APOLIPOPROTEIN E ALLELE DISTRIBUTION IN A SOUTH AFRICAN INDIAN FEMALE POPULATION: EFFECT ON THE LIPID PROFILE

by

NIRMALA GOUNDEN BSc (Hons)

Submitted in partial fulfilment of
the requirements for the degree of

MASTER OF MEDICAL SCIENCE

in the
Department of Chemical Pathology
Faculty of Medicine
University of Natal
Durban
1993

DEDICATED WITH LOVE TO MUM, DAD AND RAJIN

Discovery consists of seeing what everybody has seen
and thinking what nobody has thought.

in " The Scientist Speculates"
1962

ABSTRACT

Genetic polymorphism of apolipoprotein (apo) E has been shown to account for a significant amount of variance in plasma lipid and lipoprotein levels, thereby contributing to the incidence of cardiovascular disease across populations. In this cross-sectional study apo E genotypes were determined in a sample of 173 healthy, middle-aged Indian women using a restriction isotyping method, in which DNA was amplified by PCR and the *Cfo*I restricted DNA fragments were separated on a polyacrylamide gel, allowing unambiguous typing of the common apo E genotypes. Considering the three common alleles, $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$, a reduced frequency of the $\epsilon 2$ allele was observed in the study population in comparison to other populations around the world. This finding underlines the heterogeneity of apo E allele frequencies in different populations.

This study also investigated possible effects of apo E genotype on lipoprotein changes in this sample of women spanning the menopause. Apo E polymorphism was associated with significant differences in plasma lipid levels. Notably, total and low density lipoprotein cholesterol and more especially plasma triglyceride concentrations were increased in carriers of the $\epsilon 3/4$ genotype. Two-way analysis of variance of the effect of apo E genotype and menopausal status on the lipid profile showed no significant interaction effect, indicating that the effects of apo E genotype on the lipid profile do not differ significantly between premenopausal and postmenopausal women. Age and to a lesser extent the waist:hip ratio also correlated with lipid concentrations, but menopausal status had no apparent effect in this sample.

This study confirms the potentially deleterious effect of the $\epsilon 4$ allele, in a vulnerable population. The reduced occurrence of the E2 isoform, which is considered to offer a measure of protection against cardiovascular disease, may contribute to the relatively high incidence of coronary heart disease in the South African Indian population. However, the relatively low incidence of the $\epsilon 2$ allele may protect this population against the occurrence of type III hyperlipoproteinaemia precipitated by diabetes and obesity in $\epsilon 2/2$ homozygotes.

In this research the statistical planning and analyses, and recommendations arising from these analyses, have been done with the support of the Institute for Biostatistics of the Medical Research Council.

This study was supported by the University of Natal through a Postgraduate Scholarship to Miss N Gounden and by grants to Professor GMB Berger from the South African Medical Research Council and the Atherosclerosis Risk Factor Research Programme (founded by the Egg Control Board of South Africa)

PREFACE

This study represents original work by the author and has not been submitted in any form to another University. Where use was made of work of others it has been duly acknowledged in the text.

The research described in this thesis was carried out in the Department of Chemical Pathology, University of Natal, under the supervision of Professor GMB Berger and Dr RJ Pegoraro.

Professor GMB Berger

Date: 9/4/94

RJ Pegoraro

Dr RJ Pegoraro

Date: 9/4/94

Author

Date: 09-04-94

ACKNOWLEDGEMENTS

The author offers sincere thanks to the following people:

My supervisors, Professor GMB Berger and Dr RJ Pegoraro, for their constructive criticism during the course of this project. A special thank you to Professor GMB Berger for suggesting the project, for his advice, and for spending so much of his time reading draft versions of this work. In particular I would like to thank him for the latitude he afforded me in my studies - not having the development of this thesis rigidly dictated to me has made me far richer in research experience and technique than I would otherwise have been. Particular thanks to Dr Kogie Reddy for teaching me the ground work in PCR, to Dr Jenisa Naidoo for her constant encouragement and invaluable clinical assistance in the project and Mrs Trudy Machattie who willingly came to my assistance on numerous occasions both in the laboratory and the computer room.

I am also grateful to the nurses of Chatsmed, RK Khan and St. Aidens Hospital who gave generously of their time and interest in volunteering for this study; Miss Eleanor Gouws (Institute for Biostatistics of the Medical Research Council) and Professor Clark (Department of Statistics, University of Natal, Pietermaritzburg) for their assistance in the statistical analysis of the data; Dr Maritha Kotze (Faculty of Medicine, Department of Human Genetics, Tygerberg) for supplying me with controls of known apo E genotypes; Dr B Williamson (St. Mary's Hospital Medical School, London) for the DNA extraction method used in this study; Sara Serjeant and all the staff of Medical Illustration Unit, Faculty of Medicine, University of Natal for preparing the illustrations in this thesis; the staff of the Department of Chemical Pathology for help and encouragement.

I would especially like to express my sincere appreciation to my family and friends, and my fiance' and his family, for their love, support and the sacrifices they have made to bring me to this point in my career.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

ABBREVIATIONS

apo	apolipoprotein
BMI	body mass index
bp	base pair
CHD	coronary heart disease
°C	degrees celsius
DMSO	dimethyl sulphoxide
DNA	deoxyribonucleic acid
dNTPs	deoxynucleotide triphosphates
g	gram
g	gravitational force
HDL-C	high density lipoprotein cholesterol
HLP	hyperlipoproteinaemia
IEF	isoelectric focusing
Kb	kilobase
LDL-C	low density lipoprotein cholesterol
mRNA	messenger ribonucleic acid
mg	milligram
ml	millilitre
mmol/l	millimole per litre
min	minute
MW	molecular weight
nm	nanometre
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
sec	second
μg	microgram
μl	microlitre
μmol/l	micromole per litre
UV	ultra-violet
VLDL	very low density lipoprotein
v/v	volume per volume
WHR	waist:hip ratio

NOMENCLATURE

The nomenclature of apo E genotypes and phenotypes used in this text were proposed by Zannis *et al.*, (1982).

The alleles are designated $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$.

The isoforms are designated apo E2, apo E3 and apo E4.

The apo E phenotypes are designated as follows:
E2/2, E2/3, E2/4, E3/3, E3/4 and E4/4.

The apo E genotypes are designated as follows:
$\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 2/4$, $\epsilon 3/3$, $\epsilon 3/4$ and $\epsilon 4/4$.

'Lipid profile' is defined to include lipids, lipoproteins and apolipoproteins.

CHAPTER 1

INTRODUCTION

Lipoproteins play a central role in the development of atherosclerotic cardiovascular disease in humans. The plasma concentrations of lipoproteins and their metabolic fates are modulated by apolipoproteins on the surface of these lipid-rich particles. Genetic variation in apolipoproteins may determine part of the inter-individual variation in susceptibility to coronary heart disease (CHD). In this respect, apolipoprotein (apo) E, which is responsible for the elimination of potentially atherogenic chylomicron remnant particles (Mahley, 1988), is of particular interest.

Apo E is a glycoprotein constituent of all lipoproteins other than low density lipoprotein cholesterol (LDL-C). It plays a regulatory role in lipid metabolism, inter alia, by mediating the cellular uptake of apo E-bearing lipoproteins, notably chylomicron remnants, by two classes of receptors, an apo E-specific chylomicron remnant receptor and the LDL receptor. Apo E is mainly responsible for the hepatic uptake of dietary cholesterol via the chylomicron remnant receptor, and possibly for the removal of endogenous cholesterol contained in very low density lipoprotein (VLDL) remnants as well as for the removal of excess cholesterol from peripheral tissues through the hepatic clearance of apo E containing HDL (Davignon *et al.*, 1988). This makes apo E a central protein in directing the traffic of plasma lipoproteins and especially in the hepatic clearance of cholesterol-rich lipoproteins.

The human apo E gene has been mapped to chromosome 19 (Des *et al.*, 1985). Genetic polymorphism of apo E is controlled by three common (ϵ2, ϵ3, ϵ4) and several rare alleles at the apo E structural gene locus (Utermann, 1987). Apo E allele frequencies vary significantly between different ethnic groups (Hallman *et al.*, 1991). The apo E2 and apo E4 isoforms each differ from apo E3 by a single amino acid substitution resulting in different receptor-binding affinities. Apo E4 has been

associated with an increase in LDL-C which in turn is associated with hypercholesterolaemia (Davignon *et al.*, 1988). In contrast apo E2 is associated with a decrease in LDL-C and possibly a lower risk of CHD. Despite this generally favourable effect, apo E2 homozygosity is often seen with type III hyperlipoproteinaemia (HLP), a rare dysbetalipoproteinaemia that is characterised clinically by xanthomatosis and severe atherosclerosis of the coronary and peripheral arteries (Utermann *et al.*, 1977; Rall and Mahley, 1992). Thus, the phenotypic expression of the different apo E alleles is modified by biologic, genetic and environmental factors.

In all population studies thus far, apo E3 has been the predominant isoform, but the relative proportions of the three isoforms have shown variation among populations. Several groups have demonstrated a heterogeneity of apo E genotype frequencies between populations, for example, a reduced frequency of the mutant (E2 or E4) isoforms in certain Asiatic populations compared with Caucasians (Hallman *et al.*, 1991). Given the well established differences in the incidence of cardiovascular disease between these populations apo E polymorphism may also contribute to inter-population differences with respect to cardiovascular disease rates, though it is likely that environmental factors, especially diet, are the main determinants of risk.

The role of sex hormones in the pathogenesis or prevention of atherosclerotic CAD has long been a subject of interest and controversy. Population studies, (Kuller *et al.*, 1990; Lobo, 1991), indicate that a higher level of HDL-C in women is probably a major reason for their lower CHD incidence and mortality as compared to men. It has been suggested that this difference is due to the effect of oestrogen in premenopausal women (Lobo, 1991; Stampfer *et al.*, 1991). Contrary to this apparent anti-atherogenic role of oestrogens in females, in males low plasma testosterone and an increased oestrogen-to-testosterone ratio may be important risk factors for myocardial infarction (Sewdarsen *et al.*, 1990).

CHD is a major health problem in migrant Indian populations throughout the world. This susceptibility to CHD in expatriate Indian communities is believed to be due, in part, to a plasma lipid phenotype characterised by low levels of high density lipoprotein cholesterol (HDL-C) and by raised triglyceride and moderately elevated cholesterol levels. The Indian population of South Africa has a high rate of CHD (Seedat *et al.*, 1990). Studies by Sewdarsen and coworkers (1990, 1991) have focused on lipid and lipoprotein abnormalities in South African men. They showed that disturbances in sex hormones in young men are significantly associated with myocardial infarction. Given the high incidence of CHD, and the fact that associations in men cannot be assumed to hold for women, it is desirable that the determinants of the lipid profile and CHD in Indian women be studied.

The intention of the present study was to determine the gene frequencies for the apo E alleles in the Indian population, and to gain further insight into the influence apo E polymorphism on plasma lipid and lipoprotein levels, particularly as they affect women. The basic experimental design was a cross-sectional, population-based study on premenopausal and postmenopausal Indian nurses, working at three metropolitan hospitals, who were in good health and free from major identifiable risk factors for atherosclerosis. In addition to establishing an apo E genotyping method, the study was intended to answer chiefly the following questions:

i) the allele frequencies of the different apo E variants in the local adult Indian female population,

ii) the effect of apo E genotype on the lipoprotein profile in the study population;

iii) the influence of the menopause and other biologic variables on the lipid profile and possible interactions with apo E alleles.

