JA; Carton,D (1973): . In: Lysosomes and Storage Diseases. (Eds: Hers,HF; Van Hoof,F) Academic, New York, 347.

Kirkilionis,AJ; Riddell,DC; Spence,MW; Fenwick,RG (1991): Fabry disease in a large Nova Scotia kindred: carrier detection using leucocyte alpha-galactosidase activity and an NcoI polymorphism detected by an alpha-galactosidase cDNA clone. *J. Med. Genet.* 28, 232.

Kishimoto,Y; Hiraiwi,M; O'Brien,JS (1992): Saposins: structure, function, distribution, and molecular genetics. *J. Lipid Res.* 33, 1255.

Klausner,RD; Sitia,R (1990): Protein degradation in the endoplasmic reticulum. *Cell* 62, 611.

Klima,H; Tanaka,A; Schnabel,D; Nakano,T; Schroder,M; Suzuki,K; Sandhoff,K (1991): Characterization of the full length cDNAs and the gene coding for the human GM2 activator protein. *FEBS Lett.* 289, 260.

Kobayashi,T; Kira,J; Shinnoh,N; Goto,I; Kuroiwa,Y (1985): Fabry's disease with partially deficient hydrolysis of ceramide trihexoside. *J. Neurol. Sci.* 67, 179.

Koide,T; Ishiura,M; Iwai,K; Inoue,M; Kaneda,Y; Okada,Y; Uchida,T (1990): A case of Fabry's disease in a patient with no alpha-galactosidase A activity caused by a single amino acid substitution of Pro-40 by Ser. *FEBS Lett.* 259, 353.

Kornfeld,R; Kornfeld,S (1985): Assembly of asparagine-linked oligosaccharides. *Annu. Rev. Biochem.* 54, 631.

Kornfeld,S (1986): Trafficking of lysosomal enzymes in normal and disease states. *J. Clin. Invest.* 77, 1.

Kornfeld,S (1989): Lysosomal enzyme targeting. *Biochem. Soc. Trans.* 18, 367.

Kornfeld,S (1992): Structure and function of the mannose 6-phosphate/insulin like growth factor II receptors. *Annu. Rev. Biochem.* **61**, 307.

Kornfeld,S; Mellman,I (1989): The biogenesis of lysosomes. *Annu. Rev. Cell Biol.* **5**, 483.

Kornreich,R; Desnick,RJ; Bishop,DF (1989): Nucleotide sequence of the human alpha-galactosidase A gene. *Nucl. Acids Res.* **17**, 3301.

Kornreich,R; Bishop,DF; Desnick,RJ (1990): Alpha-galactosidase A gene rearrangements causing Fabry disease. Identification of short direct repeats at breakpoints in an Alu-rich gene. *J. Biol. Chem.* **265**, 9319.

Kornreich,R; Astrin,KH; Desnick,RJ (1992): Amplification of human polymorphic sites in the X-chromosomal region q21.33 to q24: DXS17, DXS87, DXS287 and alpha-galactosidase A. *Genomics* **13**, 70.

Kornreich,R; Desnick,RJ (1993): Fabry disease: detection of gene rearrangements in the human alpha-galactosidase A gene by multiplex PCR amplification. *Hum. Mutat.* **2**, 108.

Koster,A; von Figura,K; Pohlmann,R (1994): Mistargeting of lysosomal enzymes in Mr 46000 mannose 6-phosphate receptor-deficient mice is compensated by carbohydrate-specific endocytotic receptors. *Eur. J. Biochem.* **224**, 685.

Kozak,M (1986): Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* **44**, 283.

Kozak,M (1987): An analysis of 5'-non-coding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**, 8125.

Krall,WJ; Challita,PM; Perlmutter,LS; Skelton,DC; Kohn,DB (1994): Cells expressing human glucocerebrosidase from a retroviral vector repopulate macrophages and central nervous system microglia after murine bone marrow transplantation. *Blood* **9**, 2737.

Kramer,W; Thormann,J; Mueller,K; Frenzel,H (1984): Progressive cardiac involvement by Fabry's disease despite successful renal allotransplantation. *Int. J. Cardiol.* **7**, 72.

Krawczak,M; Cooper,DN (1991): Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. *Hum. Genet.* **86**, 425.

Krawczak,M; Reiss,J; Cooper,DN (1992): The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes. *Hum. Genet.* **90**, 41.

Krentler,C; Pohlmann,R; Hasilik,A; von Figura,K (1986): Lysosomal membrane proteins do not bind to the mannose-6-phosphate-specific receptors. *Biol. Chem. Hoppe-Seyler* **367**, 141.

Kretz,KA; Cripe,D; Carson,GS; Fukushima,H; O'Brien,JS (1992): Structure and sequence of the human alpha-L-fucosidase gene and pseudogene. *Genomics* **12**, 279.

Kreysing,J; von Figura,K; Gieselman,V (1990): Structure of the arylsulphatase A gene. *Eur. J. Biochem.* **191**, 627.

Krivit,W; Vance,DE; Desnick,R; Whitecar,JP; Sweeley,CC (1968): Red cell physiology in Fabry's disease. *J. Lab. Clin. Med.* 12, 906.

Krivit,W; Pierpont,ME; Ayaz,K; Tsai,M; Ramsay,NK; Kersey,JH; Weisdorf,S; Sibley,R; Snover,D; McGovern,MM; Schwartz,MF; Desnick,RJ (1984): Bone marrow transplantation in the Maroteaux-Lamy syndrome (Mucopolysaccharidosis type VI) - biochemical and clinical status 24 months after transplantation. *N. Eng. J. Med.* 311, 1606.

Kusiak,JW; Quirk,JM; Brady,RO; Mook,GE (1978): Purification and properties of the two major isozymes of alpha-galactosidase from human placenta. *J. Biol. Chem.* 253, 184.

Laemmli,UK (1970): Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680.

Lazzarino,DA; Gabel,CA (1988): Biosynthesis of the mannose-6-phosphate recognition marker in transport-impaired mouse lymphoma cells - demonstration of a 2-step phosphorylation. *J. Biol. Chem.* 263, 10118.

Lazzarino,DA; Gabel,CA (1989): Mannose processing is an important determinant in the assembly of phosphorylated high mannose-types of oligosaccharides. *J. Biol. Chem.* 264, 1015.

Le Donne,NC; Fairley,JL; Sweeley,CC (1983): Biosynthesis of alpha-galactosidase A in cultured Chang liver cells. *Arch. Biochem. Biophys.* 244, 186.

Lemansky,P; Bishop,DF; Desnick,RJ; Hasilik,A; von Figura,K (1987): Synthesis and processing of alpha-galactosidase A in human fibroblasts. Evidence for different mutations in Fabry disease. *J. Biol. Chem.* 262, 2062.

Levade,T; Giordano,F; Maret,A; Marguery,MC; Bazex,J; Salvayre,R (1991): Case report: different phenotypic expression in female monozygotic twins. *J. Inher. Metab. Dis.* 14, 105.

Levitt,RC (1991): Polymorphisms in the 3' untranslated region of eukaryotic genes. *Genomics* 11, 484.

Lewin,B (1994): Messenger RNA is the template. In: Genes V. Oxford University Press, Oxford, 253-278.

Li,S-C; Kihara,H; Serizawa,S; Li,Y-T; Fluharty,A; Mayes,JS; Shapiro,LJ (1985): Activator protein required for the enzymatic hydrolysis of cerebroside sulfate. Deficiency in urine of patients affected with cerebroside sulfatase activator deficiency and identity with activators for the enzymatic hydrolysis of GM1 ganglioside and globotriaosylceramide. *J. Biol. Chem.* 260, 1867.

Little,LE; Lau,MMH; Quon,DVK; Fowler,AV; Neufeld,EF (1988): Proteolytic processing of the alpha-chain of the lysosomal enzyme beta-hexosaminidase, in normal fibroblasts. *J. Biol. Chem.* 263, 4288.

Lovelock,JE; Bishop,MWH (1959): Prevention of freezing damage to living cells by dimethyl sulphoxide. *Nature* 183, 1394.

Lyon,MF (1961): Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* 190, 372.

MacDermot,KD; Morgan,SH; Cheshire,JK; Wilson,TM (1987): Anderson Fabry disease, a close linkage with highly polymorphic DNA markers DXS17, DXS87, DXS88. *Hum. Genet.* 77, 263.

Maddalena,A; Sosnoski,DM; Berry,GT; Nussbaum,RL (1988): Mosaicism for an intragenic deletion in a boy with mild ornithine transcarbamylase deficiency. *N. Eng. J. Med.* 319, 999.

Madsen,KM; Hasholt,L; Sorensen,SA; Largerstrom Fermer,M; Dahl,N (1995): Two novel mutations (L32P) and (G85N) among five different missense mutations in six Danish families with Fabry disease. *Hum. Mutat.* 5, 277.

Mahuran,DJ (1990): Characterisation of human placental beta-hexosaminidase I2: proteolytic processing intermediates of hexosaminidase A. *J. Biol. Chem.* 265, 6794.

Makalowski,W; Mitchell,GA; Labuda,D (1994): Alu sequences in the coding regions of mRNA: a source of protein variability. *Trends Genet.* 10, 188.

Mandel,J-L; Monaco,AP; Nelson,DL; Schlessinger,D; Willard,H (1992): Genome analysis and the human X chromosome. *Science* 258, 103.

Markello,TC; Bernardini,ME; Gahl,WA (1993): Improved renal function in children with cystinosis treated with cysteamine. *N. Eng. J. Med.* 328, 1157.

Matsuo,M; Masumura,T; Nishio,H; Nakajima,T; Kitoh,Y; Takumi,T; Koja,J; Nakamura,H (1991): Exon skipping during splicing of dystrophin mRNA precursor due to an intraexon deletion in the gene of Duchenne muscular dystrophy Kobe. *J. Clin. Invest.* 87, 2127.

Matthews,BW (1993): Structural and genetic analysis of protein stability. *Annu. Rev. Biochem.* 62, 139.

Mayes,JS; Schreerer,JB; Sifers,RN; Donaldson,ML (1980): Differential assay for lysosomal alpha-galactosidases in human tissues and its application to Fabry's disease. *Clin. Chim. Acta* 112, 247.

McDowell,GA; Cowan,TM; Blitzer,MG; Greene,CL (1993): Intrafamilial variability in Hurler syndrome and Sanfilippo syndrome type A: implications for evaluation of new therapies. *Am. J. Med. Genet.* 47, 1092.

McInnes,B; Potier,M; Wakamatsu,N; Melancon,SB; Klavins,MH; Tsuji,S; Mahuran,DJ (1992): An unusual splicing mutation in the HEXB gene is associated with dramatically different phenotypes in patients from different racial backgrounds. *J. Clin. Invest.* 90, 306.

Mellman,I (1987): Molecular sorting during endocytosis. *Kidney International* Suppl. 23, 184.

Migeon,BR (1994): X-chromosome inactivation: molecular mechanisms and genetic consequences. *Trends Genet.* 10, 230.

Miller,RD; Hoffman,JW; Powell,PP; Kyle,JW; Shipley,JM; Bachinsky,DR; Sly,WS (1990): Cloning and characterization of the human beta-glucuronidase gene. *Genomics* 7, 280.

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

Mitchell,ML; Erickson,RP; Schmid,D; Hieber,V; Poznanski,AK; Hicks,SP (1981): Mannosidosis: two brothers with different degrees of disease severity. *Clin. Genet.* 20, 191.

Modaressi,S; Rupp,K; von Figura,K; Peters,C (1993): Structure of the human arylsulphatase B gene. *Biol. Chem. Hoppe-Seyler* 374, 327.

Mononen,I; Fisher,KJ; Kaartinen,V; Aronson,NN,Jr (1993): Aspartylglycosaminuria: protein chemistry and molecular biology of the most common lysosomal storage disorder of glycoprotein degradation. *FASEB J.* 7, 1247.

Moore,GE; Gerner,RE; Franklin,HA (1967): Culture of normal human leukocytes. *J. Am. Med. Assoc.* 199, 519.

Morgan,SH; Cheshire,JK; Wilson,TM; MacDermot,K; d'A. Crawfurd,M (1987): Anderson-Fabry disease - family linkage studies using two polymorphic X-linked DNA probes. *Pediatr. Nephrol.* 1, 536.

Morgan,SH; Rudge,P; Smith,SJM; Bronstein,AM; Kendall,BE; Holly,E; Young,EP; d'A Crawfurd,M; Bannister,R (1990): The neurological complications of Anderson-Fabry disease (alpha-galactosidase A deficiency) - investigation of the symptomatic and presymptomatic patients. *Q. J. Med.* 75, 491.

Mori,N; Yokota,J; Akiyama,T; Sameshima,Y; Okamoto,A; Mizoguchi,H; Toyoshima,K; Sugimura,T; Terada,M (1990): Variable mutations of the RB gene in small-cell lung carcinoma. *Oncogene* 5, 1713.

Morreau,H; Bonten,E; Zhou,XY; d'Azzo,A (1991): Organisation of the gene encoding lysosomal beta-galactosidase. *DNA Cell Biol.* 10, 495.

Morris,CP; Guo,X-H; Apostolou,S; Hopwood,JJ; Scott,HS (1994): Morquio A syndrome: cloning, sequence, and structure of the human N-acetylgalactosamine 6-sulphatase (GALNS) gene. *Genomics* 22, 652.