CHAPTER 2

LITERATURE REVIEW

LIST OF CONTENTS **PAGE**

Apo E is involved in lipoprotein and lipid metabolism and this in turn influences the development of CHD. This chapter reviews lipoprotein metabolism, the relation of plasma lipoproteins to the development of CHD, the role of apo E and its genetic polymorphism in these processes and the methodology used to investigate some aspects of this relationship.

2.1 LIPOPROTEIN METABOLISM

The major physiological function of the lipoproteins is that of transporting dietary and/or endogenously synthesised lipids (cholesterol and its esters, triglycerides and phospholipids) in the plasma. All lipids are relatively insoluble in water and are transported in the plasma as micelles, solubilised by polar apolipoproteins, phospholipids, and free cholesterol. The proteins known as apolipoproteins function both to solubilise the lipids and also to interact with enzymes and cellular receptors which control the metabolism of the lipids. Lipoproteins are classified according to their density, composition, and characteristics of the main lipoprotein classes as follows (Durrington, 1989; Kroon and Powell, 1992):

i) Chylomicrons

These are large, triglyceride-rich complexes, formed from dietary lipids and intestinal apolipoproteins in the small intestine which eventually reach the systemic circulation via the thoracic duct. They transport exogenous lipid from the intestine to all cells. Although these particles contain very little protein, their apolipoprotein components especially apo B-48, are essential to their synthesis and secretion. Upon entry into the systemic circulation chylomicrons acquire apo E and apo C peptides from HDL.

ii) Very low density lipoproteins (VLDL)

These are moderately large particles, with triglyceride as the main lipid component and apo B-100 the main apolipoprotein but also including cholesterol, apo C and E peptides. They are formed in the liver and their main function is to transport

endogenously derived lipids from the liver to extrahepatic cells for storage and\or energy utilisation.

iii) Intermediate density lipoproteins (IDL)

These are transient intermediates formed during the conversion of VLDL to LDL. IDL may undergo internalisation by the liver as a result of the interaction with the LDL (apo B\E) receptor on the hepatocyte.

iv) Low density lipoproteins (LDL)

These are relatively smaller cholesterol-rich particles mainly derived from catabolism of VLDL (hydrolysis of triglyceride and removal of apo C peptides and apo E). LDL serves as a major cholesterol carrier to peripheral tissues where they are taken up by specific B\E receptors (LDL receptors). Most LDL particles are cleared by the liver rather than peripheral tissues.

v) High density lipoprotein (HDL)

HDL are the smallest and most dense of the lipoproteins. They are involved in mediating the reverse transport of cholesterol from cells in peripheral tissues to the liver for excretion from the body. The major HDL apolipoproteins are apo A-I and apo A-II. Apo E is secreted with hepatic HDL but ultimately redistributed within the HDL density range and probably to other lipoprotein fractions as well. HDL enters the hepatocyte via the LDL (apo B\E) receptor for which the E interaction appears to be the major component. Uptake may be inhibited by high plasma concentrations of VLDL which compete for the LDL receptor.

The lipid transport system can be divided into an exogenous and an endogenous cycle. The chylomicrons represent the exogenous lipid pathway. Most of the VLDL is assembled from endogenously synthesised lipids by the liver. The conversion of VLDL

to IDL and LDL and their subsequent receptor-mediated catabolism constitute the endogenous lipid pathway. Both pathways interact with HDL.

2.1.1 Exogenous lipid pathways

This pathway involves the transport of dietary lipids from the intestine to the liver and peripheral tissues. After absorption from the intestine, dietary triglyceride and cholesterol is secreted into lymph as chylomicrons, which carry apo B-48 as their structural protein. In plasma, an exchange of apolipoproteins between chylomicrons and HDL takes place, and the amount of apo E and apo C peptides on chylomicrons is increased. The plasma apo C-II activates the enzyme lipoprotein lipase (LPL), which is present on endothelial cells. This enzyme hydrolyses triglycerides from the chylomicron core, and the liberated fatty acids are taken up in adipose tissue and muscle where they are re-esterified to triglycerides and stored for energy reserve. Upon hydrolysis of the triglyceride, the chylomicrons become smaller and more enriched in cholesterol and are referred to as chylomicron remnants. Chylomicron remnants are rapidly cleared from the plasma by the liver, where the cholesterol is either used in membrane or lipoprotein biosynthesis or excreted as free cholesterol or bile acids. Apo E is the recognition site on chylomicrons responsible for the receptor-mediated removal of chylomicron remnants by the liver. The receptor responsible for remnant uptake is referred to as the or apo E (chylomicron remnant) receptor, though an unknown proportion may also be cleared through the LDL receptor.

2.1.2 Endogenous lipid pathways

In this pathway lipids are transported from the liver to peripheral tissues. In the liver, endogenous lipoproteins are synthesized, especially VLDL containing apo E and apo B-100 as the structural protein. These triglyceride-rich particles undergo a similar catabolic cascade through the influence of LPL, with apo C-II acting as cofactor. The end product of LPL action is IDL, a VLDL remnant which is depleted of triglycerides, phospholipids and apo C peptides but retains substantial amounts of apo E and all of its apo B-100. This may either be taken up by the liver cells through the LDL (apo

B/E) receptor that recognise their apo E components, or progressively further catabolised to LDL (more cholesterol rich), which only contains apo B-100 as apolipoprotein. Further uptake and metabolism of LDL by cells are mediated via the LDL receptor where apo B-100 is the ligand responsible for LDL binding to the LDL receptor.

2.1.3 HDL and reverse cholesterol transport

The route of excretion of cholesterol is in the bile. Thus, cholesterol synthesised in extrahepatic tissues has to be removed and transported to the liver, to prevent accumulation of cholesterol in these cells. A process referred to as "reverse cholesterol transport" was postulated several years ago (Glomset, 1968) and was envisioned to involve HDL as the means of transport. The enzyme lecithin cholesterol acyltransferase (LCAT) is important in this transport through its action in the removal of free cholesterol on the HDL surface to reappear as cholesterol ester in the core of the HDL particle.

The transfer of cholesterol esters to IDL, LDL and chylomicron remnants is mediated by the cholesteryl ester transfer protein (CETP). These lipoproteins are finally trapped and degraded in the liver by the LDL receptor-mediated and the chylomicron remnant pathway. In addition HDL also facilitate the uptake of cholesterol directly from peripheral cells. As the HDL become cholesterol-enriched, they acquire apo E. As shown by *in vitro* studies, the presence of apo E actually facilitates the acquisition of cholesterol by HDL (Mahley, 1988). The cholesterol-rich HDL with apo E formed in the interstitial fluid of various tissues transport cholesterol from peripheral tissues to the liver, where they are taken up by the LDL and apo E receptors. Typical HDL lacking apo E do not bind to the LDL receptors (Mahley and Innerarity, 1983; Davignon *et al.*, 1988).

2.2 PLASMA APOLIPOPROTEINS

Apolipoproteins fulfil two distinct metabolic functions: (i) they act as essential cofactors for enzymes such as LPL and LCAT; and (ii) they mediate and modulate the interaction of lipoproteins with cellular receptors. During the 1980s, the genes for ten apolipoproteins were isolated, sequenced and localised to specific chromosomes in the human genome (Table I, Pg 11). The structure-function relationships of the apolipoproteins involved in lipid metabolism are discussed in more detail below (Durrington, 1989; Breslow, 1992; Angelin and Rudling, 1992). These apolipoproteins are divided into families A, B, C, and E.

2.2.1 Apo A-I, A-II and A-IV

The principal apolipoproteins of HDL are the A apolipoproteins (A-I, A-II and A-IV). The gene for apo A-I is located on chromosome 11, where it occurs in close proximity to the genes for apo C-III and apo A-IV. Apo A-I is necessary for the synthesis and secretion of HDL. Apo A-I and apo A-IV are cofactors for LCAT activity. Apo A-I and apo A-IV can bind to cell surfaces and promote the efflux of cholesterol from peripheral cells providing substrate for the LCAT reaction. Apo A-II originates in the liver and intestine where the mature apo A-II contains 77 amino acids, but exists mainly as a dimer, the two molecules being linked by a disulphide bridge. A specific role of apo A-II has yet to be defined.

2.2.2 Apo B

Apo B is central to lipoprotein transport. Apo B-48 is synthesised in the gut and is smaller than the apo B-100 which is synthesised in the liver. Apo B-100 and apo B-48 are the product of one gene and apo B-48 is produced by a novel messenger RNA (mRNA) to yield a protein that only contains the amino-terminal half of apo B-100. This difference is important, since it is the carboxy-terminal that contains the LDL receptor binding domain of apo B-100. Apo B-100 mRNA contains 14121 nucleotides and is thus, the largest mRNA known. Chylomicrons and VLDL contain single apo B-48 and apo B-100 molecules respectively, which are essential for assembly of the

hydrophobic cores of these particles. Through linkage analysis of the apo B gene it has been shown that apo B is the disease locus in hypobetalipoproteinaemia where several variants of truncated or otherwise abnormal apo B molecules have been demonstrated. The familial defective apo B-100 mutation affects the receptor binding region of apo B-100 reducing the LDL-receptor-mediated elimination of LDL from plasma, resulting in an increase of plasma LDL levels (Angelin and Rudling, 1992).

2.2.3 Apo C-I, C-II and C-III

The C apolipoproteins are a group of peptides initially isolated from VLDL where they are most abundant although they are also constituents of chylomicrons and HDL. The genes for C-I and apo C-II, like those for apo E, are located on chromosome 19, whereas the gene for apo C-III is part of the gene cluster including apo A-I and apo A-IV on chromosome 11. Apo C-I as yet has no certain role in lipoprotein metabolism, while apo C-II plays a major role in lipoprotein metabolism through its activation of LPL. Thus, without apo C-II the triglycerides of circulating triglyceride-rich lipoproteins cannot be removed by lipolysis by tissues such as skeletal muscle and adipose tissue, which are their major sites of clearance. The autosomal recessive condition in which apo C-II is deficient produces hypertriglyceridaemia.

2.2.4 Apo E

Apo E binds to the apo E (chylomicron remnant) and the LDL receptors in the liver, thus, allowing for the normal elimination of chylomicron remnants from plasma. Apo E, initially termed the "arginine-rich apoprotein," was first identified as a lipoprotein constituent of VLDL in 1973 by Shore and Shore and was extensively characterized in several animal species after it was realized that dietary cholesterol altered its distribution in plasma. Considerable impetus to understanding the role of apo E in lipoprotein metabolism came from the observation of Havel and Kane (1973), that apo E-enriched β-VLDL accumulate in the plasma of patients with type III hyperlipoproteinaemia (HLP), a genetic disorder. This observation, in association with those obtained with cholesterol-fed animals, suggested that apo E played a key

role in cholesterol metabolism and specifically in the metabolism of chylomicron and VLDL. Utermann *et al.* in 1977 were the first to demonstrate that apo E is polymorphic. Polymorphic variation in the apo E gene is associated with quantitative variation in lipid and lipoprotein levels in the general population and contributes to CHD. This association is detailed below.