Moskowitz,SM; Tieu,PT; Neufeld,EF (1993): Mutation in Scheie syndrome (MPS IS): a G->A transition creates new splice site in intron 5 of one IDUA allele. *Hum. Mutat.* 2, 141.

Mosnier,JF; Degott,C; Bedrossian,J; Molas,G; Degos,F; Pruna,A; Potet,F (1991): Recurrence of Fabry's disease in a renal allograft eleven years after successful renal transplantation. *Transplantation* 51, 759.

Moullier,P; Merechal,V; Danos,O; Heard,JH (1993): Continuous systemic secretion of a lysosomal enzyme by genetically modified mouse skin fibroblasts. *Transplantation* 56, 427.

Murali,R; Iounnou,YA; Desnick,RJ; Burnett,RM (1994): Crystallization and preliminary X-ray analysis of human alpha-galactosidase A complex. *J. Mol. Biol.* 239, 578.

Myers,RM; Larin,Z; Maniatis,T (1985): Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. *Science* 230, 1242.

Nagao,Y; Nakashima,H; Fukuhara,Y; Shimmoto,M; Oshima,A; Ikara,Y; Mori,Y; Sakaruba,H; Suzuki,Y (1991): Hypertrophic cardiomyopathy in late-onset variant of Fabry disease with high residual activity of alpha-galactosidase A. *Clin. Genet.* 39, 233.

Naylor,JA; Green,PM; Rizza,CR; Giannelli,F (1992): Factor VIII gene explains all cases of haemophilia A. *Lancet* 340, 1066.

Nebes,VL; Schmidt,MC (1994): Human lysosomal alpha-mannosidase: isolation and nucleotide sequence of the full-length cDNA. *Biochem. Biophys. Res. Commun.* 200, 239.

Neote,K; Bapat,B; Dumbrilleross,A; Troxel,C; Schuster,SM; Mahuran,DJ; Gravel,RA (1988): Characterization of the human HexB gene encoding lysosomal beta-hexosaminidase. *Genomics* 3, 277.

Neufeld,EF (1991): Lysosomal storage diseases. *Annu. Rev. Biochem.* 60, 257.

Neurmann,JR; Schaefer-Ridder,M; Wang,Y; Hofschneider,PH (1982): Gene transfer into mouse lyoma cells by electroporation in high electric fields. *EMBO J.* 1, 841.

Nilsen,TW (1994): RNA-RNA interactions in the spliceosome: unraveling the ties that bind. *Cell* 78, 1.

Novikoff,AB; Beaufay,H; de Duve,C (1956): Electron microscopy of lysosome rich fractions from rat liver. *J. Biophys. Biochem. Cytol.* 2, 179.

Ogata,S; Fukuda,M (1994): Lysosomal targeting of LIMP-II membrane glycoprotein requires a novel Leu-Ile motif at a particular position in the cytoplasmic tail. *J. Biol. Chem.* 269, 5210.

Ogawa,K; Sugamata,K; Funamoto,N; Abe,T; Sato,T; Nagashima,K; Ohkawa,S (1990): Restricted accumulation of globotriaosylceramide in the hearts of atypical cases of Fabry's disease. *Hum. Pathol.* 21, 1067.

Oishi,M; Cosloy,SD (1972): The genetic and biochemical basis of the transformation of Escherichia coli K12. 49, 1568.

Orita,M; Suzuki,Y; Sekiya,T; Hayashi,K (1989): Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* 5, 874.

Oshima,A; Yoshida,K; Itoh,K; Kase,R; Sakaruba,H; Suzuki,Y (1994): Intracellular processing and maturation of mutant gene products in hereditary beta-galactosidase deficiency (beta-galactosialidosis). *Hum. Genet.* 93, 109.

Owada,M; Neufeld,EF (1982): Is there a mechanism for introducing acid hydrolases into liver lysosomes that is independent of mannose 6-phosphate recognition? Evidence from I-cell disease. *Biochem. Biophys. Res. Commun.* 105, 814.

Park,H; Fisher,KJ; Aronson,NN (1991): Genomic structure of human lysosomal glycosylasparaginase. *FEBS Lett.* 288, 168.

Peltola,M; Chiatayat,D; Peltonen,L; Jalanko,A (1994): Characterization of a point mutation in aspartylglucosaminidase gene: evidence for a read-through of a translational stop codon. *Hum. Mol. Gen.* 3, 2237.

Peltz,SW; Brown,AH; Jackobson,A (1993): mRNA destabilisation triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes Dev.* **7**, 1737.

Peters,C; von Figura,K (1994): Biogenesis of lysosomal membranes. *FEBS Lett.* **346**, 108.

Philippart,M; Franklin,SS; Gordon,A (1972): Reversal of an inborn sphingolipidosis (Fabry's disease) by kidney transplantation. *Ann. Intern. Med.* **77**, 195.

Pisoni,RL; Thoene,JG (1991): The transport systems of mammalian lysosomes. *Biochim. Biophys. Acta* **1071**, 351.

Ploos van Amstel,JK; Jansen,RPM; de Jong,JGN; Hamel,BCJ; Wevers,RA (1994): Six novel mutations in the alpha-galactosidase A gene in families with Fabry disease. *Hum. Mol. Gen.* **3(3)**, 503.

Poduslo,SE; Dean,M; Kolch,U; O'Brien,JS (1991): Detecting high resolution polymorphisms in human coding loci by combining PCR and single-strand conformation polymorphisms (SSCP) analysis. *Am. J. Hum. Genet.* **49**, 106.

Price,NC; Stevens,L (1988): The structure of enzymes. In: Fundamentals of enzymology. (:) Oxford University Press, Oxford, 90-93.

Proia,RL; Soravia,E (1987): Organization of the gene encoding the human beta-hexosaminidase alpha-chain. *J. Biol. Chem.* **262**, 5677.

Puck,JM; Pepper,AE; Bedard,PM; Laframboise,R (1995): Female germline mosaicism as the origin of a unique IL-2 receptor gamma chain mutation causing X-linked severe combined immunodeficiency. *J. Clin. Invest.* **95**, 859.

Pyeritz,RE; Ullman,MD; Moser,AB; Braine,HG; Moser,HW (1980): Plasma exchange removes glycosphingolipid in Fabry disease. *Am. J. Med. Genet.* **7**, 301.

Radcliffe,KW; Evans,BA (1990): Anderson-Fabry disease (Angiokeratoma corporis diffusum universale). *Genitourin Med.* **66**, 399.

Rietra,PJGM; Molenaar,JL; Hamers,MN; Tager,JM; Borst,P (1974): Investigation of the alpha-galactosidase deficiency in Fabry's disease using antibodies against the purified enzyme. *Eur. J. Biochem.* **46**, 89.

Rietra,PJGM; Brouwer-Kelder,EM; DeGroot,WP; Tager,JM (1976): The use of biochemical parameters for the detection of carriers of Fabry's disease. *J. Molec. Med.* **1**, 237.

Robertson,DA; Freeman,C; Nelson,PV; Morris,CP; Hopwood,JJ (1988): Human glucosamine-6-sulfatase cDNA reveals homology with steroid sulfatase. *Biochem. Biophys. Res. Commun.* **157**, 218.

Romeo,G; Migeon,BR (1970): Genetic inactivation of the alpha-galactosidase locus in carriers of Fabry's disease. *Science* **170**, 180.

Romeo,G; Urso,M; Pisacane,A; Blum,E; DeFalco,A; Ruffilli,A (1975): Residual activity of alpha-galactosidase A in Fabry's disease. *Biochem. Genet.* **13**, 615.

Rommerskirch,W; Fluharty,AL; Peters,C; von Figura,K; Gieselmann,V (1991): Restoration of arylsulfatase A activity in Human metachromatic leukodystrophy fibroblasts via retroviral vector-mediated gene transfer. *Biochem. J* 280, 459.

Ropers,H; Wienker,TF; Grimm,T; Schroetter,K; Bender,K (1977): Evidence for preferential X-chromosome inactivation in a family with Fabry disease. *Am. J. Hum. Genet.* 29, 361.

Rorman,EG; Scheikner,V; Grabowski,GA (1992): Structure and evolution of the human prosaposin cloned gene. *Genomics* 13, 312.

Rosenfeld,MG; Kreibich,G; Popov,D; Kato,K; Sabatini,DD (1982): Biosynthesis of lysosomal hydrolases: their synthesis in bound polysomes and the role of co- and post-translational processing in determining their subcellular distribution. *J. Biol. Chem.* 93, 135.

Rothman,JE (1989): Polypeptide chain binding proteins - catalysts of protein folding and related processes in cells. *Cell* 59, 591.

Rowe,JW; Gilliam,JI; Warthin,TA (1974): Intestinal manifestations of Fabry's disease. *Ann. Intern. Med.* 81, 628.

Ruano,G; Kidd,KK (1991): Coupled amplification and sequencing of genomic DNA. *Proc. Natn. Acad. Sci. USA* 88, 2815.

Saad,FA; Halliger,B; Muller,CR; Roberts,RG; Danieli,GA (1994): Single base substitutions are detected by double strand conformation analysis. *Nucl. Acids Res.* 22, 4352.

Saiki,RK; Scharf,S; Faloona,F; Mullis,KB; Horn,GT; Erlich,HA; Arnheim,N (1985): Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230, 1350.

Sakuraba,H; Oshima,A; Fukuhara,Y; Shimmoto,M; Nagao,Y; Bishop,DF; Desnick,RJ; Suzuki,Y (1990): Identification of point mutations in the alpha-galactosidase A gene in classical and atypical hemizygotes with Fabry disease. *Am. J. Hum. Genet.* 47, 784.

Sambrook,J; Fritsch,EF; Maniatis,T (Eds.) (1989): Molecular cloning, a laboratory manual. 2nd ed. Vol. 1-3. Cold Spring Harbor Press, New York.

Sandhoff,K; van Echten,G; Schroder,M; Schnabel,D; Suzuki,K (1992): Metabolism of glycolipids: the role of glycolipid-binding proteins in the function and pathobiochemistry of lysosomes. *Biochem. Soc. Trans.* 20, 695.

Sandhoff,K; Klein,A (1994): Intracellular trafficking of glycosphingolipids: role of sphingolipid activator proteins in the topology of endocytosis and lysosomal digestion. *FEBS Lett.* 346, 103.

Sanger,F; Nicklen,S; Coulson,AR (1977): DNA sequencing with chain terminating inhibitors. *Proc. Natn. Acad. Sci. USA* 74, 5463.

Sarkar,G; Sommer,SS (1990): The "megaprimer" method of site-directed mutagenesis. *BioTechniques* 8, 404.

von Scheidt,W; Eng,CM; Fitzmaurice,TF; Erdmann,E; Hubner,G; Olsen,GJ; Christomanou,H; Kandolf,R; Bishop,DF; Desnick,RJ (1991): An atypical variant of Fabry's disease with manifestations confined to the myocardium. *N. Eng. J. Med.* 324, 395.

Schlinder,D; Kanzaki,T; Desnick,RJ (1990): A method for the rapid detection of urinary glycopeptides in alpha-N-acetylgalactosaminidase deficiency and other lysosomal storage diseases. *Clin. Chim. Acta* 190, 81.

Schuchman,EH; Levran,O; Pereira,LV; Desnick,RJ (1992): Structural organization and complete nucleotide sequence of the gene encoding human acid sphingomyelinase. *Genomics* 12, 197.

Scott,HS; Guo,X-H; Hopwood,JJ; Morris,CP (1992): Structure and sequence of the human alpha-L-iduronidase gene. *Genomics* 13, 1311.

Scriver,CR; Beaudet,AL; Sly,W; Valle,D (Eds.) (1989): The metabolic basis of inherited disease. 6th ed. McGraw-Hill, New York. 3006 pages.

Seglen,PO; Bohley,P (1992): Autophagy and other vacuolar protein degradation mechanisms. *Experientia* 48, 158.

Shapiro,MB; Senapathy,P (1987): RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucl. Acids Res.* 15, 7155.

Sharp,PA (1994): Split genes and RNA splicing. *Cell* **77**, 805.

Sheffield,VC; Cox,DR; Lerman,LS; Myers,RM (1989): Attachment of a 40 base-pair G + C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. *Proc. Natn. Acad. Sci. USA* 86, 232.

Sheffield,VC; Beck,JS; Kwitek,AE; Sandstrom,DW; Stone,EM (1993): The sensitivity of single-strand conformation polymorphism analysis for the detection of single base substitutions. *Genomics* 16, 325.

Sheth,KJ; Good,TA; Murphy,JV (1981): Heterozygote detection in Fabry disease utilizing multiple enzyme activities. *Am. J. Med. Genet.* 10, 141.

Shipley,JM; Miller,RD; Wu,BM; Grubb,JH; Christensen,SG; Kyle,JW; Sly,WS (1991): Analysis of the 5' flanking region of the human beta-glucuronidase gene. *Genomics* 10, 1009.

Simon,SM; Blobel,G (1991): A protein-conducting channel in the endoplasmic-reticulum. *Cell* 65, 371.

Smooker,PM; Cotton,RGH (1993): The use of chemical reagents in the detection of DNA mutations. *Mutation Research* 288, 65.

Sonderfeld-Fresco,S; Proia,RL (1989): Analysis of the glycosylation and phosphorylation of the lysosomal enzyme, beta-hexosaminidase B, by site-directed mutagenesis. *J. Biol. Chem.* 264, 7692.

Southern,E (1975): Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* 98, 503.

Spence,MW; MacKinnon,KE; Burgess,JK; d'Entremont,DM; Belitsky,P; Lannon,SG; MacDonald,AS (1976): Failure to correct the metabolic defect by renal allotransplantation in Fabry's disease. *Ann. Intern. Med.* 84, 13.

Spence,MW; Goldbloom,AL; Burgess,JK; d'Entremont,D; Ripley,BA; Weldon,KL (1977): Heterozygote detection in angiokeratoma corporis diffusum (Anderson-Fabry disease). *J. Med. Genet.* 14, 91.

Spinardi,L; Mazars,R; Theillet,C (1991): Protocols for improved detection of point mutations by SSCP. *Nucl. Acids Res.* 19, 4009.

Steingrimsdottir,H; Rowley,G; Dorado,G; Cole,J; Lehmann,AR (1992): Mutations which alter splicing in the human hypoxanthine-guanine phosphoribosyltransferase gene. *Nucl. Acids Res.* 20, 1201.

Stevens,RL; Fluharty,AL; Kihara,H; Kaback,MM; Shapiro,LJ; Marsh,B; Sandhoff,K; Fischer,G (1981): Cerebroside sulfatase activator deficiency induced metachromatic leukodystrophy. *Am. J. Hum. Genet.* 33, 900.

Sweatman,AK; Bradley,LAD; Lovering,RC; O'Reilly,M-AJ; Levinsky,RJ; Kinnon,C (1994): Physical mapping in the region of the Bruton's tyrosine kinase and alpha-galactosidase A gene loci in proximal Xq22. *Hum. Genet.* 94, 624.

Sweeley,CC; Klinosky,B (1963): Fabry disease: Classification as a sphingolipidosis and partial characterization of a novel glycolipid. *J. Biol. Chem.* 238, 3148.

Taylor,SAM; Deugau,KV; Lillicrap,DP (1991): Somatic mosaicism and female-to-female transmission in a kindred with hemophilia B (factor IX deficiency). *Proc. Natn. Acad. Sci. USA* 88, 39.

Touraine,JL; Malik,MC; Perrot,H; Maire,I; Revillard,JP; Grosshans,E; Traeger,J (1979): Maladie de Fabry: deux maladies ameliores par la greffe de cellules de foie foetal. *Nouv. Press Med.* 8, 1499.

Urlaub,G; Mitchell,PJ; Ciudad,CJ; Chasin,LA (1989): Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. *Mol. Cell. Biol* 9, 2868.

van Beusechem,VW; Kukler, A; Einerhand, MVM; Bakx, TA; van der Eb, AJ; van Bekkum, DW; Valerio, D (1990): Expression of human adenosine deaminase in mice transplanted with hematopoietic stem cells infected with amphotrophic retroviruses. *J. Exp. Med.* 172, 729.

van der Meulen,MA; van der Meulen,MJP; te Meerman,GJ (1995): Recurrence risk for germinal mosaics revisited. *J. Med. Genet.* 32, 102.

van Essen,AJ; Abbs,S; Baiget,M; Bakker,E; Boileau,C; Broeckoven,C; Bushby,K; Clarke,A; Claustres,M; et al. (1992): Parental origin and germline mosaicism of deletions and duplications of the dystrophin gene: a European study. *Hum. Genet.* 88, 249.

Vetrie,D; Bobrow,M; Harris,A (1993): Construction of a 5.2-megabase physical map of the human X chromosome at Xq22 using pulsed-field gel electrophoresis and yeast artificial chromosomes. *Genomics* 15, 631.

Vetrie,D; Kendall,E; Coffrey,A; Hassock,S; Collins,J; Todd,C; Lehrach,H; Bobrow,M; Bentley,DR; Harris,A (1994): A 6.5 Mb yeast artificial chromosome contig incorporating 33 DNA markers on the human X chromosome at Xq22. *Genomics* 19, 42.

von Figura,K; Hasilik,A (1986): Lysosomal enzymes and their receptors. *Annu. Rev. Biochem.* 55, 167.

von Heijne,G (1986): A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* 14, 4683.

Walkley,SU; Thrall,MA; Dobrenis,K; Huang,M; March,PA; Siegel,DA; Wurzelmann,S (1994): Bone marrow transplantation corrects the enzyme defect in neurons of the central nervous system in a lysosomal storage disease. *Proc. Natn. Acad. Sci. USA* 91, 2970.

Walter,P; Gilmore,R; Blobel,G (1984): Protein translocation across the endoplasmic reticulum. *Cell* 38, 5.

Wang,AM; Bishop,DF; Desnick,RJ (1990): Human alpha-N-acetylgalactosaminidase - molecular cloning, nucleotide sequence and expression of full-length cDNA. *J. Biol. Chem.* 265, 21859.

Wang,AM; Desnick,RJ (1991): Structural organization and complete sequence of the human alpha-N-acetylgalactosaminidase gene - homology with the alpha-galactosidase A gene provides evidence for evolution from a common ancestral gene. *Genomics* 10, 133.

Wang,L-M; Weber,DK; Johnson,T; Sakaguchi,AY (1988): Supercoil sequencing using unpurified templates produced by rapid boiling. *BioTechniques* 6, 839.

Wareham,KA; Lyon,MF; Glenister,PH; Williams,ED (1987): Age related reactivation of an X-linked gene. *Nature* 327, 725.

Watson,MEE (1984): Compilation of published signal sequences. *Nucl. Acids Res.* 12, 5145.

Weatherall,DJ (1991): The new genetics and clinical practice. 3rd ed. Oxford University Press, Oxford.

Weaver,DT; DePamphilis,ML (1982): Specific sequences in native DNA that arrest synthesis by DNA polymerase alpha. *J. Biol. Chem.* 257, 2075.

Weitz,G; Proia,RL (1992): Analysis of the glycosylation and phosphorylation of the alpha-subunit of the lysosomal enzyme, beta-hexosaminidase A, by site-directed mutagenesis. *J. Biol. Chem.* 267, 10039.

Wherret,JR; Hakomori,S (1973): Characterization of a blood group B glycosphingolipid accumulating in the pancreas of a patient with Fabry disease. *J. Biol. Chem.* 218, 3046.

White,MB; Carvalho,M; Derse,D; O'Brien,SJ; Dean,M (1992): Detecting single base substitutions as heteroduplex polymorphisms. *Genomics* 12, 301.

Willemsen,R; van Dongen,JM; Ginns,EI; Sips,HJ; Schram,AW; Tager,JM; Barranger,JA; Reuser,AJJ (1987): Ultrastructural localization of glucocerebrosidase in cultured Gaucher's disease fibroblasts by immunocytochemistry. *Neurology* **234**, 41.

Williams,MA; Fukuda,M (1990): Accumulation of membrane glycoproteins in lysosomes require a tyrosine residue at a particular position in the cytoplasmic tail. *J. Cell Biol.* **111**, 955.

Wilson,PJ; Morris,CP; Anson,DS; Occhiodoro,T; Bielicki,J; Clements,PR; Hopwood,JJ (1990): Hunter syndrome: isolation of an iduronate-2-sulphatase cDNA clone and analysis of patient DNA. *Proc. Natn. Acad. Sci. USA* **87**, 8531.

Wilson,PJ; Meaney,CA; Hopwood,JJ; Morris,CP (1993): Sequence of the human iduronate 2-sulphatase IDS gene. *Genomics* **17**, 773.

Winchester,Bryan (1992): The Lysosome: structure, organization and function. In: Fundamentals of medical cell biology, Membranology and subcellular organelles. Vol. 4. JAI Press Inc. 305-361.

Winchester,B; Young,E; Geddes,S; Hurst,J; Middleton-Price,H; Williams,N; Webb,N; Habel,A; Malcolm,S (1992): Hunter disease due to non-random inactivation of the X-chromosome: a consequence of twinning. *Am. J. Med. Genet.* **44**, 834.

Yamamoto,K; Tanouchi,J; Kagiya,T; Iwai,K; Fujii,K; Ozaki,H; Matsuyama,T; Mishima,M; Hori,M; Kitabatake,A; Kamada,T (1989): A case of Fabry's disease manifested only with myocardial lesion. *Shinzo (Tokyo)* **3**, 326.

Yang,M; Allen,H; DiCiocco,RA (1992): A mutation generating a stop codon in the alpha-L-fucosidase gene of a fucosidosis patient. *Biochem. Biophys. Res. Commun.* **189**, 1063.

Yang,M; Allen,H; DiCioccio,RA (1993): Pedigree analysis of alpha-L-fucosidase gene mutations in a fucosidosis family. *Biochim. Biophys. Acta* **1182**, 245.

Yap,EPH; McGee,JOD (1993): Nonisotopic discontinuous phase single-strand conformation polymorphism (DP-SSCP): genetic profiling of D-loop of human mitochondrial (mt) DNA. *Nucl. Acids Res.* **21**, 4155.

Yokoi,T; Shinoda,K; Ohno,I; Kato,K; Miyawaki,T; Taniguchi,N (1991): A 3' splice site consensus sequence mutation in the intron 3 of the alpha-galactosidase A gene in a patient with Fabry disease. *Jpn. J. Hum. Genet.* **36**, 245.

Yoshida,A (1983): The existance of atypical blood group galactosyltransferase which causes an expression of A2 character in A1B red blood cells. *Hum. Genet.* **35**, 1117.

Zhang,Z-X; Wakamatsu,N; Mules,EH; Thomas,GH; Gravel,RA (1994): Impact of premature stop codons on mRNA levels in infantile Sandhoff disease. *Hum. Mol. Gen.* **3**, 139.

Zlotogora,J (1987): Intrafamilial variability in lysosomal storage diseases. *Am. J. Med. Genet.* **27**, 633.

Zuhlsdorf,M; Imort,M; Hasilik,A; von Figura,K (1983): Molecular forms of beta-hexosaminidase and cathepsin D in serum and urine of healthy subjects and patients with elevated activity of lysosomal enzymes. *Biochem. J* **213**, 733.

J Med Genet 1993; 30: 658–663

Sequence variations in the first exon of α-galactosidase A

Joanna P Davies, Bryan G Winchester, Sue Malcolm

Abstract

The α-galactosidase A gene (GALA), which is deficient in males with Anderson-Fabry disease, is shown to be remarkably polymorphic in the 5′ untranslated region. GALA contains seven exons. The first exon contains 60 bp of 5′ untranslated sequence before the methionine initiation codon. Single strand conformation polymorphism (SSCP) screening has shown three polymorphic variants from the published sequence within the 60 base pairs. The sequence changes involved are C to T at −10, G to A at −12 (which removes an *Msp*I site), and G to A at −30 (which removes a *Sac*II site). The combined frequency of these is 10%. A further insertion-deletion polymorphism is detected by SSCP of a 400 bp fragment including exon 3. Both polymorphisms can be easily detected using small polyacrylamide gels and ethidium bromide staining. Nine of 20 women were informative for one of these polymorphisms and this simple SSCP analysis should be of great assistance in family studies of Anderson-Fabry disease. Such a high level of polymorphism has not been previously reported in the 5′ untranslated region of a human gene and is unusual in any such short stretch of DNA.

(*J Med Genet* 1993;30:658–63)

α-galactosidase (EC 3.2.1.22) is an X linked lysosomal hydrolase involved in the catabolism of glycoconjugates. It catalyses the hydrolysis of α-galactosidic linkages at the non-reducing terminals of naturally occurring glycosphingolipids and glycoproteins, including blood group B substance, and of synthetic substrates.[1] A detergent is required for the hydrolysis of glycosphingolipid substrates in vitro whereas in vivo this function is achieved by sphingolipid activator protein 1 (SAP-1, another lysosomal protein).[2] α-galactosidase is synthesised as a precursor which undergoes post-translational modification en route to lysosomes from its site of synthesis on the rough endoplasmic reticulum.[3] The mature native enzyme has a molecular weight of approximately 100 000 and is composed of two apparently identical subunits containing 7 to 15% carbohydrate.

In Fabry disease, a genetic deficiency of α-galactosidase leads to the excretion in the urine and the accumulation within lysosomes of a variety of tissues of glycosphingolipids containing terminal α-galactose residues.[4] The characteristic clinical symptoms in affected males include angiokeratoma, pain in the extremities, renal failure, and heart and cardiovascular disease. These can be attributed to the progressive accumulation of storage products within the vascular endothelium. Death most often occurs around the age of 40 years. The definitive diagnosis is the demonstration of a deficiency of α-galactosidase in leucocytes or plasma.[5] The range of α-galactosidase activity in heterozygous females is very wide, overlapping that in both affected males and normal controls. This makes enzymic detection of carriers very difficult. Most heterozygotes are asymptomatic but may show some clinical symptoms, such as corneal opacity or angiokeratoma in later life. In a few heterozygotes the symptoms have been reported to be as severe as in affected males.[6] This variability in biochemical and clinical expression owing to random X inactivation emphasises the need for molecular genetic tests for heterozygosity.