Table I: The functions of plasma apolipoproteins and chromosomal location

Apolipo-protein	Major sources	Lipoprotein association	Function	Chromosome location
A-I	Intestine, liver	HDL. chylomicrons	Activates LCAT	11
A-II	Intestine, liver	HDL, chylomicrons	Unknown (possible hepatic lipase cofactor)	1
A-IV	Intestine, liver	HDL, chylomicrons	Activates LCAT	11
B-48	Intestine	Chylomicrons	Chylomicron synthesis, secretion	
B-100	Liver	VLDL, LDL	VLDL synthesis, secretion; LDL receptor recognition	
C-I	Liver	Chylomicrons, VLDL, HDL	Unknown (possible activator of LCAT)	19
C-II	Liver	Chylomicrons, VLDL, HDL	Activates lipoprotein lipase	19
C-III	Liver	Chylomicrons, VLDL, HDL	Inhibits receptor mediated catabolism of triglyceride-rich lipoproteins	11
E	Liver	Chylomicrons, VLDL, IDL, HDL	Ligand for LDL receptor and chylomicron remnant receptor; lipid transport in nervous system.	19

2.3 APO E GENE STRUCTURE

The genetics and mode of inheritance of apo E, as first described by Utermann *et al.*, (1977) and later refined by Zannis and Breslow (1981), were pivotal in understanding the basis of apo E polymorphism. The apo E gene locus has multiple alleles that give rise to a protein polymorphism that was initially detected by isoelectric focusing (Utermann *et al.*, 1975). Three frequent and various rare alleles have been identified. The three major isoforms of apo E, termed apo E2, apo E3, and apo E4 are the products of three alleles $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$ respectively. Six common genotypes, three homozygous ($\epsilon 2/2$; $\epsilon 3/3$; $\epsilon 4/4$;) and three heterozygotes ($\epsilon 2/3$; $\epsilon 2/4$; $\epsilon 3/4$), result from the expression of any two of the three major alleles. Population studies have shown that the apo E3 is by far the most commonly observed form of the E apolipoprotein and therefore, represents the parent form of the apo E. The other common isoforms, apo E2 and E4, are considered mutant forms.

Apo E gene is located on the long arm of chromosome 19 (Des *et al.*, 1985) as part of a 50 Kb apolipoprotein gene cluster that includes apo C-I and apo C-II (Smit *et al.* 1988). The gene for the LDL receptor has also been mapped to this chromosome. The complete nucleotide sequence of the human apo E gene and its proximal flanking sequences are shown in Fig. 1 (Pg 13) (Paik *et al.*, 1985). Apo E gene consists of four exons and three introns which is similar to that of the apo A-I and apo C-II genes. The overall length of the apo E gene is 3597 nucleotides, which encodes a mRNA of 1163 nucleotides. The mRNA yields a primary translation product of 317 amino acids whose 18 amino-terminal signal peptide is co-translationally cleaved, giving rise to a 299 amino acid mature protein (relative molecular mass of 34 200 daltons).

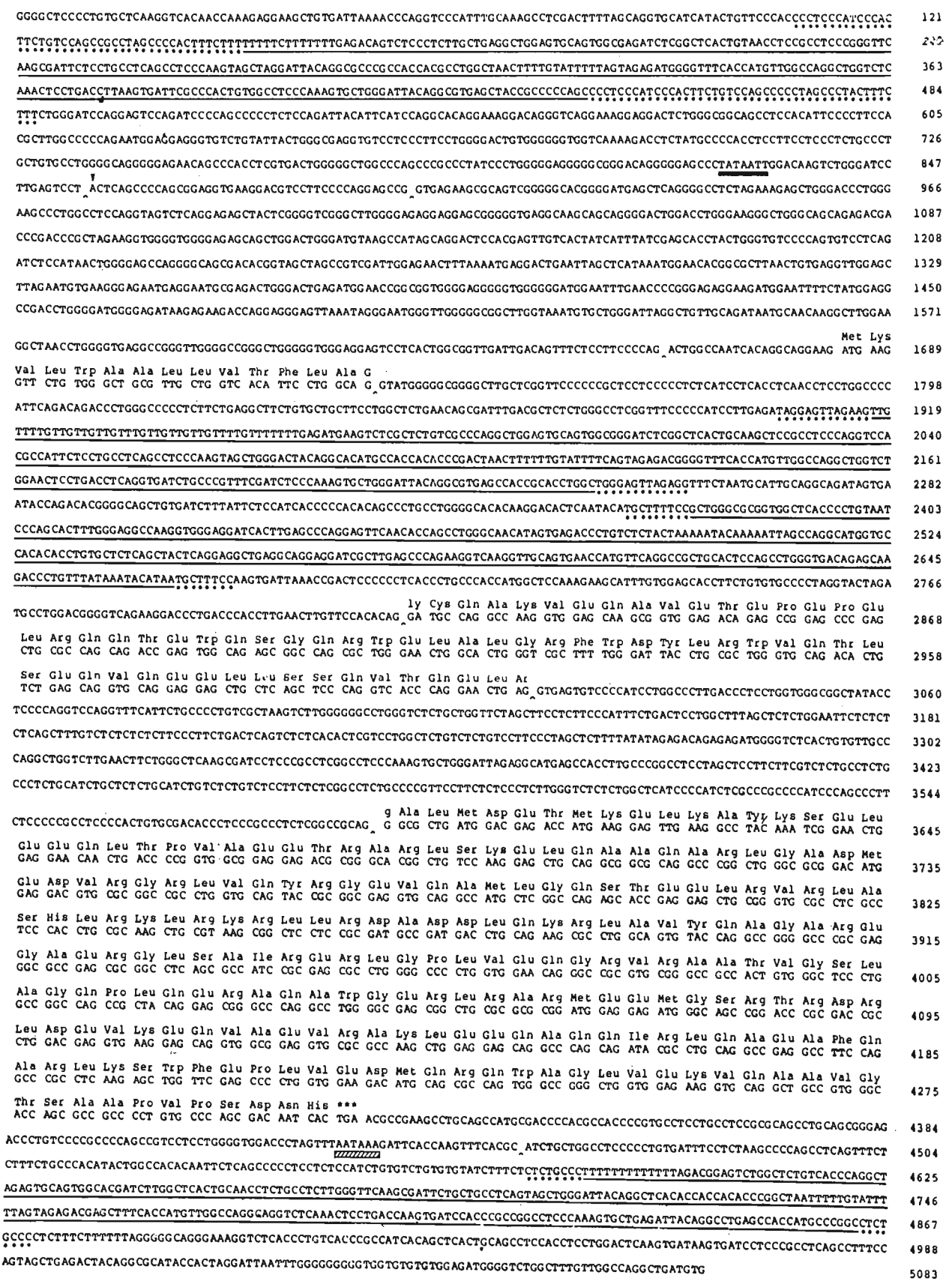


Fig. 1: Complete nucleotide sequence of the human apo E gene as well as the 856 nucleotides of the 5′ flanking region and 629 nucleotides of the 3′ flanking region are shown. Beginning and end of the gene, corresponding to the final mRNA product, as well as the exon-intron junctions are indicated by the marks (Λ) below the sequence. The major transcription initiation site is indicated at the beginning of the first exon (▾). The group of three asterisks indicates the translation termination codon in the fourth exon. The TATA box element in the 5′ flanking region (■) and locations of the *Alu* family members (—) and their flanking repeated sequences (....) are underlined (Paik *et al.*, 1985).

Detailed amino acid and sequence analyses of apo E have revealed that the major isoforms of apo E differ from one another by single amino acid substitutions at two different sites, namely 112 and 158, on the 299 amino acid chain of the mature apo E molecule. (Rall *et al.*, 1982). The amino acid substitution in these mutants involve arginine for cysteine interchanges and account for the known charge differences observed in isoelectric focusing. Apo E3 has cysteine at site 112 and arginine at site 158, while E4 has arginine and E2 cysteine at both sites (Fig. 2, Pg 14). Other rare apo E mutations have been reported and these will be discussed later (section 2.8, Pg 20).

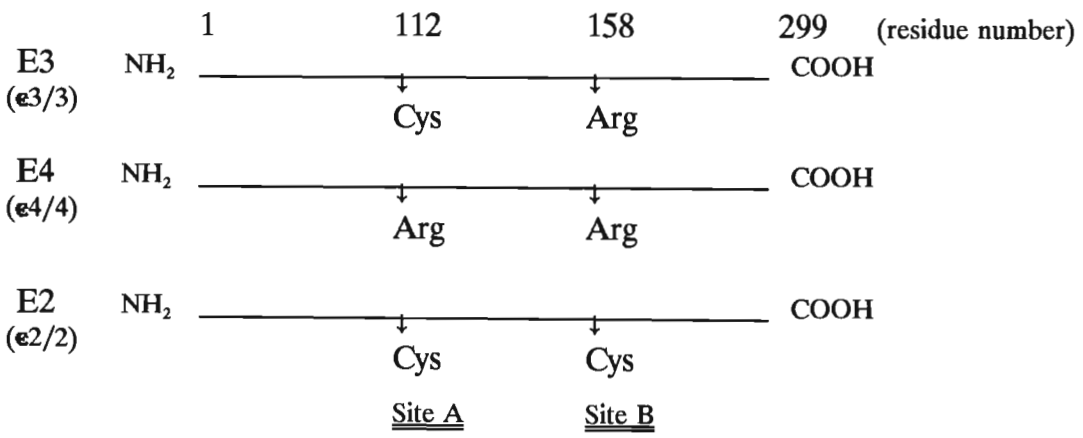


Fig. 2: Diagram illustrating the structural differences among the common apo E isoforms. The sequences of E2, E3 and E4 are identical except at site A (residue 112) and site B (residue 158).

2.4 SITE OF APO E SYNTHESIS

The largest quantity of apo E mRNA is found in the liver, which is the major source of apo E, and accounts for approximately two-thirds of the plasma apo E. The second largest concentration of mRNA is found in the brain (Mahley, 1988). Significant quantities of apo E mRNA are also detected in the spleen, lung, adrenal, ovary, kidney, and muscle in several different species. One notable exception is the epithelium of the intestine, which does not appear to produce apo E. The participation of several cell types in the widespread production of apo E indicates the

importance of apo E in lipid transport and possibly in roles unrelated to lipid transport.

2.5 FUNCTION OF APO E IN LIPID TRANSPORT

Apo E plays a central role in lipid homeostasis as it serves as a ligand for the apo E (chylomicron remnant) and the LDL (apo B/E) receptor, thereby enabling the hepatic uptake of VLDL, IDL, chylomicron remnants, and a subclass of HDL (Mahley and Innerarity, 1983; Mahley, 1988; Davignon *et al.*, 1988). The region of apo E spanning residues 140-160 binds to the LDL receptor. The molecular conformation in this region is important for normal receptor binding, but substitutions outside this immediate region also could have an effect on the binding domain by altering the conformation of the protein. The receptor binding of apo E may represent an ionic interaction between the basic residues in this region and acidic residues of the ligand binding domain of the LDL and apo E (chylomicron remnant) receptor (Mahley and Innerarity, 1983).

2.5.1 The LDL (apo B/E) receptor

Approximately 75% of LDL is removed from the circulation by the liver and most of this is mediated by the LDL receptor. The LDL receptors are found in fibroblasts, smooth muscle cells, leucocytes, adrenal cortex, testes, ovaries and the liver. Both apo B-100 and apo E contain a recognition site for the LDL receptor. The receptor binds specifically to LDL, HDL with apo E, chylomicron remnants and VLDL, containing either apo E or Apo B as the ligand (Mahley and Innerarity, 1983). The receptor-binding sites of apo B and apo E can be blocked by chemical modification such as methylation and glycosylation. Essentially this receptor functions to regulate LDL levels, redistribute cholesterol by apo B and apo E lipoproteins to various tissues and cholesterol utilization for membrane and hormone production. IDL binds to the LDL receptor with much higher affinity than LDL (Kroon and Powell, 1992). This has been attributed to the presence of multiple copies of apo E on each IDL particle, while each LDL particle contains only a single copy of apo B-100.