α-galactosidase A has been cloned[7] and the full length cDNA sequences reported.[8,9] The gene contains seven exons extending over 12 kb.[9] The complete nucleotide sequence has been reported.[10] α-galactosidase has a 60 base pair 5′ non-translated region, but is unusual in having no 3′ non-translated region. The polyadenylation signal is in the coding region 12nt from the termination codon, which is followed by the poly(A) tract. The α-galactosidase gene has been assigned to the long arm of the X chromosome (Xq22) using somatic cell hybrids and linkage analysis.[4]

Only two RFLPs have been reported for GALA. An *Nco*I polymorphism is located 3′ to the gene, approximately 10 kb downstream from the termination codon, with allele frequencies 0·87/0·13. A rare *Sac*I polymorphism with allele frequencies 0·92/0·08 is located in intron 4.[11,12] Closely linked polymorphic DNA sequences have also been used in carrier detection for Fabry disease, using markers around Xq22, such as DXS17 (a *Taq*I polymorphism) and DXS88 (*Bgl*II RFLP).[13] However, few specific single base changes within the gene have been identified and out of 12 point mutations reported,[14–20] only two are reported as non-disease causing. At exon 1, codon 8, an A to G change is silent in a triplet coding for leucine,[18] and in exon 2, codon 66 a Glu to Gln base change is non-causative for Fabry's disease[14] as shown by expression of this mutant in CD5 cells. The others are described as disease causing either because whole cDNA[14–18,20] or exon sequencing[19] has shown no other mutations or because expression studies show that the mutation causes a reduction in enzyme activity.[14,18] No specific nucleotide variation

Units of Molecular Genetics and Enzymology, Division of Biochemistry and Genetics, Institute of Child Health, 30 Guilford Street, London WC1N 1EH, UK.
J P Davies
B G Winchester
S Malcolm

Correspondence to Dr Malcolm.

Received 2 November 1992.
Revised version accepted 23 February 1993.

changes have been described in the 5′ untranslated region of the gene.

SSCP is a convenient method for screening stretches of DNA for sequence changes.[21] It is based on the principle that denaturation and rapid cooling, single stranded DNA molecules will adopt unique sequence dependent conformations which can be separated by electrophoresis in native polyacrylamide gels.

Materials and methods

RADIOLABELLED SSCP ANALYSIS OF EXONS 1 AND 3

DNA samples were prepared from anticoagulated whole blood (in EDTA) in a process involving the salting out of cellular proteins by dehydration and precipitation with saturated ammonium acetate.[22] Red blood cells were lysed by adding water to 5 to 10 ml of whole blood to give a final volume of 50 ml and centrifuging at 2300 rpm for 20 minutes. The nuclear pellet was then washed with 25 ml of 0·1% NP40 and centrifuged for a further 20 minutes. The nuclei were lysed and proteinase K digested as previously described. Then 1 ml of saturated ammonium acetate (9·6 mol/l) was added and the sample centrifuged for 20 minutes to precipitate protein. The DNA was precipitated from the supernatant by the addition of 2 volumes of absolute ethanol and was redissolved in 1 ml of 10 mmol/l Tris, 1 mmol/l EDTA solution.

Oligonucleotide primers were designed using the PRIMER program provided by the HGMP resource centre based at MRC Clinical Research Centre, Harrow, UK. The exon 1 primers were 5′GTCCCAGTTGCCAGAGAAAC3′ and 5′AAAGGGAAGGGAGTACCCAA3′, located from nucleotide position 1054–1073 and 1403–1384.[10] The exon 3 primers were 5′GATGATTTTGGGGGTTTGTG3′ and 5′GATTGGTTCTTTGGCTCAGC3′ located at 7123–7142 and 7522–7503. Primers were synthesised on an Applied Biosystems 381A DNA synthesiser and 5′-biotinylation was carried out during oligonucleotide synthesis using DMT-biotin-C-6-PA(CRB) according to the manufacturer's instructions.

PCR was conducted in a total volume of 50 µl and each reaction mixture contained genomic DNA, 40 pmol of each primer, 0·2 mmol/l dATP, ddGTP, dTTP, 0·02 mmol/l dCTP, 1 µCi of ^{32}P dCTP, in 1·5 mmol/l $MgCl_2$, 50 mmol/l KCl, 10 mmol/l Tris-HCl (pH 8·8), 1% Triton X-100. Mineral oil (50 µl) was added to each reaction mixture before an initial 94°C denaturing step, followed by the addition of 1 unit of *Taq* Polymerase (Promega) at 58°C. Amplification was carried out by 30 cycles of elongation at 72°C (one minute), denaturation at 94°C (30 seconds), and annealing at 58°C (30 seconds). The amplification was completed by a final 10 minute elongation step at 72°C. The products (350 and 400 base pairs for exons 1 and 3 respectively) were checked by running 5 µl of each product on a 2% agarose gel with ethidium bromide (1 mg/ml) staining.

A volume of 5 µl of PCR product was diluted by the addition of 40 µl of 0·1% SDS, 10 mmol/l EDTA solution and 2 µl of diluted product was mixed with 2 µl of 95% formamide, 0·05% xylene cyanol, 0·05% bromophenol blue, 20 mmol/l EDTA solution (Sequenase kit stop solution, USB). The samples were then denatured by heating at 94°C for three minutes and placed on ice before loading onto a 0·4 mm thick 30 × 40 cm, 6% polyacrylamide gel (Protogel: 30% acrylamide to 0·8% bisacrylamide, National Diagnostics), with 5% glycerol.

Electrophoresis was carried out in 0·09 mmol/l Tris-borate, 0·002 mmol/l EDTA (pH 8·3) buffer (TBE), at 340 V at room temperature for approximately 15 hours, until the xylene cyanol dye had migrated 31 cm. The gel was dried onto 3MM Whatman paper under vacuum at 80°C and exposed (without intensifying screens) to *x* ray film (Kodak, IBI Ltd) for one to five days at −70°C.

NON-RADIOLABELLED SSCP

PCR was conducted in a volume of 50 µl as described above but using unlabelled dNTPs (0·2 mmol/l of each). A volume of 15 µl of each product was mixed with 5 µl of loading buffer (15% sucrose, 0·05% xylene cyanol) before denaturing (three minutes at 94°C) and placing on ice.

The entire volume of each sample was loaded onto a 1 mm thick, 16 × 20 cm polyacrylamide gel and electrophoresis was carried out using the Protean II gel system (BioRad, UK) in TBE buffer. For exon 1 a 6% polyacrylamide gel (Protogel: 30% acrylamide, 0·8% bisacrylamide, National Diagnostics) with 5% glycerol was run at 100 V overnight. For exon 3, the gel was 12% polyacrylamide (Protogel), without glycerol run at 200 V overnight.

SEQUENCE ANALYSIS

PCR amplification of exon 1 was performed as previously described, except that 5 pmol of each primer was used in 100 µl total reaction mixture volume and unlabelled dNTPs were used (0·2 mmol/l of each). The 3′ primer was biotinylated at its 5′ end allowing production of single stranded DNA using magnetic streptavidin coated beads (Dynal UK). Sequencing was carried out as follows: 50 µl of PCR product was mixed with 30 µl of magnetic Dynal W280 streptavidin beads and incubated for five minutes allowing the double stranded DNA to bind via a streptavidin-biotin bond. The supernatant was removed while the PCR product was attracted to a magnet (Promega) and the beads (with DNA) were then washed twice with 100 µl of TES (10 mmol/l Tris-HCl, pH 8·0, 1 mmol/l EDTA, 0·1 mol/l NaCl). After a five minute denaturing step with 0·15 mol/l NaOH at room temperature, the non-biotinylated sense strand was removed and discarded, leaving the single stranded, antisense DNA bound to the bead and magnet. This was again washed twice with

TES before resuspending in water, to give a final volume of 7 µl.

The sequencing reaction was conducted using 1 pmol of the 5′ primer and 7 µl of the resuspended Dynal bead bound DNA, using the Sequenase II kit (USB), as described by the manufacturer. The Sequenase enzyme was diluted eight fold and the labelling mixture was diluted fifteen fold. The Sequenase elongation step was two minutes and the termination step four minutes.

The four termination mixtures were denatured at 85°C for three minutes and placed on ice until loading onto a 6% polyacrylamide (Accugel: 40% (19:1) acrylamide to bisacrylamide, National Diagnostics), 8·3 mol/l urea, denaturing gel. Electrophoresis was performed at 55 W for about two hours in TBE buffer using the BRL S2 sequencing gel apparatus.

The gel was fixed in 10% methanol and 10% acetic acid solution for 10 to 30 minutes before blotting and drying as described previously for the SSCP gels.

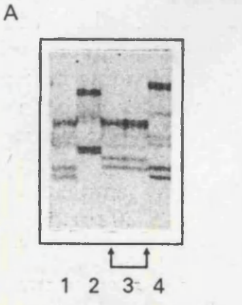


Figure 1 Exon 1 SSCP variants. SSCP analysis on 6% polyacrylamide gels with 5% glycerol. Three exon 1 polymorphisms and the common pattern in normal males are shown. 1, 2, 3, and 4 are polymorphisms 1P1, 1P2, 1P3, and the normal pattern, respectively. M is a DNA marker (1 kb ladder, Gibco BRL). In (A) radiolabelled PCR products were analysed at 340 V overnight and at room temperature; in (B) non-labelled DNA was run at 100 V overnight with water cooling to about 15°C and stained with ethidium bromide.

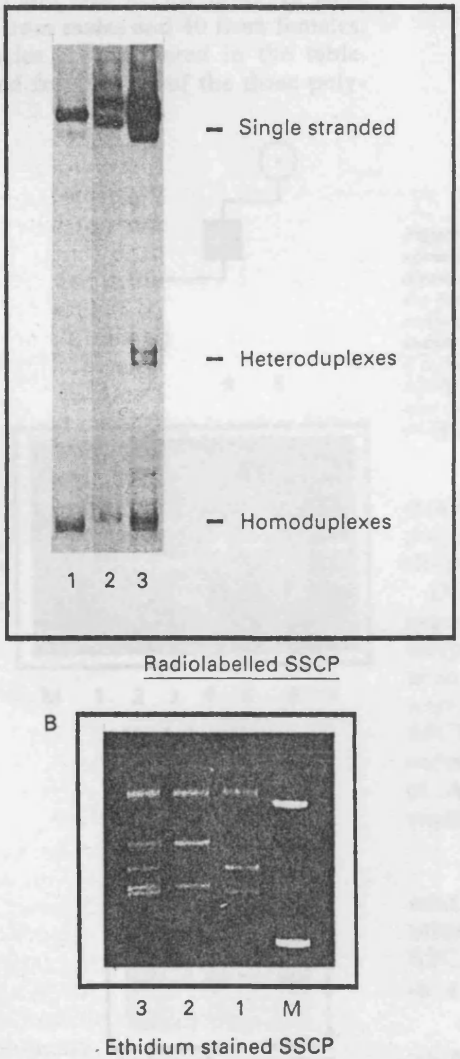


Figure 2 Exon 3 SSCP variants. SSCP analysis of the exon 3 deletion polymorphism and the normal pattern. 1, 2, and 3 are the polymorphism 3P1, the normal allele, and a female carrier of the two alleles (3P1/normal), respectively. M is a DNA marker (1 kb ladder, Gibco BRL). In (A) radiolabelled PCR products were analysed at 340 V overnight at room temperature and on a 6% polyacrylamide, 5% glycerol gel. In (B) non-labelled PCR products were run at 100 V overnight with water cooling (15°C) on a 12% polyacrylamide (with no glycerol) gel and then ethidium bromide stained.

RESTRICTION ENZYME ANALYSIS

PCR amplification of exon 1 was carried out as previously described for non-radioactive SSCP analysis. A volume of 10 µl of PCR product was digested in a total volume of 20 µl, using *Msp*I or *Sac*II and the appropriate enzyme buffer solution (NBL). The entire sample was analysed using a 3% (2% agarose, 1% NuSieve agarose) for *Msp*I or 4% (2% agarose, 2% NuSieve agarose) for *Sac*II. The gel was stained with ethidium bromide (1 mg/ml).

Results

SSCP ANALYSIS OF EXONS 1 AND 3

Exon 1 of α-galactosidase A was amplified from DNA of 61 unrelated normal males and 20 normal females. Four band patterns were observed using SSCP analysis (fig 1). DNA from a subject with each of the three minor band patterns was sequenced.

Two SSCP patterns were observed in PCR amplified material from exon 3 (fig 2). Both polymorphic systems could be visualised using 16 × 20 cm polyacrylamide gels stained with ethidium bromide.

SEQUENCE ANALYSIS OF POLYMORPHIC VARIANTS

Three single base changes were observed. These were G to A at nucleotide 30 before the ATG initiation codon (1P3), G to A at −12 (1P1), and C to T at −10 (1P2) (fig 3).

1P1 removes a restriction enzyme site for *Msp*I. This can be detected by *Msp*I digestion of PCR products. Fig 4 illustrates the *Msp*I

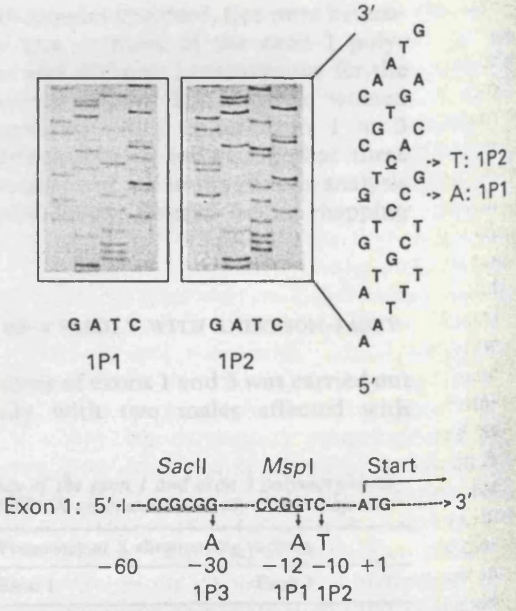


Figure 3 Sequence analysis of exon 1 SSCP variants. Direct sequencing of exon 1 polymorphisms using Dynal bead technology (Dynal, UK) and the Sequenase II kit (USB). Samples were run on a 6% polyacrylamide denaturing (8·3 mol/l urea) gel at 55 W. Sequencing gels for the sense strand of DNA from normal males with polymorphisms 1P1 and 1P2 are shown.

site variant segregating in a family with Anderson-Fabry disease. 1P3 removes a *Sac*II site and segregation of this polymorphism is illustrated in a normal family in fig 5.