2.5.2 The apo E (chylomicron remnant) receptor

In addition to the LDL receptor (apo B/E) receptor, the liver also has the apo E (chylomicron remnant) receptor which recognises apolipoprotein E-containing lipoproteins (Mahley and Innerarity, 1983; Davignon *et al.*, 1988). This receptor is also referred to as the LDL receptor-related protein/α_2 macroglobulin (LRP/α_2-MR) and differs from the LDL receptor mainly by replacing the O-linked sugar domain in the LDL receptor by six EGF-like repeats (Schneider and Nimpf, 1993). Chylomicron remnants (activated by apo E enrichment) and α_2-macroglobulin (activated by serum proteases) both bind avidly to this molecule (Brown *et al.*, 1991).

The apo E (chylomicron remnant) receptor does not appear to be expressed by extrahepatic tissues and its major role is to remove from the circulation the remnants of chylomicron metabolism. There are several reasons for postulating the existence of this receptor; one reason is that individuals with familial hypercholesterolaemia-that is, patients with absent or defective LDL receptors, are capable of clearing chylomicron remnants from the plasma. This does not mean that the LDL receptor does not normally play a role in chylomicron remnant uptake. In vitro, chylomicron remnants do, in fact, bind with high affinity to the LDL receptors. The apo E (chylomicron remnant) receptor exists on the membranes of hepatic parenchymal cells, where it binds most avidly to chylomicron remnants, IDL and to an HDL subspecies rich in apo E.

This receptor does not bind strongly to newly secreted chylomicrons or to VLDL, despite the presence of apo E, which is believed to be masked from the receptor until significant lipolysis and triglyceride removal has occurred (Durrington, 1989). It has been suggested that, of the surface components lost during this process, the loss of apo C-III may be important for the exposure of apo E to the receptor binding site. The remnant receptor, unlike the LDL receptor, is not down-regulated as the intrahepatic cholesterol pool increases. The apo E (chylomicron remnant) receptor appears to

provide a rapid and efficient means of removal of chylomicron remnant particles until the hepatocyte is literally bursting with cholesterol.

2.6 OTHER FUNCTIONS OF APO E

In addition to its role in lipid transport, apo E is involved in the mobilisation and redistribution of cholesterol in repair, growth, and maintenance of myelin and neuronal membranes during development or after injury (Davignon *et al.*, 1988). The apo E synthesized and secreted by macrophages may convey the state of activation or inactivation of the macrophages to other cells within the local environment. Apo E inhibits mitogenic stimulation of lymphocytes by binding to specific sites on the surface of the lymphocytes.

It has been suggested that apo E may assist in the transport of cholesterol out of cells such as macrophages and thus, be important at an early stage in reverse cholesterol transport (Mahley, 1988). Also, the accumulation of cholesterol in macrophage may lead to the development of fatty streaks in arterial walls, which may go on to produce atheromatous plaques. Apo E is unique among apolipoproteins in that it has a special relevance to nervous tissue. Astrocytes contain apo E, which may be important in transporting lipids along their processes to the cells they nourish.

Apo E has been found to bind to the β-amyloid deposited in the brain of patients suffering from Alzheimer's disease, the degenerative brain disorder. An association between the apo E4 protein variant and late-onset Alzheimer's disease has recently been reported (Poirier *et al.*, 1993; Beardsley, 1993). Apo E4 is enriched in amyloid plaques and forms a preferential association over apo E3 with the β-amyloid peptide under oxidising conditions, indicating a specific role for apo E in Alzheimer's disease. The arginine in apo E4 at position 112 rather than the cysteine in apo E3 indicates a primary role for this structural change in the generation of this devastating disorder.

2.7 EFFECTS OF APO E POLYMORPHISM ON CHD

The apo E polymorphism has been shown to be an inherited determinant of total and LDL cholesterol (Kuusi *et al.*, 1989). Its status as a risk factor for CHD either through a casual relation with LDL cholesterol level or independently, is less clearly understood. Apo E polymorphism appears to affect the efficiency of cholesterol absorption and may by this mechanism contribute to the variation in plasma total and LDL cholesterol concentration (Davignon *et al.*, 1988). The association of elevated LDL cholesterol levels with CHD susceptibility is well documented and heightens the importance of understanding the influence of apo E genotype on LDL levels. Apo E is not a constituent of LDL and its influence must be indirect. Presently available data are suggestive of a beneficial influence of the $\epsilon 2$ allele on CHD, while the $\epsilon 4$ allele appears to predispose to CHD (Rall and Mahley, 1992; Eichner *et al.*, 1993).

This proposal is based on the under-representation of the $\epsilon 4$ allele in octogenarians, the over-representation in patients with documented cardiovascular disease (Cumming *et al* 1984; Kuusi *et al.*, 1989), and a tendency for subjects with the $\epsilon 4$ allele to experience myocardial infarction at a younger age (Eichner *et al.*, 1993). The cholesterol raising effect, in part due to more efficient absorption of dietary cholesterol, would tend to implicate apo E4 as an atherogenic isoform. Population studies (Ehnholm *et al.*, 1986; Hallman *et al.*, 1991) have demonstrated that apo $\epsilon 4/4$ homozygotes and $\epsilon 3/4$ heterozygotes indeed have significantly higher mean total cholesterol, LDL-C and apo B concentrations than matched apo $\epsilon 3/3$ subjects.

In normolipidaemic subjects with the $\epsilon 2/2$ genotype, which is present in about 1% of the population, the LDL cholesterol levels are reduced. The lowering of LDL may be a consequence of an increased expression of the hepatic LDL receptors secondary to the reduced uptake of cholesterol into chylomicrons. The cholesterol lowering effect of the $\epsilon 2/2$ homozygote has led to suggestions that the E2 isoform may offer a measure of protection against CHD. However, the cardioprotective effect can be corrupted by the presence of additional genetic defects or environmental factors that

result in increased production of triglyceride-rich lipoproteins and their consequent catabolic remnants, thus exposing the reduced receptor-binding capacity of the apo $\epsilon 2/2$ (Rall and Mahley, 1992). This diminished capacity is manifested in its extreme form as type III HLP, a dyslipoproteinaemia that is clinically characterized by xanthomatosis and severe atherosclerosis of the coronary arteries.

Hypertriglyceridaemia and hypercholesterolaemia in type III HLP individuals is caused by the accumulation of β-VLDL, particles that are thought to constitute remnants of chylomicrons, VLDL and IDL. More that 90% of type III HLP patients are homozygous for apo E2; however, only 2% of apo E2 homozygotes suffer from the disorder (Assmann *et al.*, 1991). If the apo E defect is the only problem, the accumulation is not of significant magnitude to raise plasma total triglyceride and cholesterol concentrations because apo B-100 is present on these lipoproteins as an alternative mediator of clearance of these particles via the LDL receptor. The extremely low degree of penetrance of type III HLP in $\epsilon 2/2$ subjects indicates that other genetic and/or environmental factors are necessary for expression of the hyperlipidaemia, such as familial combined hyperlipidaemia, age, obesity, hypothyroidism, diabetes and (in women) oestrogen status (Rall and Mahley, 1992).

Numerous studies on the effects of the apo E polymorphism on lipid metabolism have shown that individuals with at least one $\epsilon 2$ allele tend to have lower levels of total plasma cholesterol than do individuals who are homozygous for the $\epsilon 3$ allele. Individuals with at least one $\epsilon 4$ allele tend to have higher levels of total plasma cholesterol than do $\epsilon 3$ homozygotes (Davignon *et al.*, 1988). Despite the large differences in mean cholesterol levels among populations studied, the same tendency of apo E gene effects on plasma lipoprotein was observed with cholesterol concentrations increasing in the following order: 2/2, 2/3, 3/3, 3/4, 4/4 (Hallman *et al.*, 1991).

2.8 FUNCTIONAL DIFFERENCES OF APOLIPOPROTEIN E VARIANTS

Both common and rare variants of apo E have been described. As already discussed the common variants apo E2 and apo E4 have a significant impact on interindividual variation of lipid and lipoprotein levels in normal subjects. The common apo E2 mutant and a number of rare variants are defective in binding to the LDL receptor, and are causally associated with the lipid disorder type III HLP. The mutations involve mainly single amino acid substitutions, in which neutral amino acids are substituted for basic residues, and the substitutions occur in the vicinity of residues 140 to 160. Table II (Pg 22) summarises the nature of the mutations and their association with type III HLP.

Type III HLP is usually inherited as a recessive trait and is most commonly associated with the $\epsilon 2/2$ genotype. While homozygosity is generally but not invariably required for expression, not all $\epsilon 2/2$ subjects develop hyperlipidaemia. On the other hand, it is now clear that the dominant expression of the disorder occurs in individuals who have the Cys142, Cys146 and apo E_{Leiden} mutations, and that hyperlipidaemia occurs in subjects heterozygotes for the defective variants. These data reinforce the concept that cysteine substitution at residue 158 has a secondary effect on determining receptor binding and that the binding activity can be modulated by a variety of other environmental or genetic factors; the end result is variably defective apo E, ranging from very defective to nearly normal.

The great majority of rare apo E variants cause type III HLP in an autosomal dominant way with high penetrance. The clinical picture of these patients is not different from type III HLP patients homozygous for apo E2. In the literature patients with rare apo E variants include the mutation of apo E2 which takes place at the amino acid residue 146, whereby the amino acid lysine is replaced by glutamine, E2 (lys_{146}→gln). In addition there are other rare mutations like E3 (cys_{112}→arg; arg_{142}→cys); $E1_{Harrisburg}$ (lys_{146}→glu); E2 (arg_{145}→cys) and $E2_{Christchurch}$ (arg_{136}→ser) (Table II, Pg 22). Lohse *et al.*, (1991) reported apo $E4_{Philaidelphia}$ (Glu_{13}→lys; arg_{145}→cys) to be associated

with severe type III HLP. Wenham *et al.* (1993) identified the E2/E1 mutation which substitutes gly_{127}→asp which appears to be recessive with respect to the expression of type III HLP, although it may be somewhat more potent than the parent $\epsilon 2$ allele.

There is one situation in which type III HLP has been described that is not associated with a receptor binding-defective apo E variant. This condition is apo E deficiency, in which apo E is virtually undetectable in the plasma (Schaefer *et al.*, 1986). Affected subjects are homozygous for a mutation in the acceptor splice site of the third intron of the apo E gene, a mutation that leads to an abnormal and unstable apo E protein. These individuals have many of the usual characteristics of type III HLP, including premature cardiovascular disease, xanthomas, beta-VLDL (lacking apo E) and hyperlipidaemia.

Another case of apo E deficiency, also associated with type III HLP, has been described by Lohse *et al.* (1989), in which plasma apo E levels were less then 1% of normal. The molecular defect in this latter case is a nucleotide substitution (G→T) in the apo E gene that introduces a premature stop codon, resulting in a truncated 209 amino acid apo E protein. The association of type III HLP with apo E deficiency highlights the importance of apo E in the metabolism of triglyceride-rich lipoprotein remnants.

Table II: Known Apo E Genetic Variants

Apo E Isoforms	Gene Defects	Protein Defects	Associated with type III HLP	Mode of inheritance of type III HLP	Reference
E1	-	Gly_{127}→Asp Arg_{158}→Cys	Yes	Recessive	Rall and Mahley (1992)
$E1_{Harrisburg}$		Lys_{146}→Glu	Yes	Dominant	Mann *et al.* (1989)
E2E1	G→A	Gly_{127}→Asp	Yes	Recessive	Wenham *et al.* (1993)
E2	C→T	Arg_{158}→Cys	Yes	Recessive	Utermann and Steinmetz (1977)
E2*	C→T	Arg_{145}→Cys	Yes	Unknown	Rall and Mahley (1992)
E2**	A→G	Lys_{146}→Gln	Yes	Dominant	Smit *et al.* (1990)
$E2_{Christchurch}$	C→A	Arg_{136}→Ser	Yes	Unknown	Wardell *et al.* (1987)
E3*	T→C	Cys_{112}→Arg Arg_{142}→Cys	Yes	Dominant	Havel *et al.* (1983)
$E3_{Leiden}$	-	Tandem repeat of residues 121 to 127	Yes	Dominant	Wardell *et al.* (1989)
E4	T→C	Cys_{112}→Arg	No	-	Utermann (1987)
$E4_{Philadelphia}$	A→G C→T	Glu_{13}→Lys Arg_{145}→Cys	Yes	Unknown	Lohse *et al.* (1991)
E5	A→G	Glu_{13}→Lys	No	-	Rall and Mahley (1992)
E7	G→A	Glu_{244}→Lys Glu_{245}→Lys	No	-	Tajima *et al.* (1989)
$E3_{Washington}$	-	Premature stop codon in apo E gene	Yes	Unknown	Lohse *et al.* 1992

2.9 FEMALE SEX HORMONES AND CHD

The relatively lower incidence of CHD and death in young and middle-aged women than in men is generally attributed to a protective effect exerted by female sex hormones (Bush *et al.*, 1988). Epidemiological evidence suggests an association between natural or surgical menopause and an increased risk of coronary heart disease. The only recognised distinct effect the menopause has on standard cardiovascular risk factors is on lipids and lipoproteins which are major determinants of CHD (Jensen *et al.*, 1990). It is well established that LDL cholesterol increases progressively after menopause. This increase in LDL cholesterol levels in postmenopausal women is greater than that in age-matched men and has been related to low levels of oestrogen (hypoestrogenism) (Bush *et al.*, 1988). Both exogenous and endogenous sex hormones affect plasma lipoprotein metabolism profoundly and, in turn, influence atherosclerosis and CHD (Moreno and Manson, 1993).

Changes in lipoproteins after the menopause are due to a combination of aging, hormonal changes, and behavioral factors such as weight gain, exercise, smoking, and alcohol consumption. Jensen *et al.*, (1990), reported in a longitudinal examination that an increase in total cholesterol, triglycerides and LDL-C as a consequence of the menopause occurred within 3-6 months of cessation of the menses. The effect of menopause on HDL-C are inconsistent. The Framingham study (Kannel *et al.*, 1976) and the Lipid Research Clinics Prevalence study (Bush *et al.*, 1984) reported no differences in the HDL-C as a consequence of the menopause, whereas, the longitudinal study by Jensen *et al.* (1990) showed a significant decrease in HDL-C occurring gradually over the 2 years preceding the cessation of menses. The controversy might be explained by the weak ability of a cross-sectional design to detect the evidently small and gradual changes in HDL-C that occur around the menopause.

2.10 CHD IN SOUTH AFRICAN INDIANS

CHD is a major problem in migrant Indians (Asians) throughout the world (Editorial 1986). Despite their great cultural and geographical diversity, migrant Indians have a high mortality from CHD in comparison with other ethnic groups locally. The Indian population in South Africa had emigrated from India and settled in this country

since 1860. It is now well established that the Indian population in South Africa has a high rate of CHD that is now reaching "epidemic" proportions (Wyndham, 1982; Seedat *et al.*, 1990; Sewdarsen *et al.*, 1990; 1991).

The important risk factors in South African Indian men are hypercholesterolaemia, hypertriglyceridaemia, diabetes, smoking (Seedat *et al.*, 1990; Sewdarsen *et al.*, 1991) and abnormalities in sex hormones (Sewdarsen *et al.*, 1990). The important risk factors for women are diabetes, hypercholesterolaemia and hypertriglyceridaemia. The risk factors leading to CHD in Indians are similar to the Afrikaans-speaking whites, except that diabetes is common in the Indian population.

Seedat and coworkers (1990) reported a highly significant association between diabetes and CHD history in the South African Indian population. A high prevalence of diabetes has been reported in migrant Indians (both male and female) in South Africa (Omar *et al.*, 1985) as well as in migrant Indians in Fiji, Trinidad, Singapore, and London. The prevalence of diabetes amongst Indians in India was initially believed to be low but recent reports have shown that it is now comparable to the high prevalence seen in migrant Indian populations (Seedat *et al.*, 1990). Despite a high rate of CHD in South African Indian women (Seedat *et al.*, 1990), little information is available about the distributions of lipid concentrations and the frequency of hyperlipidaemia in this population group.

2.11 APO E POLYMORPHISM IN DIFFERENT ETHNIC GROUPS

Although the association of the $\epsilon 2$ allele with lower total and LDL cholesterol and the $\epsilon 4$ allele with higher total and LDL cholesterol in the general population is well established, there appears to be considerable heterogeneity with respect to the magnitude of the allelic effects estimated in different populations. Hallman *et al.* (1991), attempted to evaluate the allelic effects of the apo E polymorphism in nine populations (Tyrolean, Sudanese, Indian, Chinese, Japanese, Hungarian, Icelandic, Finnish and Malay) of different ethnicity from widely varying locations, where the above mentioned associations were observed. The effect of the $\epsilon 2$ allele on

cholesterol levels was negative in all populations with a mean value of -0.37 mmol/l, while the effect of the $\epsilon 4$ allele was positive in all populations except for Malays, with a mean value, over all populations, of 0.21 mmol/l.

The studies on apo E genes in relation to plasma lipids in Germans and Finns had suggested that apo E gene effects are independent of ethnic and cultural backgrounds; however, studies extended to a population sample including Japanese, Chinese, Indians, and Malays from Singapore modified this view. The effect of the $\epsilon 2$ allele was seen to the same extent in all populations studied. Plasma total cholesterol in $\epsilon 2$ homozygotes and heterozygotes was lower that in $\epsilon 3/3$ homozygotes regardless of the ethnic and cultural background of the population. On the contrary, the association of the $\epsilon 4$ allele with high cholesterol was highly significant only in Finns, moderate in Germans, but less significant or absent in the Japanese and Singapore populations. Hence, the effect of E4 on plasma cholesterol and apo B is not independent of cultural and/or ethnic background.

Differences in dietary habits are of particular interest. Dietary habits also influence plasma lipid levels as does apo E polymorphism, and some studies suggest an interaction between dietary habits and apo E isoforms in this respect (Tikkanen *et al.*, 1990), while others do not (Xu *et al.*, 1990). Hallman *et al.* (1991) concluded that the effects of the apo E polymorphism are not different across populations and therefore, there is no significant interaction between apo E genotype and dietary cholesterol (reflected in mean plasma cholesterol levels in the sample). However, Manttari *et al.* (1991), found evidence of interaction between apo E and diet in a large study of men in the Helsinki Heart study.

Carriers of the $\epsilon 4$ allele appear to be more sensitive to diet than other individuals. In particular, $\epsilon 4$-carriers respond significantly to the relatively high dietary intake of saturated fat and cholesterol typical of northern Europe (Manttari *et al.*, 1991; Xu *et al.*, 1990). Such an interaction may help to explain more of the interpopulation variability in CHD -incidence rates related to apo E polymorphism. It is not known

at present why Finns exhibit a strong association of the E4 allele with high cholesterol concentrations, whereas the Singapore population does not (Hallman *et al.*, 1991). One hypothesis is that these populations differ significantly in nutritional habits (e.g., fat intake) and that individuals with the E4 allele may acquire high cholesterol levels only while on a high-fat diet. Fat intake is high in Finns, not quite so high in Germans, and low in Japanese and in the population of Singapore. Hence the association of the $\epsilon 4$ allele with high cholesterol in Finns may be an example of an ectogenetic interaction. Furthermore, dissimilarities in apo E allele frequencies among Caucasian populations are comparable to dissimilarities between some Caucasian and Asian populations (Gerdes *et al.*, 1992). Notably, the frequency of the $\epsilon 4$ appears to be higher in the Northern regions of Europe (Scotland, Germany, and the Netherlands) than in southern regions (Switzerland, Tyrol, France, Italy and Spain) (Eto *et al.*, 1986; Davignon *et al.*, 1988; Gerdes *et al.*, 1993)

Hence, apo E genes contribute to the normal variance of plasma lipids and apolipoproteins within populations. The influence of apo E alleles may vary between populations depending on genetic and dietary factors prevalent within different people. Because apo E allele frequencies vary considerably between different ethnic groups, the apo E gene locus may be a significant genetic factor contributing to differences in lipid levels between populations.

2.12 SCREENING STRATEGIES FOR IDENTIFICATION OF APOLIPOPROTEIN E MUTANTS:

The apo E genotypes can be identified by a number of phenotyping and genotyping techniques. Traditionally, apolipoprotein mutants have been identified at the protein level using electrophoretic techniques. DNA-based analytical techniques, especially PCR, have replaced some of the more time-consuming laboratory techniques and have greatly aided the analysis of mutations in apolipoprotein genes.

2.12.1 Phenotyping of the apo E polymorphism

Originally the apo E polymorphism was detected by the separation of VLDL

apolipoproteins using isoelectric focusing (IEF) (Utermann *et al.*, 1975). This method is based on charge differences between the apo E isoproteins. The isolation of VLDL by ultracentrifugation and subsequent delipidation makes this technique very labour intensive. It is known from several studies that apo E dissociates from lipoprotein particles upon shear stress induced by ultracentrifugation and hence, only apo E tightly bound to lipoprotein is characterized after isolation (Steinmetz 1987). The use of chemical precipitation of VLDL removed the need for an ultracentrifuge, but both methods are laborious and require at least 1 ml of serum (McDowell *et al.*, 1989).

The need for a more rapid and reliable technique led to the application of immunoblotting to the analysis of the apo E polymorphism from unfractionated serum samples (Steinmetz, 1987). The combination of cysteamine treatment of apo E in plasma and isoelectric focusing of apolipoproteins with immunoblotting obviated the need for isolating apo E-containing lipoproteins or for two-dimensional gel electrophoresis of serum (Steinmetz 1987; McDowell *et al.*, 1989). Serum is digested with sialidase, delipidated, and redissolved in urea. Electrofocusing is carried out in agarose, followed by immunoblotting with a monoclonal antibody to apo E and an anti-immunoglobulin-peroxidase conjugate. This two-dimensional procedure is useful in detecting apo E mutants that differ in their cysteine residues or cofocus with other apolipoproteins and thus, escape detection by conventional one-dimensional techniques. Drawbacks result from the limited commercial availability of anti-apo E antibodies and sometimes poor specificity.

In addition the above techniques suffer from potential interference due to the post-translational addition of carbohydrate residues (both sialic acid and glucose) in the circulation. The sialated forms of apo E focus at more anodal positions which can lead to an apparent increase in the proportion of apo E2. It is therefore necessary to remove sialic acid residues by pre-treatment with neuraminidase before performing IEF. IEF is also time consuming and is often unreliable due to variation in isoform staining intensity and distance of migration during the electrophoresis procedure

(McDowell *et al.*, 1989).

With isoelectric focusing, only charge mutants whose total charge exceeds the limits of the E4 or E2 position on a focusing gel can be detected. For example, a mutation having the same charge shift as E2/2, would only lead to the misclassification of the isoprotein but not to its identification as a mutant on IEF. As IEF of VLDL apolipoproteins is based on charge differences between the apo E isoproteins, it is not possible to identify neutral amino acid substitutions nor to locate the site of the mutation. However, it is now possible to solve these problems by application of DNA typing methods.