The polymorphism in the amplified product from exon 3 was shown to be an insertion-deletion polymorphism by *Bst*NI digestion of the PCR product. Predicted band sizes were 39 bp, 31 bp, 144 bp, and 186 bp. The 31 bp fragment, which falls 80 bp into intron 2, was reduced in size in subjects carrying the polymorphism. This was confirmed by *Hin*fI digestion of the fragment. DNA sequencing showed a deletion of 7 base pairs between nucleotides 7184 and 7193[10] (CCCCCAGCC) presumably owing to replication stuttering across the stretch of Cs. It is not possible to determine the exact breakpoints because of the Cs at either end.

POLYMORPHISM FREQUENCY

One hundred and one X chromosomes were analysed, 61 from males and 40 from females. The frequencies are presented in the table. The combined frequencies of the three polymorphisms in exon 1 were 10/101 (10%). The polymorphism in exon 3 was present in 11/101 chromosomes (12%).

Of the 20 females analysed, five were heterozygous for one or more of the exon 1 polymorphisms and six were heterozygous for the exon 3 polymorphism. Nine of the women were informative using either exon 1 or 3 SSCP polymorphisms, indicating that these variants would be of use in the genetic analysis of Anderson-Fabry disease or in mapping studies.

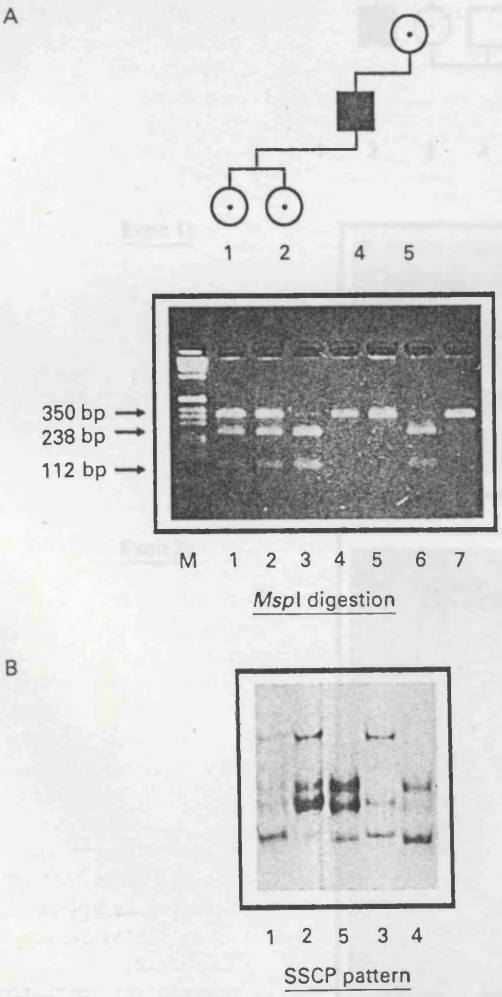


Figure 4 Segregation of the MspI (1P1) variant in a Fabry family. (A) PCR amplification of exon 1, MspI digestion, and agarose (3%) gel electrophoresis confirms the presence of the 1P1 polymorphism shown by radiolabelled SSCP(B). M is a DNA marker (1 kb ladder, Gibco BRL). The pedigree of 1, 2, 4, and 5 is illustrated, 3 and 6 have normal digestion pattern (112 bp and 238 bp products), and 7 is the undigested PCR product (350 bp).

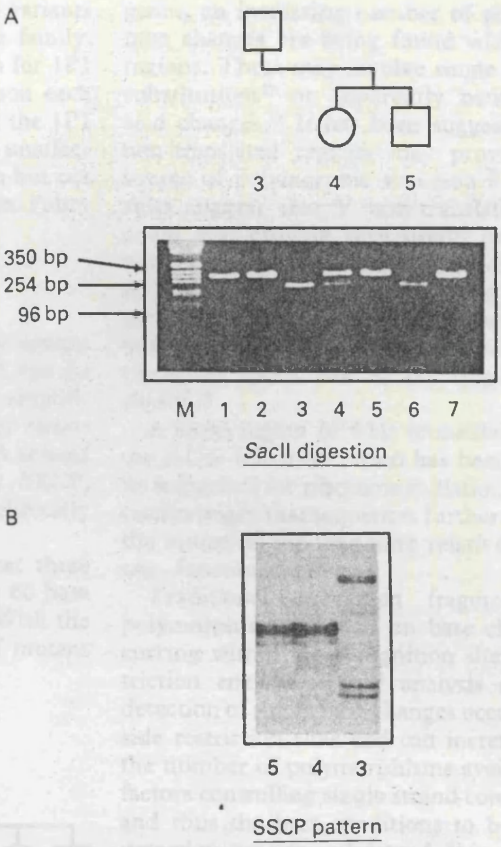


Figure 5 Segregation of the SacII (1P3) variant in a normal family. (A) PCR amplification of exon1, SacII digestion, and agarose (4%) gel electrophoresis confirms the presence of the 1P3 polymorphism shown by radiolabelled SSCP(B). M is a DNA marker (1 kb ladder, Gibco BRL). The pedigree for 3–5 is illustrated, 6 is the normal SacII digestion pattern (96 bp and 254 bp products), 2 is an unrelated normal control who also carries the 1P3 polymorphism, and 1 and 7 are the undigested PCR product (350 bp).

ANALYSIS OF A FAMILY WITH ANDERSON-FABRY DISEASE

SSCP analysis of exons 1 and 3 was carried out in a family with two males affected with

The frequency of the exon 1 and exon 3 polymorphisms in a total of 101 X chromosomes in unrelated subjects.

Frequency of X chromosome variants			
Exon 1		Exon 3	
Variant	Frequency	Variant	Frequency
1P1	3	3P1	11
1P2	5	Normal	90
1P3	2		
Normal	91		
Total	101		101

Anderson-Fabry disease (fig 6). Two variants were found to be segregating in the family. The carrier mother was heterozygous for 1P1 and 3P1. Her affected brother and son each carried the mutant 3P1 allele but not the 1P1 allele. Her daughter has inherited the unaffected allele (with the 1P1 polymorphism but not 3P1) and is not a carrier of Anderson-Fabry disease.

Discussion

We describe a multiallelic polymorphic system within the α-galactosidase A gene. It can be readily visualised using a single PCR amplification and SSCP analysis, with either radioactive or ethidium bromide staining. A second two allele system, also detected by SSCP, increases the informativeness to a clinically useful level.

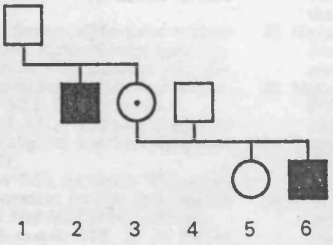
The variation is remarkable in that three point mutations are found within the 60 base pairs of 5′ non-translated sequence. With the increased emphasis on sequencing of mutant genes, an increasing number of polymorphic base changes are being found within coding regions. These may involve single base silent substitutions[23] or apparently neutral amino acid changes.[24] It has been suggested that 3′ non-translated regions may provide a rich source of polymorphic variation.[25] These results suggest that 5′ non-translated regions could also provide very useful polymorphic variation. Few such variants have been described, a striking exception being the highly variable stretch of (CGG) repeats within the 5′ non-translated region of FMR-1 which, when expanded, gives rise to the fragile X syndrome.[26]

A short region of 9 bp immediately before the AUG initiation codon has been shown to be important for ribosome initiation.[27 28] These results imply that sequences further away from the initiation site may have relatively little, if any, functional effect.

Traditional restriction fragment length polymorphisms depend on base changes occurring within the recognition site for a restriction enzyme. SSCP analysis allows the detection of single base changes occurring outside restriction sites and can increase greatly the number of polymorphisms available. The factors controlling single strand conformation, and thus the best conditions to be used for detection, are not understood. It is noteworthy that three base changes occur within 20 bases of each other (two within 2 base pairs) but are all readily detected and each gives a distinctive pattern. This shows that the factors controlling the preferred conformation must be subtle.

Very few polymorphisms have previously been reported within the α-galactosidase A gene, making DNA studies in families with Anderson-Fabry disease difficult. Biochemical detection of female heterozygotes by the measurement of the α-galactosidase activity, usually relative to an unaffected lysosomal enzyme such as β-D-galactosidase or N-acetyl-β-D-hexosaminidase, in plasma, leucocytes, or fibroblasts[5] is unreliable because of the random inactivation of the X chromosome. It is possible to show the cellular mosaicism associated with females heterozygous for an X linked disorder by assaying α-galactosidase in cloned skin fibroblasts. However, this is a very time consuming and technically difficult procedure. Alternatively, α-galactosidase may be measured in individual hair follicles. The detection of hair roots with normal, negligible, and intermediate levels of activity from a female is diagnostic for a heterozygote.[29] This is also a time consuming procedure and does not give complete discrimination, especially when the activities are low with fine hair. Therefore, molecular genetic analysis will provide an important additional detection method. The polymorphisms described here are easy to use and sufficiently frequent to be of clinical use in carrier detection. The example shown in fig 6, where family member 5 can be excluded from being a carrier, illustrates this.

In the past, the majority of SSCP detection methods have used radiolabelled PCR pro-



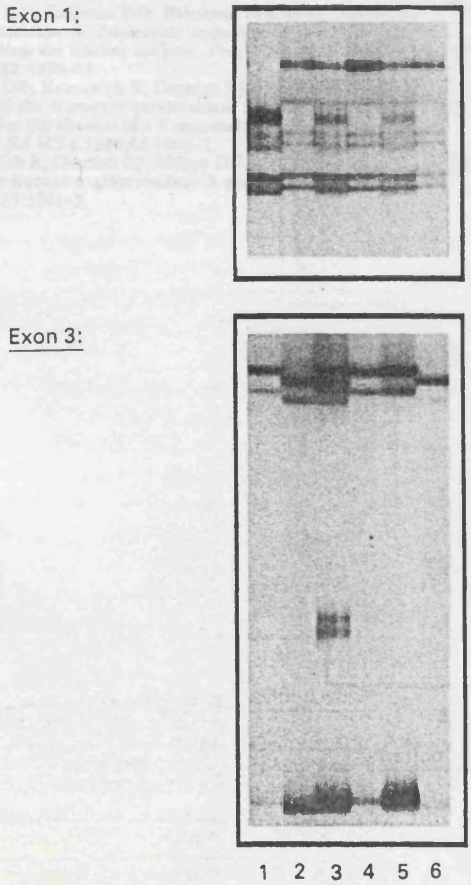


Figure 6 Segregation of 1P1 and 3P1 in a Fabry family. Radiolabelled SSCP analysis of exons 1 and 3 on a 6% polyacrylamide, 5% glycerol gel run at 340 V overnight. The family pedigree for 1–6 is illustrated.

ducts. However, from the results shown here it is clear that single base changes can be effectively distinguished, even when changes are in close proximity to each other (such as polymorphisms 1P1 and 1P2) without the need for radiolabelling. Results can be obtained immediately after running a polyacrylamide gel by ethidium bromide staining and on a small gel system, thus increasing the speed and safety of SSCP as a mutation detection method.

JD is a Wellcome prize student. We gratefully acknowledge financial support from the Research Trust for Metabolic Diseases in Children and the Fabry's Disease Research Fund. We wish to thank Paul Rutland for oligonucleotide synthesis.