2.12.2 Genotyping of the apo E polymorphism

The sequencing of the apo E gene by Paik and co-workers (1985) made it possible for apo E genotyping to be performed. Funke *et al.* (1986) developed a procedure to monitor for a sequence heterogeneity in the apo E gene using four oligonucleotides which allow the discrimination between a CGC triplet coding for arginine and a TGC triplet coding for cysteine in positions 112 and 158 of the apo E amino acid sequence, respectively. In the first few attempts at analysis, genomic DNA was digested with restriction enzymes such as *Apa*I, separated by electrophoresis in agarose gels, transferred to nitro-cellulose or nylon membranes by Southern blotting and hybridised with allele specific oligonucleotide probes (Funke *et al.*, 1986). The hybridisation of such prepared DNA with all four oligonucleotides can either be done one after the other with the same membrane in rehybridisation procedures or all at the same time with four different membranes. In the case where neither of the two alternative oligonucleotides made for each of the two polymorphic regions binds to the genomic DNA, the existence of at least one rare mutation within the 21-base region of the gene, complementary in sequence to the respective oligonucleotide, can be concluded (Main *et al.*, 1991; Assmann *et al.*, 1987). The invention of the polymerase chain reaction (PCR) simplified this process, making it possible to amplify only the region of interest in the apo E gene.

2.12.2.1 The polymerase chain reaction (PCR)

The PCR procedure is performed entirely *in vitro* without a need for cloning in bacteria. The method and principle have been extensively reviewed (Mullis and Faloon, 1987; Innis *et al.*, 1990; Blair and Zajdel, 1992). The principle of PCR is amazingly simple. It is closely patterned on the natural replication of the genetic material which occurs whenever a cell divides to form two new cells. It depends upon the fact that when DNA is heated to a certain temperature, its two component strands separate, and that when cooled down again, the strands come together or anneal. Extremely small amounts of nucleic acid, originating perhaps only from a few cells, are required for PCR and the purity of the nucleic acid need not be high.

To amplify nucleic acid, an enzyme called DNA polymerase is used. This enzyme is used *in vivo* by cells to make copies of DNA or RNA during replication, and it efficiently extends the nucleotides of the developing chain using the existing single strand of nucleic acid as a template. Furthermore, the polymerase has a property on which PCR technology capitalises: the polymerase fulfils its biological function only if it finds a "starting-block" - a short sequence of intact, double-stranded DNA. Synthetic sequences, known as primers, of single stranded DNA no more then 20-30 bases in length, which flank the fragment of interest, are used in this technique to mark the beginning of the target DNA on the two long single strands (Fig. 3, Pg 31). When these primers form molecular hybrids with their complementary strands, they are extended by the DNA polymerase. The extension of the newly developing nucleic acid chain will only proceed in a 5′ to 3′ direction and will thus extend from each of the 3′ ends of the template DNA to complete the complementary copy of the target sequence (see Fig. 3, Pg 31).

Therefore, by changing the temperature from 90-95°C for denaturation, to produce a single strand of DNA, dropping the temperature to 40-60°C for annealing of the primers and then raising it again to 67-72°C for the polymerase to extend the DNA strand, a cycle of replication of the target sequence can be achieved (Fig. 3, Pg 31).

The result is an exponential accumulation of the specific target fragment, approximately 2^n copies where n is the number of cycles of amplification performed. The PCR technique is now largely automated. This has been made possible by the development of the thermal cycler and the stability of the enzyme *Taq* DNA polymerase.

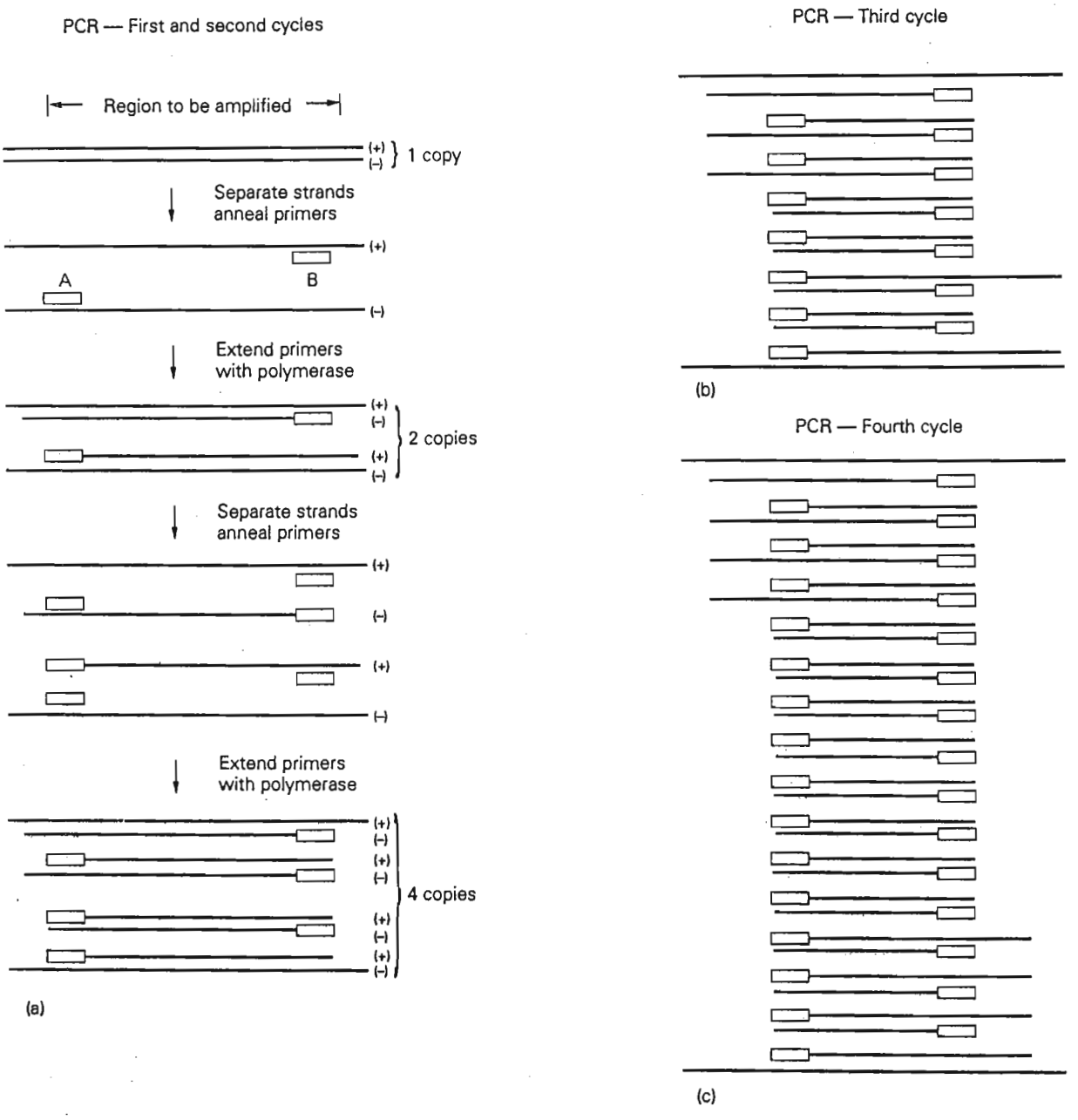


Fig. 3: General scheme of the PCR (Blair and Zajdel, 1992). A and B represent the two primers. DNA synthesis proceeds across the region between the primers.

2.12.2.1.1 Factors influencing the PCR

Because of the complex interactions amongst the components of a PCR it is unlikely that there ever will be one set of amplification conditions optimal for all situations. Minor adjustments to the PCR parameters often transform a marginal PCR into one with excellent specificity and yield. Hence, although the PCR is simple in principle, in practice the procedure may be complicated by several factors:

i) Magnesium Concentration

The composition of the buffer affects the yield and specificity of the PCR. In particular, the concentration of magnesium may influence primer annealing, strand separation, product specificity, formation of primer-dimers and *Taq* polymerase activity (Innis and Gelfand, 1990). The usual concentration of magnesium is 1.5 mmol/l but this may vary, in part according to the concentration of dNTPs in the reaction buffer: an excess of about 0.5 mmol/l magnesium concentration in relation to total dNTP concentration is generally desirable.

ii) Deoxynucleotide triphosphates (dNTPs)

The dNTPs (dATP; dCTP; dGTP and dTTP) are usually present in amounts of 50-200 μmol/l each in the reaction mix. The specificity and accuracy of PCR is increased by using lower concentrations of dNTPs (Innis and Gelfand, 1990). The dNTPs present at higher than optimal concentrations in the reaction mix may cause misincorporations by the *Taq* DNA polymerase.

iii) *Taq* DNA polymerase

Taq DNA polymerase, isolated from the heat stable bacterium *Thermus acquaticus*, has the ability to withstand repeated exposure to high temperatures. This simplifies automation and obviates the need for the addition of polymerase after each cycle as was previously necessary with the enzyme isolated from *E.coli.* The recommended concentration range for *Taq* DNA polymerase is between 1 and 2.5 units per 100 μl

reaction medium when other parameters are optimum. However, *Taq* DNA polymerase from various suppliers may behave differently because of individual formulations, assay conditions or unit definitions. Although the *Taq* DNA polymerase has a very limited ability to synthesise DNA at temperatures above 90°C, the enzyme is relatively stable and is not denatured irreversibly by exposure to high temperatures. In a PCR mix, *Taq* DNA polymerase retains about 50% of activity after 120 min, 40 min and 5 min at 92.5, 95 and 97.5°C, respectively.

iv) Oligonucleotide primers

The design of primers is crucial to the success of PCR reactions. They are usually approximately 18-30 nucleotides long with a guanine and cytosine content of not more than 50-60%. Primers are checked against each other for complementarity to reduce the formation of primer dimers. Primer concentrations between 0.2 and 1 μmol/l are generally optimal. Higher primer concentrations may promote mispriming, the accumulation of non-specific products and the generation of primer-dimers.

v) Other components

Some protocols include dimethyl sulfoxide (DMSO) in the PCR mix but this has been shown to inhibit the enzyme *Taq* DNA polymerase, while others advocate the addition of BSA and gelatin to stabilise the enzyme (Gelford and White, 1990). Addition of Triton-X-100 to a PCR is believed to enhance enzyme specificity (Blair and Zajdel, 1992).

vi) Temperature Cycle conditions

The three steps which make up the PCR, denaturation, annealing and extension, are extremely temperature dependent. When setting up a PCR reaction, the annealing and extension temperatures are varied alternately in 5°C steps until an amplified fragment is obtained following electrophoresis. The reaction can be "fine tuned" by varying the temperatures 1°C at a time.

The most likely cause for failure of an amplification is incomplete denaturation of the target template and/or the PCR product. Almost any temperature between 91°C and 96°C results in denaturation. However, as already mentioned after 30 cycles of denaturation at 96°C, there is much less enzyme activity remaining than after 30 cycles at 92°C.

The choice of annealing temperature in the PCR is a compromise. It has to be high enough to avoid non-specific primer binding to other areas of the DNA template and low enough to allow the primers to anneal to the desired template. The temperature and length of time required for primer annealing depend upon the base composition, length, and concentration of the amplification primers.

The *Taq* DNA polymerase functions at its optimum at 75°C (Gelford and White, 1990). It is therefore desirable for the extension temperature to be close to this. The extension temperature used in most PCR experiments is 72°C. The time allowed for extension depends upon the length and concentration of the target sequences and upon temperature of the reaction.

The average number of cycles for most reactions varies between 25 and 40. After a certain number of cycles the amplification reaction gradually declines and enters a linear or stationary phase. This is called the "plateau effect" and is influenced by exhaustion of substrates (dNTPs or primers), stability of enzyme, end product accumulation and incomplete denaturation at high product concentration (Innis *et al.*,1990).