1 Dean KJ, Sweeley CC. Studies on human liver alpha-galactosidase. *J Biol Chem* 1979;**262**:2062–5.
2 Gartner S, Conzelmann E, Sandhoff K. Activator protein for the degradation of globotriaosylceramide by human alpha-galactosidase. *J Biol Chem* 1983;**254**:9994–10000.
3 Lemansky P, Bishop DF, Desnick RJ, *et al.* Synthesis and processing of alpha-galactosidase A in human fibroblasts. Evidence for different mutations in Fabry disease. *J Biol Chem* 1987;**262**:2062–5.
4 Desnick RJ, Bishop DF. Fabry disease: alpha-galactosidase A deficiency; Schindler disease: alpha-N-acetylgalactosaminidase deficiency. In Scriver CR, Beaudet AL, Sly WS, Valle D, eds. *The metabolic basis of inherited disease.* 6th ed. New York: McGraw-Hill, 1989:1751–96.
5 Desnick RJ, Allen KY, Desnick SJ, *et al.* Fabry's disease: enzymatic diagnosis of hemizygotes and heterozygotes. *Lab Clin Med* 1973;**81**:157–71.
6 Rietra PJGM, Brouwer-Kelder EM, de Groot WP, *et al.* The use of biochemical parameters for the detection of carriers of Fabry's disease. *J Mol Med* 1976;**1**:237–42.
7 Calhoun DH, Bishop DF, Bernstein HS, *et al.* Fabry disease: isolation of a cDNA clone encoding human α-galactosidase A. *Proc Natl Acad Sci USA* 1985;**82**:7364–8.
8 Bishop DF, Calhoun DH, Bernstein HS, *et al.* Human α-galactosidase A: nucleotide sequence of a cDNA clone encoding the mature enzyme. *Proc Natl Acad Sci USA* 1986;**83**:4859–63.
9 Bishop DF, Kornreich R, Desnick RJ. Structural organization of the human α-galactosidase A gene: further evidence for the absence of a 3′ untranslated region. *Proc Natl Acad Sci USA* 1988;**85**:3903–7.
10 Kornreich R, Desnick RJ, Bishop DF. Nucleotide sequence of the human α-galactosidase A gene. *Nucleic Acids Res* 1989;**17**:3301–3.
11 Desnick RJ, Bernstein HS, Astrin KH, *et al.* Fabry disease: molecular diagnosis of hemizygotes and heterozygotes. *Enzyme* 1987;**38**:54–64.
12 Kirkilionis AJ, Riddell DC, Spence MW, Fenwick RG. Fabry disease in a large Nova Scotia kindred: carrier detection using leucocyte α-galactosidase activity and an *Nco*I polymorphism detected by an α-galactosidase cDNA clone. *J Med Genet* 1991;**28**:232–40.
13 MacDermot KD, Morgan SH, Cheshire JK, *et al.* Anderson-Fabry disease, a close linkage with highly polymorphic DNA markers DXS17, DXS87 and DXS88. *Hum Genet* 1987;**77**:263–6.
14 Ishii S, Sakaruba H, Suzuki Y. Point mutations in the upstream region of the alpha-galactosidase A gene exon 6 in an atypical variant of Fabry disease. *Hum Genet* 1992;**89**:29–32.
15 Yokoi T, Shinoda K, Ohno I, *et al.* A 3′ splice site consensus sequence mutation in the intron 3 of the alpha-galactosidase A gene in a patient with Fabry disease. *Jpn J Hum Genet* 1991;**36**:245–50.
16 Sakaruba H, Oshima A, Fukuhara Y, *et al.* Identification of point mutations in the alpha-galactosidase gene in classical and atypical hemizygotes with Fabry disease. *Am J Hum Genet* 1990;**47**:784–9.
17 Sakaruba H, Eng CM, Desnick RJ, *et al.* Invariant exon skipping in the human alpha-galactosidase A pre-mRNA: a g+1 to t substitution in a 5′-splice site causing Fabry disease. *Genomics* 1992;**12**:643–50.
18 Koide T, Ishiura M, Iwai K. A case of Fabry's disease in a patient with no alpha-galactosidase A activity caused by a single amino acid substitution of Pro-40 by Ser. *FEBS Lett* 1990;**259**:353–6.
19 Bernstein HS, Bishop DF, Astrin DF, *et al.* Fabry disease: six gene rearrangements and an exonic point mutation in the α-galactosidase gene. *J Clin Invest* 1989;**83**:1390–9.
20 Scheidt WV, Eng CM, Fitzmaurice TG, *et al.* An atypical variant of Fabry's disease with manifestations confined to the myocardium. *N Engl J Med* 1991;**324**:395–9.
21 Hayashi K. PCR-SSCP: a simple and sensitive method for detection of mutations in genomic DNA. *PCR: methods and applications* 1991;**1**:34–8.
22 Miller SA. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res* 1988;**16**:1215.
23 Trofatter JA, Pratt VM, Dlouhy SR, *et al.* AhaII polymorphism in human X-linked proteolipid protein gene (PLP). *Nucleic Acids Res* 1991;**19**:6057.
24 Petty EM, Carstens R, Bale AE. Ornithine transcarbamylase polymorphism detected by PCR introduction of DraI site. *Nucleic Acids Res* 1991;**19**:690.
25 Levitt RC. Polymorphisms in the transcribed 3′ untranslated region of eukaryotic genes. *Genomics* 1991;**11**:484–9.
26 Fu YH, Kuhl DPA, Pizzuti A, *et al.* Variation of the CGG repeat at the fragile X site results in genetic instability: resolution of the Sherman paradox. *Cell* 1991;**67**:1047–58.
27 Kozak M. Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. *Cell* 1986;**44**:283–92.
28 Kozak M. An analysis of 5′-noncoding sequences from 669 vertebrate messenger RNAs. *Nucleic Acids Res* 1987;**15**:8125–48.
29 Ejiofor A, Robinson D, Hamers M, *et al.* Anderson-Fabry disease; rapid detection carriers by hair bulb analysis. *J Inherited Metab Dis* 1978;**1**:71–7.

Human Molecular Genetics, 1993, Vol. 2, No. 7 1051–1053

Mutation analysis in patients with the typical form of Anderson–Fabry disease

Joanna P.Davies, Bryan G.Winchester and Sue Malcolm*

Units of Molecular Genetics and Enzymology, Division of Biochemistry and Genetics, Institute of Child Health, 30 Guildford Street, London WC1N 1EH, UK

Received March 22, 1993; Revised and Accepted April 9, 1993

Anderson–Fabry disease is an X-linked disorder caused by the deficiency of the lysosomal hydrolase, alpha-galactosidase A (EC 3.2.1.22). This enzyme catalyses the hydrolysis of alpha-galactosidic linkages at the non-reducing terminals of glycosphingolipids in normal individuals (1). In Fabry patients, loss of this activity results in the progressive intra-lysosomal accumulation of these glycoconjugates in a variety of tissues. In particular, the vascular endothelium is affected and in severe cases clinical symptoms include angiokeratoma, pain in the extremities, renal failure and cerebro- and cardiovascular manifestations.

Fabry disease can be diagnosed in severely affected males by demonstration of a deficiency of alpha-galactosidase activity in plasma and leukocytes (2). Identification of heterozygous females is, however, often inconclusive, due to the random X-inactivation phenomenon (3, 4). Linkage studies using restriction fragment length polymorphisms are possible, although they are not always informative. This problem can be overcome by molecular analysis of the gene, which is now possible, since the sequence of the alpha-galactosidase A gene is known (5).

The alpha-galactosidase A gene, which is located at Xq22 (1) is 12kb in size and consists of 7 exons, each of which is less than 300 bp long. It is therefore an ideal structure for exon by exon single-strand conformation polymorphism (SSCP) analysis (6). This allows the identification of sequence changes, including point mutations, which alter the electrophoretic migration pattern of single stranded DNA through a polyacrylamide gel.

Here, we report the identification of 7 point mutations and a one base pair insertion in families with the classical form of Fabry disease. Three of the point mutations result in the generation of stop codons and four cause substitution of an amino acid residue. The one base pair insertion causes a frameshift and premature termination in protein translation. These mutations have been identified by polymerase chain reaction (PCR) amplification of each exon using newly designed primers, followed by SSCP analysis and sequencing.

Intronic oligonucleotide primers were designed using the PRIMER program provided by the HGMP resource centre based at MRC Clinical Research Centre, Harrow, UK (Table 1). Genomic DNA was isolated from whole blood by an ammonium acetate salting out procedure, as previously described (7). All seven exons and a small portion of flanking intron sequences from the genomic DNA of eight affected males and one manifesting female in unrelated Fabry families were amplified by PCR. No exons failed to amplify. PCR was conducted in a 50 μl reaction mixture volume, containing genomic DNA, 40 pmol of each primer (Table 1), 0.2 mM dATP, dGTP, dTTP and 0.02 mM dCTP, 1 μCi of ^{32}P-dCTP, 1 unit of Promega Taq polymerase and in Promega enzyme buffer with 1.5 mM $MgCl_2$. The conditions for amplification were as follows: 30 cycles of denaturation at 94°C (30 seconds), an annealing step of 30 seconds (at 58°C for exons 1 and 3, at 59°C for exons 2 and 4 and 61°C for exons 5, 6 and 7) and an elongation step at 72°C (1 minute). The amplification was completed by a final 10 minute elongation step at 72°C.

SSCP analysis was conducted for each exon as follows. A volume of 5 μl of each PCR product was diluted by adding 40 μl of 0.1% SDS, 10 mM EDTA solution and 2 μl of this was mixed with 2μl of 95% formamide, 0.05% xylene cyanol, 0.05% bromophenol blue, 20 mM EDTA solution. The samples were then denatured by heating at 94°C for 3 minutes and placed on ice before loading onto a 0.4mm thick 30×40cm, polyacrylamide gel (Protogel: 30% acrylamide to 0.8% bisacrylamide, National Diagnostics), with 5% glycerol. Electrophoresis was carried out in 0.09 M Tris-Borate, 0.002 M EDTA (pH 8.3) buffer (TBE), at 340V and room temperature for approximately 15 hours, until the xylene cyanol had migrated 31 cm. The gel was dried onto 3MM Whatman paper under vacuum at 80°C and exposed to X-ray film (Kodak, IBI Ltd) for 1–3 days at −70°C. All samples showing an SSCP variation were sequenced directly, using Dynal bead (Dynal, UK) DNA strand separation and sequencing with the Sequenase II kit (USB), as previously described (7).

In order to exclude common polymorphic sequence changes, DNA from twenty normal male individuals was also amplified and analysed for SSCP variants in each exon and short sections of the adjacent intron sequences. A further 80 normal X chromosomes were analysed in this manner for exons 1, 3, 5, 6 and 7. Three previously described polymorphic variants were identified in the 5′untranslated region of exon 1 and one intronic variant was found in the exon 3 PCR product (7). Single intronic polymorphic variants were also found in the 5′sequence adjacent to exon 5 (an A to G transition at nucleotide 10115 of the gene) and to exon 7 (a C to T transition at nucleotide 10956 of the gene). No SSCP changes were observed in the PCR products for exons 2, 4 or 6 and no polymorphisms were found in any part of the amino acid coding region of the gene.

The following unique sequence changes were identified in patients with the classical form of Fabry disease Figure 1. They consist of substitutions at the following nucleotide positions of the cDNA (8): 644 (A→T, Asn^{215}→Ser, exon 5); 679 (C→T, Arg^{227}→Stop, exon 5); 806 (T→C, Val^{269}→Ala, exon 6); 861 (G→A, Trp^{287}→Stop, exon 6); 979 (C→A, Gln^{327}→Lys, exon 6); 1024 (C→T, Arg^{342}→Stop, exon 7); and 1081 (G→A, Gly^{361}→Arg, exon 7). A one base insertion of an adenine nucleotide between nucleotides 716 and 717 of the normal cDNA sequence was also observed in another patient (results not shown).

* To whom correspondence should be addressed

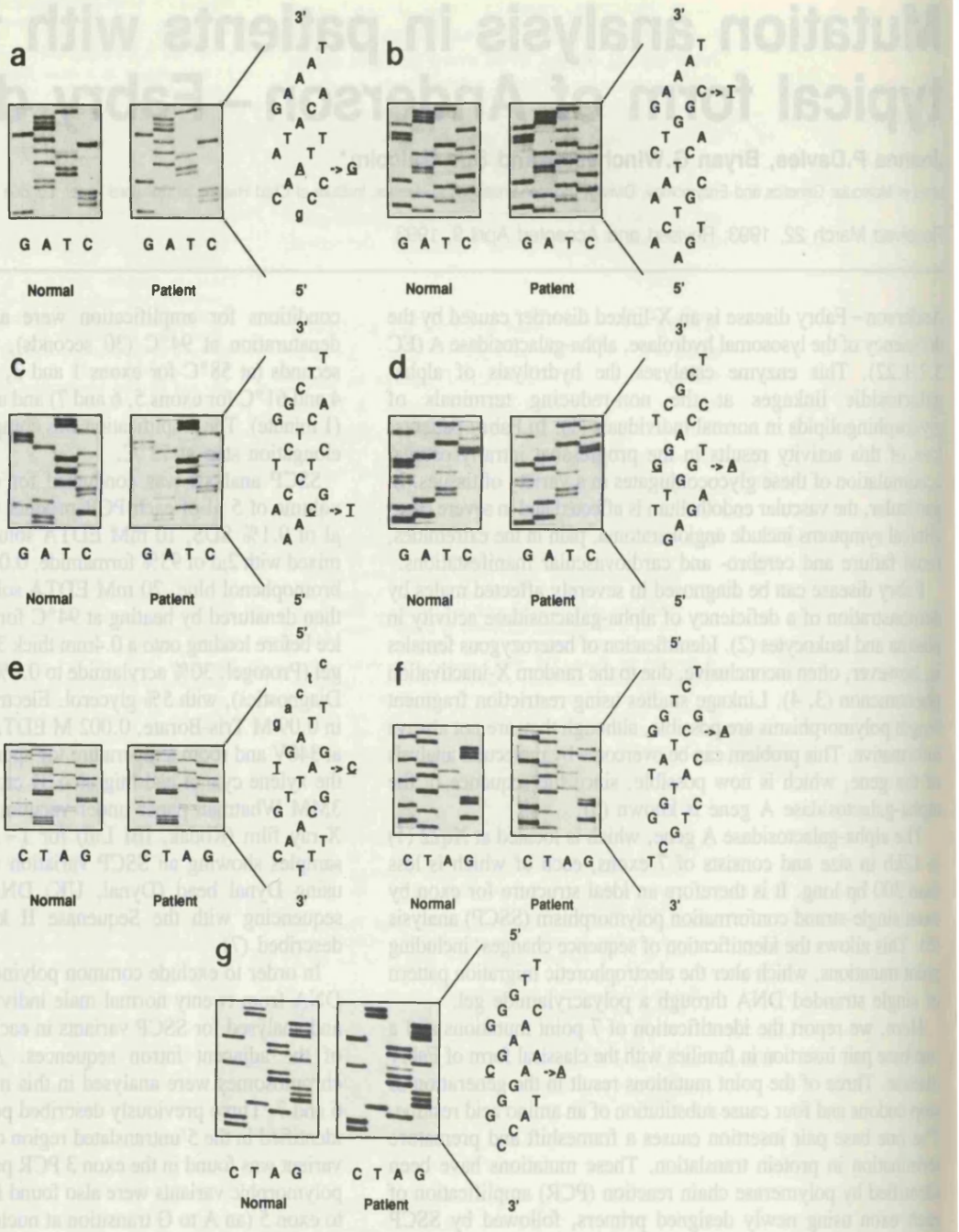


Figure 1. Sequence analysis of exons showing an altered SSCP pattern in six affected males (a)–(f) and one manifesting female (g) at cDNA nucleotide positions (8): (a) 644 (exon 5, Asn^{215}→Ser); (b) 679 (exon 5, Arg^{227}→stop); (c) 1024 (exon 7, Arg^{342}→stop); (d) 1081 (exon 7, Gly^{361}→Arg); (e) 806 (exon 6, Val^{269}→Ala); (f) 861 (exon 6, Trp^{287}→stop); and (g) 979 (exon 6, Gln^{327}→Lys).