2.12.2.1.2 Application of PCR to the genotyping of apo E.

The fact that the two variable nucleotides are located only 138 base pairs from each other renders the polymorphic area of the apo E gene an especially favourable target

for *in vitro* DNA amplification with the PCR. Smeets *et al.* (1988) used a procedure based on selective hybridization with allele-specific oligonucleotides for genotyping apo E variants from DNA. Following amplification, the DNA is immobilised onto nylon membranes by means of a slot blotting apparatus. The immobilised DNA is then hybridised with allele-specific oligonucleotides (ASOs) and the genotype determined.

However, the purchase of a slot blotting apparatus is expensive and raises the possibility of cross-contamination between runs unless a stringent and time-consuming procedure is used (Houlston *et al.*, 1990). Houlston *et al.*, (1990) found the use of agarose gel electrophoresis and the Southern blotting technique to immobilise the amplified DNA obviates the need to purchase specialised equipment and also offers little risk of cross-contamination, intrinsic in other techniques used to mobilise DNA. However, all hybridisation-based assays are time consuming and technically demanding. They may suffer from nonspecific background reactions and depend on the use of radioactivity.

These disadvantages have been overcome by the amplification refractory mutation system (ARMS) (Wenham *et al.*, 1991). Using oligonucleotide specific primers, the technique is based on the observation that if a primer is mismatched at the 3' nucleotide, amplification will not occur. In principle, the wild type primer amplifies only wild type sequence and the mutant primer amplifies the mutant sequence using specific primers in separate amplification experiments. The genotype is determined by agarose gel electrophoresis and obviates the need for hybridization with allele specific oligonucleotide probes.

In 1990, Hixson and Vernier reported restriction isotyping (restriction enzyme isoform genotyping) as a simpler and faster method for typing the common apo E isoforms. Restriction isotyping uses PCR amplification, but avoids the use of costly and time-

consuming hybridisation and sequencing techniques. It relies on amplification of discrete DNA regions followed by the use of restriction enzymes to identify the apo E polymorphism which is discussed in detail below.

2.12.2.2 Restriction enzyme analysis

Restriction endonucleases are found in bacteria where they are used to guard against foreign DNA. Restriction enzymes bind specifically to and cleave double-stranded DNA at specific sites within or adjacent to a particular sequence known as the recognition sequence and are used extensively *in vitro* to recognise a particular DNA sequence or gene mutation. The majority recognise sequences 4-6 nucleotides in length. Over five hundred different restriction endonucleases have been characterised and purified for the use in molecular biology (Brown, 1986).

2.12.2.2.1 Application to apo E genotyping

A very simple and reliable technique to detect the two variable nucleotides at codons 112 and 158 of the apo E gene is restriction-fragment length analysis of PCR amplified DNA developed by Hixson and Vernier (1990). This method takes advantage of the fact that nucleotide substitutions at these two loci generate restriction fragment length polymorphisms that can be detected with the enzyme *Hha*I. Because the nucleotide substitution that result in Arg-Cys interchanges at position 112 and 158 also alter *Hha*I cleavage sites, each genotype can be distinguished by unique combinations of *Hha*I fragment sizes in all homozygotic and heterozygotic combinations. The base sequence present around the nucleotide position 3745 and 3883 in exon 4 of the apo E gene creates the restriction enzyme site GCGC when Arg is coded for. *Hha*I recognises this sequence and cleaves as follows:

G C G ↓C
C ↑G C G.

The endonuclease cleaves at GCGC encoding Arg_{112} (E4) and Arg_{158} (E3, E4), but does not cut at GTGC encoding Cys_{112} (E2, E3) and Cys_{158} (E2). Fig. 4 (Pg 38) shows the *Hha*I cleavage sites (GCGC) in the amplified E4 sequence, including *Hha*I sites

at codons for arginine residues (GCGC) at positions 112 and 158. There are six recognition sequences for *Hha*I in the 244 bp amplified apo E4 product, four of them remote from the mutant sites. This results in digested fragments ranging from 16 to 91 base pairs.

The major advantage of this method is the simplicity of the detection method requiring only polyacrylamide gel electrophoresis, in which no radioactive materials are required. When comparing the *Hha*I restriction assay with the ARMS technique the restriction isotyping method has the advantage of being less expensive despite the longer time required for analysis (Assmann *et al.*, 1991).

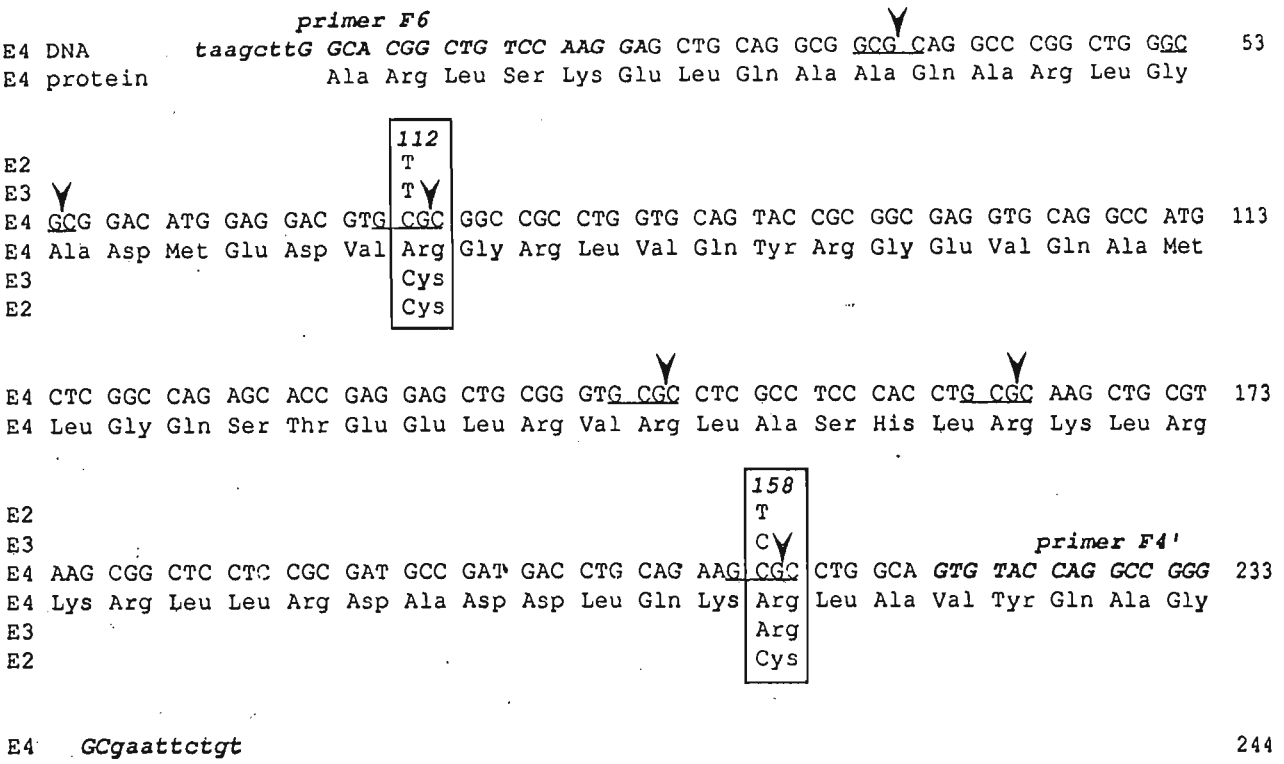


Fig. 4: DNA and protein sequences of amplified regions encoding common apo E isoforms and locations of *Hha*I cleavage sites (Hixson and Vernier, 1990). The amplified E4 nucleotide sequence (244 bp numbered to the right) is shown above the E4 amino acid sequence. The sequences of amplification primers (F6 and F4′, the reverse component of F4) are also shown (upper case italics are apo E sequences, lower case italics are synthetic cleavage sites). Nucleotide substitutions that distinguish E2 and E3 isoforms are shown above the E4 nucleotide sequence, and amino acid substitutions are shown below the E4 amino acid sequence (substitution sites at codons 112 and 158 are boxed). The sites for *Hha*I cleavage in the E4 nucleotide sequence are underlined and marked by arrows.

2.12.3 Apo E phenotyping verses genotyping

Most diagnostic laboratories determine apo E phenotype by isoelectric focusing (IEF) of plasma or of VLDL, followed by protein staining or immunoblotting (Steinmetz 1987; McDowell *et al.,* 1989). Experiments comparing different techniques to identify the common apo E isoform generally support genotyping by DNA analysis. Kontula *et al.*, (1990) compared the genotyping by restriction enzyme digestion to the slot-blot hybridisation technique using a pre-chosen panel of samples with six different genotypes. Although the latter system yielded identical results, it is technically more demanding, and even the slightest deviations of temperature during the final filter-washing step can produce erroneous results. They also found discrepant results between data from protein studies using IEF and genotyping by PCR followed by hybridisation allele specific oligonucleotide probes: a sample judged to represent an E2/3 phenotype by the former technique was classified as a $\epsilon 3/3$ genotype by DNA analysis.

Snowden *et al.,* (1991), compared apo E phenotypes (using IEF) and genotypes (determined by hybridisation with allele specific oligonucleotide probes) in 95 non-insulin-dependent diabetics and found discrepancies in 16%. They advocate the use of genotyping instead of phenotyping to identify apo E polymorphism. Studies employing phenotyping to identify the apo E polymorphism may in fact have over estimated the frequencies of the $\epsilon 2$ and $\epsilon 3$ alleles whilst underestimating those of the $\epsilon 4$ allele. Wenham *et al.* (1991) reported that IEF can lead to the erroneous assignment of apo E phenotype even after pretreatment with neuraminidase. The study used IEF and immunoblotting of delipidated plasma (McDowell *et al.*, 1989) to determine phenotypes. Genotypes were determined by the restriction isotyping method (Hixson and Vernier, 1990; Wenham *et al.*, 1991). Discrepancies between phenotype and genotype were observed in 13% of the diabetics and 17% of the non-diabetic controls. They suggested that genotyping by DNA analysis is the method of choice in determining apo E status. Hence, restriction isotyping may replace apo E phenotyping techniques for routine purposes, providing a simpler alternative method that does not involve hybridization techniques.

CHAPTER 3

PATIENTS AND METHODS

CONTENTS PAGE

3.1 INTRODUCTION

This chapter describes the study design as well as details of the protocol and the development and validation of the laboratory methods used. This study was designed to address the following questions:

1) Whether the frequencies of the common apo E alleles in the South African Indian population are similar to those in other populations?

2) What effect does apo E polymorphism have on the lipid profile in premenopausal and postmenopausal Indian women?

3.2 STUDY DESIGN

The present project was part of a larger study into the metabolic and clinical effects of the menopause in Indian women. The subjects were healthy nurses working at 3 major hospitals within the Durban metropolitan region. Possible sources of bias in selection of the study sample as a representative of the South African Indian population is discussed in Chapter 5. A cross-sectional design was adopted for reasons of practicality. The nursing community is a stable, educated and health conscious section of the population and is drawn from a moderately well-off social class. All these factors made voluntary participation easier to obtain, compliance more reliable and eliminated the impact of gross socio-economic inequality on the plasma lipoprotein profile. All participation was entirely voluntary and the protocol was approved by the Ethics committee of the Faculty of Medicine at the University of Natal.

3.3 STUDY PROTOCOL

3.3.1 Selection and exclusion criteria

The target group was Indian female nurses between the ages of 35 and 55 years working at the RK Khan, Chatsmed and St. Aidens hospitals in Durban. One hundred and ninety eight nurses volunteered to participate in the study of which twenty five women were excluded according to one or more of the following criteria:

i) surgically induced menopause (oophorectomy, hysterectomy or both),

ii) perimenopausal status with irregular cycles or borderline gonadotrophin levels.

iii) previous diagnosis of non-insulin dependent diabetes mellitus (NIDDM), hypertension, hepatic or renal disease, significant angina or myocardial infarction, hyperlipidaemia and hypothyroidism.

iv) significant recent illness, injury or surgical procedure.

v) taking oral contraceptive agents or hormone replacement therapy or any other medication known to influence lipid and lipoprotein metabolism.