The stop mutation in exon 5 (at nucleotide 679, in a CpG site) was found in two unrelated Fabry families. We have failed to find mutations, using this method, in four families. These sequence changes are excluded as polymorphisms and presumed to be disease-causing because: (a) none was found in DNA from 100 chromosomes; (b) no other non-polymorphic SSCP changes were seen in the other six exons; and (c) the exon containing the unique change was fully sequenced and did not contain any other mutations.

Four of our mutations can be detected by restriction enzyme digestion. The mutations in Figure 1 (a), (d) and (f) remove restriction enzyme recognition sites for TspE I, Sau96 I and Bgl I, respectively. The insertion mutation (result not shown) in exon 5 creates an Mse I site.

Three of our point mutations (Figure 1 (b), (c), (f)) result in the generation of a stop codon and are predicted to cause premature termination of protein translation. These are presumed to be the disease-causing mutations, accounting for the severe clinical picture in each of these cases. The insertion of one base in exon 5 is predicted to cause a frameshift mutant in which amino acids 241–249 (SILDWTSFN), in the normal become EYLGLDIF-stop in the patient. This is predicted to cause

Table 1. Intronic oligonucleotide primers designed for amplifying and sequencing the seven exons of the alpha-galactosidase A gene.

Primer pair for exon:	Sequence	Nucleotide position (5)
1	5'-GTCCCAGTTGCCAGAGAAAC-3'	1054-1073
	5'-AAAGGGAAGGGAGTACCCAA-3'	1403-1384
2	5'-ATGGGAGGTACCTAAGTGTTCA-3'	5033-5054
	5'-GTCCTCTGAATGAACAAGAACA-3'	5306-5285
3	5'-GATGATTTTGGGGGTTTGTG-3'	7123-7142
	5'-GATTGGTTCTTTGGCTCAGC-3'	7522-7503
4	5'-ATAGCCCCAGCTGGAAATTC-3'	8262-8281
	5'-CAGTTCTATTGGATTCTGGGC-3'	8445-8425
5	5'-AGAAGGCTACAAGTGCCTCC-3'	10042-10061
	5'-GCAGGGTCTTGAACAAGGAG-3'	10327-10308
6	5'-GGATGCTGTGGAAAGTGGTT-3'	10413-10432
	5'-GGCCCAAGACAAAGTTGGTA-3'	10769-10750
7	5'-GAATGCCAAACTAACAGGGC-3'	10912-10931
	5'-GCATGAGCCACCTAGCCTT-3'	11352-11334

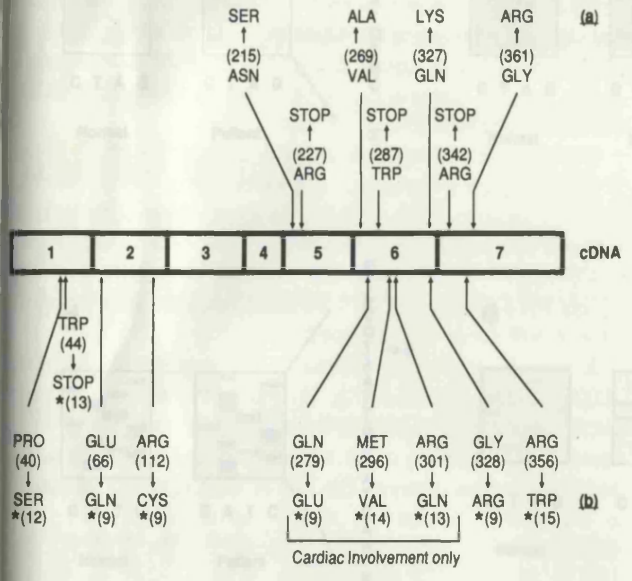


Figure 2. Diagram indicating the distribution of single base substitutions in the cDNA of alpha-galactosidase A: (a) our mutations; and (b) previously reported mutations. '*' indicates reference sources.

premature termination at amino acid 249) and consequently the severe phenotype.

Four other mutations result in amino acid substitutions, two of which cause a change in the charge of the amino acid side chain (Figure 1 (d) and (g)) are therefore very likely to affect protein structure and/ or function. The other two substitutions are an asparagine to serine (Figure 1(a)) and a valine to alanine (Figure 1(e)), neither of which leads to a change in charge.

The asparagine to substitution in exon 5 of a patient with classic Fabry disease symptoms but with no kidney involvement (Figure 1 (a)) is of particular interest as it disrupts a putative protein glycosylation site (Asn-Tyr-Thr). Therefore, this mutation may alter the pattern of glycosylation in the alpha-galactosidase A of this patient.

The valine to alanine substitution (Figure 1 (e)) in the upstream region of exon 6 in a patient with the full symptoms of Fabry's disease is also of interest. This is presumed to be disease-causing, based on the criteria described above. Such a change could disrupt specific hydrophobic interactions essential for the function of the enzyme. A recent study (9) suggests that mutations in this region, including those which cause a change in the amino acid charge (see Figure 2), cause a less severe, asymptomatic form of the disease associated with impaired cardiac function as the only symptom. These observations suggest that mutations in the upstream region of exon 6 of alpha-galactosidase A can have diverse effects on the properties of the enzyme.

Seven putative disease-causing single base substitutions and one small insertion mutation have therefore been identified in nine unrelated families with the classical form of Fabry disease. This nearly doubles the number of previously published point mutations, emphasising the heterogeneous nature of the disease. Figure 2 indicates the position of our seven substitutions in relation to known point mutations in the cDNA. Two splice-site point mutations have also been reported previously (10, 11).

The effects of these new mutations on the structure and function of alpha-galactosidase A will be investigated by expression in cells. Collation of these results with other mutation analyses will hopefully lead to a better understanding of the correlation of genotype with phenotype in this genetically and phenotypically heterogeneous disease.

ACKNOWLEDGEMENTS

J.D. thanks the Wellcome Trust for a prize research studentship. The support of the Fabry's Disease Research Fund (UK) is also gratefully acknowledged. We thank Mrs Elisabeth Young for helpful discussions and Dr George Gray, Prof. Rodney Harris, Dr T.Leslie, Dr Helen Michelakakis and Dr Kay MacDermot for providing us with material from Fabry patients.

ABBREVIATIONS

PCR — polymerase chain reaction
SSCP — single strand conformation polymorphism

REFERENCES

1. Desnick,R.J. and Bishop,D.F. (1989) In, Scriver,C.R., Beaudet,A.L., Sly,W.S. and Valle,D. (eds), *The Metabolic Basis of Inherited Diseases*. McGraw Hill, New York, 6th Ed, Vol II, pp. 1751–1796.
2. Desnick,D.F, Allen,K.Y., Desnick,S.J., Raman,M.K., Bernlohr,W. and Krivit,W. (1973) *J. Lab. Clin. Med.* **81**, 157–171.
3. Lyon,M.F. (1961) *Nature* **190**, 372–373.
4. Romeo,G. and Migeon,B.R. (1970) *Science* **170**, 180–181.
5. Kornreich,R., Desnick,R.J. and Bishop,D.F. (1989) *Nucleic Acids Res.* **17**, 3301–3302.
6. Hayashi,K. (1991) *PCR: methods and applications* **1**, 34–38.
7. Davies,J.P., Winchester,B.G. and Malcolm,S. (1993) *J. Med. Genet.* In press.
8. Bishop,D.F., Komreich,R.,Eng,C.M., Ioannou,Y.A., Fitzmaurice,T.F. and Desnick,R.J. (1991) In, Salvayre,R., Douste-Blazy,L. and Gatt,S. (eds), *Lipid Storage Disorders — Biological and Medical Aspects*. Plenum press, New York, pp. 809–822.
9. Ishii,S., Sakaruba,H. and Suzuki,Y. (1992) *Hum. Genet* **89**, 29–32.
10. Yokoi,T., Shinoda,K., Ohno,I., Kato,K., Miyawaki,T. and Taniguchi,N. (1991) *Jpn. J. Hum. Genet.* **36**, 245–250.
11. Sakaruba,H., Eng,C., Desnick,R.J. and Bishop,D.F. (1992) *Genomics* **12**, 643–650.
12. Koide,T., Ishiura,M., Iwai,K., Inoue,M., Kaneda,Y., Okada,Y. and Uchida,T. (1990) *FEBS Lett.* **259**, 353–356.
13. Sakuraba,H., Oshima,A., Fukuhara,Y., Shimmoto,M., Nagao,Y., Bishop,D.F., Desnick,R.J. and Suzuki,Y. (1990) *Am. J. Hum. Genet.* **47**, 784–789.
14. von Scheidt,W., Eng,C.M., Fitzmaurice,T.F., Erdmann,E., Hubner,G., Olsen,G.J., Christomanou,H., Kandolf,R., Bishop,D.F. and Desnick,R.J. (1991) *N. Eng. J. Med.* **324**, 395–399.
15. Bernstein,H.S., Bishop,D.F., Astrin,K.H., Korneich,R., Eng, C.M.,Sakuraba, H. and Desnick,R.J. (1989) *J. Clin. Invest.* **83**, 1390–1399.

Human Molecular Genetics, 1994, Vol. 3, No. 4 667–669

Detection of 8 new mutations in the α-galactosidase A gene in Fabry disease

Joanna Davies, Helen Christomanou[1], Bryan Winchester* and Susan Malcolm

Units of Biochemistry and Molecular Genetics, Division of Biochemistry and Genetics, Institute of Child Health (University of London), 30 Guilford Street, London WC1N 1EH, UK and [1]Laboratory of Neurochemistry and Molecular Biology, Oceanographic Biochemical Society, 2 Lampsakou Street, 115 28 Athens, Greece

Received January 20, 1994; Revised and Accepted February 21, 1994

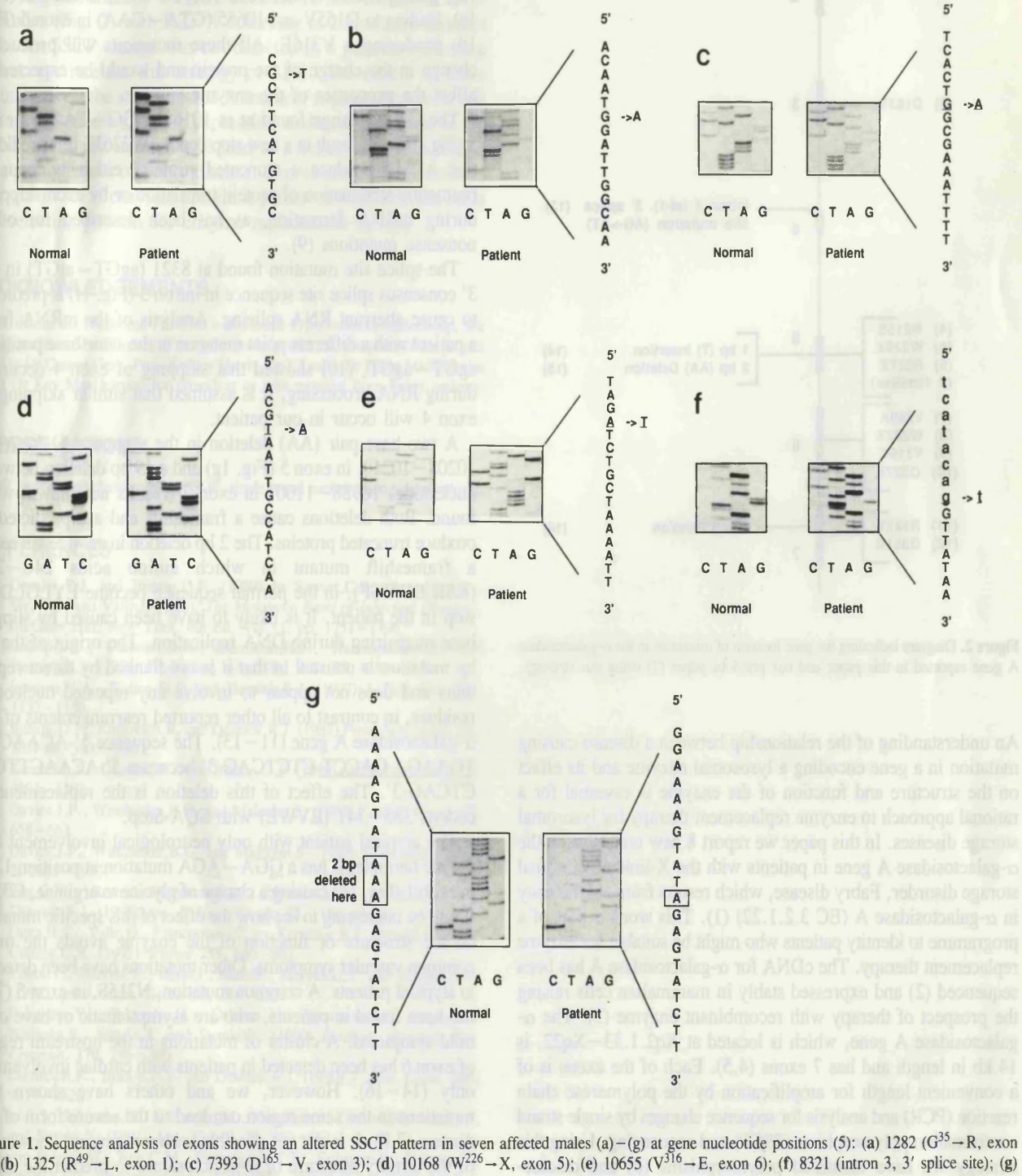


Figure 1. Sequence analysis of exons showing an altered SSCP pattern in seven affected males (a)–(g) at gene nucleotide positions (5): (**a**) 1282 (G^{35}→R, exon 1); (**b**) 1325 (R^{49}→L, exon 1); (**c**) 7393 (D^{165}→V, exon 3); (**d**) 10168 (W^{226}→X, exon 5); (**e**) 10655 (V^{316}→E, exon 6); (**f**) 8321 (intron 3, 3′ splice site); (**g**) 2 bp (AA) deletion between nucleotides 10208–10210.

* To whom correspondence should be addressed

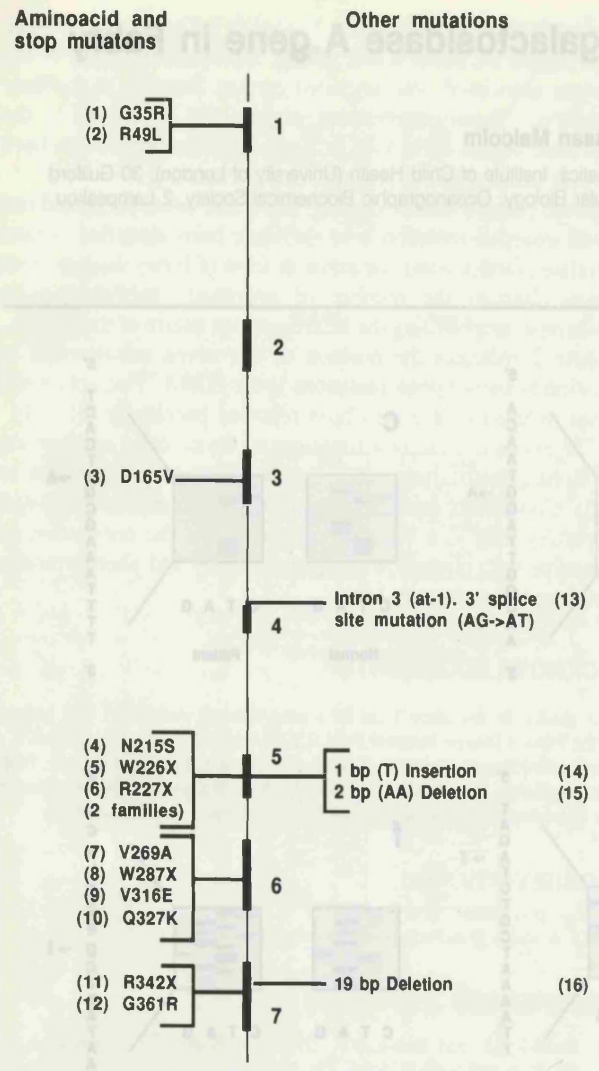


Figure 2. Diagram indicating the gene location of mutations in the α-galactosidase A gene reported in this paper and our previous paper (7) using our strategy.

An understanding of the relationship between a disease-causing mutation in a gene encoding a lysosomal enzyme and its effect on the structure and function of the enzyme is essential for a rational approach to enzyme replacement therapy for lysosomal storage diseases. In this paper we report 8 new mutations in the α-galactosidase A gene in patients with the X-linked lysosomal storage disorder, Fabry disease, which results from a deficiency in α-galactosidase A (EC 3.2.1.22) (1). This work is part of a programme to identity patients who might be suitable for enzyme replacement therapy. The cDNA for α-galactosidase A has been sequenced (2) and expressed stably in mammalian cells raising the prospect of therapy with recombinant enzyme (3). The α-galactosidase A gene, which is located at Xq2.1.33−Xq22, is 14 kb in length and has 7 exons (4,5). Each of the exons is of a convenient length for amplification by the polymerase chain reaction (PCR) and analysis for sequence changes by single strand conformation polymorphism (SSCP) and sequencing. Using this strategy we have identified polymorphisms (6) and disease-causing mutations (7) in the α-galactosidase gene. Mutations discovered by other groups are summarised in reference 8.

Seven of the patients analysed by the same strategy in this report had the classical severe form of the disease, characterised by pain in the extremities, angiokeratoma, renal failure and cerebral and cardiovascular manifestations (1). The eighth patient had an atypical form of the disease with neurological involvement and spastic hemiplegia only.

Three missense and one nonsense mutation, two small deletions and a mutation in a 3′-splice site consensus sequence were found in families with the classical form of the disease (Fig. 1). The missense mutations occurred at the following nucleotide positions in the gene structure (5):- 1325 (CGC→CTC) in exon 1 (Fig. 1a), giving rise to R49L; 7393 (GAT→GTT) in exon 3 (Fig. 1e), leading to D165V and 10655 (GTA→GAA) in exon 6 (Fig. 1d) producing a V316E. All these mutations will produce a change in the charge of the protein and would be expected to affect the properties of the enzyme.

The G→A change found at nt 10168 (TGG→TAG) in exon 5 (Fig. 1c) will result in a new stop codon, W226X. It is predicted that it will produce a truncated protein, either by causing premature termination of protein translation or by exon skipping during mRNA formation, as has been described for other nonsense mutations (9).

The splice site mutation found at 8321 (agGT→atGT) in the 3′ consensus splice site sequence in intron 3 (Fig. 1f) is predicted to cause aberrant RNA splicing. Analysis of the mRNA from a patient with a different point mutation at the same base position, agGT→aaGT, (10) showed that skipping of exon 4 occurred during RNA processing, It is assumed that similar skipping of exon 4 will occur in our patient.

A two base pair (AA) deletion in the sequence 5′-AAA-3′, 10208−10210, in exon 5 (Fig. 1g) and a 19 bp deletion between nucleotides 10988−11007 in exon 7 (results not shown) were found. Both deletions cause a frameshift and are predicted to produce truncated proteins. The 2 bp deletion in exon 5 will cause a frameshift mutant in which amino acids 240−248 (KSILDWTSF), in the normal sequence become EYLGLDIF-stop in the patient. It is likely to have been caused by slipped base mispairing during DNA replication. The origin of the 19 bp mutation is unusual in that it is not flanked by direct repeat units and does not appear to involve any repeated nucleotide residues, in contrast to all other reported rearrangements of the α-galactosidase A gene (11−13). The sequence 5′-ACAACTT TGAAG---GACCT CTCTCAG-3′ becomes 5′-ACAACTTCT-CTCAG-3′. The effect of this deletion is the replacement of codons 338−341 (EVWE) with SQA-Stop.

The atypical patient with only neurological involvement and spastic hemiplegia has a GGA→AGA mutation at position 1282 in exon 1 (Fig. 1b) causing a change of glycine to arginine, G35R. It will be interesting to see how the effect of this specific mutation on the structure or function of the enzyme avoids the more common vascular symptoms. Other mutations have been detected in atypical patients. A common mutation, N215S, in exon 5 (7,8) has been found in patients, who are asymptomatic or have only mild symptoms. A cluster of mutations in the upstream region of exon 6 has been detected in patients with cardiac involvement only (14−16). However, we and others have shown that mutations in the same region can lead to the severe form of the disease (7,8). It is not yet possible with our present knowledge of the tertiary structure of α-galactosidase to predict the effect of missense mutations on the structure and function of the enzyme.

All these sequence changes are excluded as polymorphisms and presumed to be disease-causing because: (a) none was found in DNA from 100 chromosomes in normal individuals; (b) no other non-polymorphic SSCP changes (6) were seen in the remaining six exons; and (c) the exon containing the unique change was fully sequenced and did not contain any other mutations. Five of these new mutations should be be detectable by restriction enzyme digestion. The mutations, R49L (Fig. 1a), W226X (Fig. 1c), V316E (Fig. 1d) and D165E (Fig. 1e) remove restriction enzyme recognition sites for HhaI, MaeI, MaeII and BglII, respectively and the 19 bp deletion removes an EcoNI site.

Using this strategy we have detected 16 new mutations in 17 families (Fig. 2) with R227X occurring in 2 unrelated families (7). An SSCP change was not detected in 4 patients indicating a success rate for this strategy of 81%. The effects of these mutations on the structure and function of α-galactosidase A will be investigated by protein expression studies and analysis of the RNA. This information will lead to a better understanding of the pathogenesis of Fabry disease. Expression of these mutants *in vivo* to see if enzymic protein is produced will form the basis of selecting patients for enzyme replacement therapy.

ACKNOWLEDGEMENTS

J.D. thanks the Wellcome Trust for a Wellcome Prize research studentship. We are grateful to Mrs Elisabeth Young for her help with many aspects of the project and to Dr George Gray, Prof. Rodney Harris, Dr T.Leslie, Dr Helen Michelakakis and Dr Kay MacDermott for providing us with material from Fabry patients.

ABBREVIATIONS

POR, polymerase chain reaction; SSOP, single strand conformation polymorphism.

REFERENCES

1. Desnick,R.J. and Bishop,D.F. (1989) In Scriver,C.R., Beaudet,A.L., Sly,W.S. and Valle,D. (eds), *The Metabolic Basis of Inherited Diseases*. McGraw Hill, New York, 6th Ed, Vol 11, pp. 1751–1796.
2. Bishop,D.F., Calhoun,D.H., Bernstein,H.S., Hantzopoulous,P. and Desnick,R.J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 4859–4863.
3. Ioannou,Y.A., Bishop,D.F. and Desnick,R.J. (1992) *J. Cell Biol.* **119**, 1137–1150.
4. Bishop,D.F., Kornreich,R. and Desnick,R.J. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 3903–3907.
5. Kornreich,R., Desnick,R.J. and Bishop,D.F. (1989) *Nucleic Acids Res.* **17**, 3301–3302.
6. Davies,J.P., Winchester,B.G. and Malcolm,S. (1993) *J. Med. Genet.* **30**, 658–663.
7. Davies,J.P., Winchester,B.G. and Malcolm,S. (1993) *Hum. Mol. Genet.* **2**, 1051–1053.
8. Eng,C.M., Resnick-Silverman,L.A., Niehaus,D.J., Astrin,K.H. and Desnick,R.J. (1993) *Am. J. Hum. Genet.* **53**, 1186–1197.
9. Dietz,H.C., Valle,D., Francomano,C.A., Kendzior,R.J., Pyeritz,R.E. and Cutting,G.R. (1993) *Science* **259**, 660–663.
10. Yokoi,T., Shinoda,K., Ohno,I., Kato,K., Miyawaki,T. and Taniguchi,N. (1991) *Jpn J. Hum. Genet.* **36**, 245–250.
11. Fukuhara,Y., Sakaruba,H., Oshima,A., Shimmoto,M., Nagao,Y., Nadaoka,S., Suzuki,T. and Suzuki,Y. (1990) *Biochem. Biophys. Res. Commun.* **170**, 296–300.
12. Kornreich,R., Bishop,D.F. and Desnick,R.J. (1990) *J. Biol. Chem.* **265**, 9319–9326.
13. Ishii,S., Sakuraba,H., Shimmoto,M., Minamikawi-Tachino,R., Suzuki,T. and Suzuki,Y. (1991) *Ann. Neurol.* **29**, 560–564.
14. von Scheidt,W., Eng,C.M., Fitzmaurice,T.F., Erdmann,E., Hubner,G., Olson,E.G.J., Christomanou,H., Kandolf,R., Bishop,D.F. and Desnick,R.J. (1991) *N. Eng. J. Med.* **324**, 359–399.
15. Ishii,S., Sakuraba,H. and Suzuki,Y. (1992) *Hum. Genet.* **89**, 29–32.
16. Sakuraba,H., Oshima,A., Fukuhara,Y., Shimmoto,M., Nagao,Y., Bishop,D.F., Desnick,R.J. and Suzuki,Y. (1990) *Am. J. Hum. Genet.* **47**, 784–789.