The nurses were grouped into premenopausal and postmenopausal on the basis of their menstrual history and serum gonadotrophin levels. One hundred women were classified as premenopausal and seventy three women were postmenopausal on the following criteria:

i) absence of menses for a minimum of six months,

ii) elevated gonadotrophin (FSH > 40 IU/l) levels, and

iii) in some instances the presence of menopausal symptoms.

3.3.2 Clinical Interviews

The nurses filled out a questionnaire from which a detailed menstrual, medical, surgical, drug, and social history was obtained. Clinical and anthropometric measurements, including, body mass (kg), height (cm), waist (cm) and hip (cm) circumference and blood pressures (mmHg) were recorded by a clinician. In addition a smoking history, details of alcohol consumption and a family history of cardiovascular disease, diabetes mellitus and hypertension were also recorded.

3.3.3 Sample collection

After an overnight 14-hour fast, blood was collected by venepuncture with minimal venestatis into tubes (Vacutainer®,) containing disodium ethlyenediamine tetraacetate (EDTA) (final concentration 1 mg/ml) or without anticoagulant. The blood samples were centrifuged in a laboratory centrifuge at 2000 *g* for 20 min and the serum stored

at -70°C for subsequent analysis. These samples were batched to minimise long term analytical drift. The blood cells from the EDTA tubes were collected and stored at -20°C for subsequent DNA extraction. Prior to freezing, storage buffer A (20 mmol/l citric acid; 40 mmol/l sodium citrate; 80 mmol/l glucose) was added to the cells in a 6:1 ratio of cells/buffer A (v/v).

3.4 LABORATORY METHODOLOGY

3.4.1 Quantitative lipid and lipoprotein measurements

The lipids and lipoproteins were measured by standard laboratory methods. The assays were carried out using an automated, discrete, chemistry analyser, Technicon RA 1000® (Technicon Instruments Corporation, New York, USA). Total cholesterol was assayed by the cholesterol oxidase method (Siedel *et al.*, 1983) using a commercial kit (Monotest® CHOD-PAP, High performance, Enzymatic colorimetric method; Boehringer Mannheim GmbH Diagnostica). Triglyceride concentrations were determined by the enzymatic hydrolysis of triglycerides with subsequent determination of the liberated glycerol by colorimetry using a commercial kit (Triglycerides GPO-PAP® Boehringer Mannheim GmbH Diagnostica). HDL-C was measured in the supernatant, after the precipitation of apo B containing lipoproteins with dextran sulphate/Mg^{2+} (Warnick *et al.*, 1982). The same cholesterol method was used but sample volume was increased 2-fold to compensate for the low cholesterol concentration in the HDL sample. The interassay coefficients of variation for total cholesterol, triglyceride and HDL-C were 3.0%, 2.7% and 4.8%, respectively. LDL cholesterol was calculated using the Friedewald formula, adapted for the expression of results as mmol/l:

$$\text{LDL-C} = \text{total cholesterol} - (\text{HDL-C} + \text{triglyceride}/2.2)$$

Apo A and apo B were measured by an automated immunoturbidimetric method (Technicon®, USA). Calibrators, controls and patient samples were added to apo A-I and apo B antiserum (Technicon, USA) forming insoluble antigen-antibody complexes. The absorbance of the resulting solution was measured at 340 nm after a 5 min

incubation period.

3.4.2 Apo E genotpying

Polymorphism at the apo E gene locus was discussed in detail in Chapter 2, together with the different analytical approaches to genotyping. It was decided that the restriction isotyping method would be the most practical technique for conducting this population study to identify apo E gene defects. The apo E genotyping was performed without knowledge of the individual's lipid and lipoprotein levels.

3.4.2.1 DNA extraction

The classical method for DNA extraction is the initial removal of proteins and other impurities from a solution of nucleic acids by extraction with phenol/chloroform, followed by the precipitation of the DNA (Sambrook *et al.*, 1989). Apart from this method being both laborious and expensive, phenol is highly corrosive and consequently is cumbersome when extracting DNA from a large number of samples. Fortunately PCR is one of the few techniques currently employed in molecular genetics that can be performed on fairly impure DNA templates reducing the need for extensive purification.

Many extraction methods can provide representative DNA templates for PCR. The specific choice involves some trade-off between template quality and technical complexity. Several simple methods have been tried, such as lysis of cells in hypotonic media (Longmire *et al.*, 1987), and the direct introduction of unmodified biological fluid into PCRs (Mercier *et al.*, 1990). Although these methods have been reported to work in some instances, the PCR yield may be suboptimal because of poor release of amplifiable templates or the presence of inhibitory substances.

The DNA isolation procedure utilised in this study is a rapid and inexpensive boiling

method which avoids the use of the hazardous organic solvents, phenol and chloroform, as well as expensive proteinases such as protease K. This method yields DNA which has permitted consistently efficient amplification (Williamson, personal communication).

3.4.2.1.1. Protocol

In an eppendorf tube, 800 μl of freshly prepared 170 mmol/l NH_4Cl was added to 200 μl thawed EDTA-cells/buffer A solution. The solution was then mixed for 20 min by rotating in a rotary mixer and centrifuged for 2 min in a microfuge at 15 000 g to obtain a white pellet. The supernatant was removed and the pellet was washed carefully with 300 μl of a mixture of 10 mmol/l NaCl and 10 mmol/l EDTA (pH 7.5 with HCl). This washing step was repeated until all visible haem was removed. The white cell pellet was resuspended in 500 μl of 50 mmol/l NaOH by vortexing for 10 sec followed by boiling for 20 min in a water bath. The solution was neutralised with 100 μl of 100 mmol/l Tris HCl (pH 7.5) and vortexed for 5 sec, followed by centrifugation for 15 sec to remove cell debris. Quantitation of the DNA was not necessary since 8 μl of DNA prepared in this way was empirically found to be sufficient for a single PCR reaction. DNA was aliquoted into 100 μl amounts and stored at -20°C until use in the PCR.

3.4.2.2 Polymerase chain reaction (PCR)

The principle and the factors influencing the PCR are outlined in Chapter 2. In this chapter the protocol used to amplify the apo E gene and the precautions taken to prevent contamination of the PCR are dealt with.

3.4.2.2.1 Validation of the PCR for apo E amplification

i) Primer Design

Oligonucleotides were chosen which flank the two known mutation sites (on either side of nucleotide positions 3745 and 3883) in exon 4 (see Fig 5, Pg 46). This results in

amplification of the apo E gene irrespective of the particular genotype as explained in detail in Chapter 2. The primers used to amplify this region of the apo E gene were termed F4 and F6 as first described by Emi and coworkers (1988).

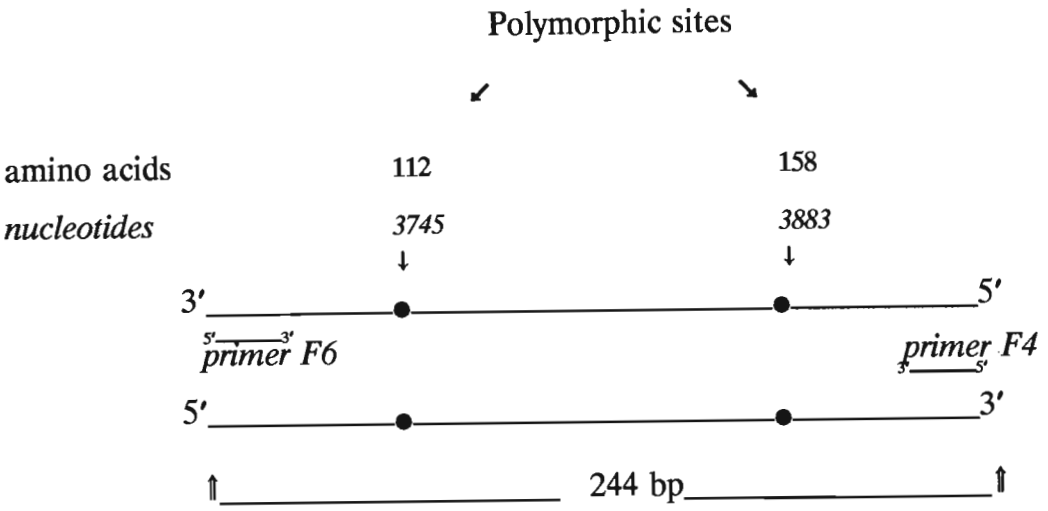


Fig. 5: The polymorphic region of the apo E gene amplified by PCR. ● denotes the polymorphic Cfo I cutting site. The PCR results in a 244 bp fragment irrespective of these polymorphic sites.

The primers have the following sequences:

F4 5′-ACAGAATTCGCCCCGGCCTGGTACAC-3′

F6 5′-TAAGCTTGGCACGGCTGTCCAAGGA-3′

Primer F6 is a sense strand sequence which attaches to the complementary anti-sense DNA strand running from 3′ to 5′ and primer F4 is an anti-sense sequence and attaches to the complementary sense DNA strand running from 5′ to 3′ (Fig. 5, Pg 46). The 3′- termini of the nucleotides are 3892 and 3699 for F4 and F6, respectively, according to the apo E sequence determined by Paik *et al.* (1985). Primer F4 is 26 nucleotides long and primer F6 is 25 nucleotides long. These primers contain artificially designed non complementary 5′ ends in which several A and T nucleotides have been substituted for G and C to avoid problems inherent in GC rich primers (Blair and Zajdel, 1992). Mismatches between the 5′ end of the primers and the

target DNA do not significantly affect the efficiency or the specificity of the amplification reaction. The primers used in this study were synthesised in a Milligen Oligonucleotide Synthesizer in the Department of Biochemistry, University of Cape Town.

ii) DMSO concentration

It has been reported that a 10% concentration of dimethyl sulphoxide (DMSO) inhibits DNA synthesis, mainly by reducing the thermal activity profile of *Taq* DNA polymerase by approximately 50% (Scharf *et al.*, 1986; Filichkin and Gelvin, 1992). DMSO has also been postulated to affect the melting temperature of the primers and the degree of product strand separation achieved at a particular "denaturation" or upper-limit temperature (Gelford and White, 1990). Filichkin and Gelvin (1992) found that, at least in some cases, a 10-fold lower concentration is sufficient to suppress mispriming and/or to initiate DNA amplification.

Due to the uncertainty regarding the requirement for DMSO in the PCR it was decided to validate its use in the amplification of the apo E gene in the Hixson and Vernier protocol. This was done in an experiment using DMSO in concentrations from 0 to 16% (Plate 1, Pg 49). No amplification products were observed at DMSO concentrations $<4\%$ and $>15\%$. Visual inspection of the gel suggested that a DMSO concentration of 10% gave maximal intensity of the amplification products with minimal primer dimer formation, thus confirming the work of Hixson and Vernier, (1990). A DMSO concentration of 10% was therefore chosen as optimal.

iii) dNTP

One of the experimental problems encountered in this study was the use and storage of dNTPs. PCR experiments using dNTPs that had been diluted and stored at -20°C showed no positive results, indicating that diluted dNTPs apparently undergo degradation on storage. Subsequently, on receipt from the manufacturer, the four

dNTPs (dATP, dTTP, dGTP and dCTP) were therefore individually aliquoted and stored at -20°C. Aliquots of each dNTP were pooled (at equivalent concentrations) and diluted to a working concentration just before use.

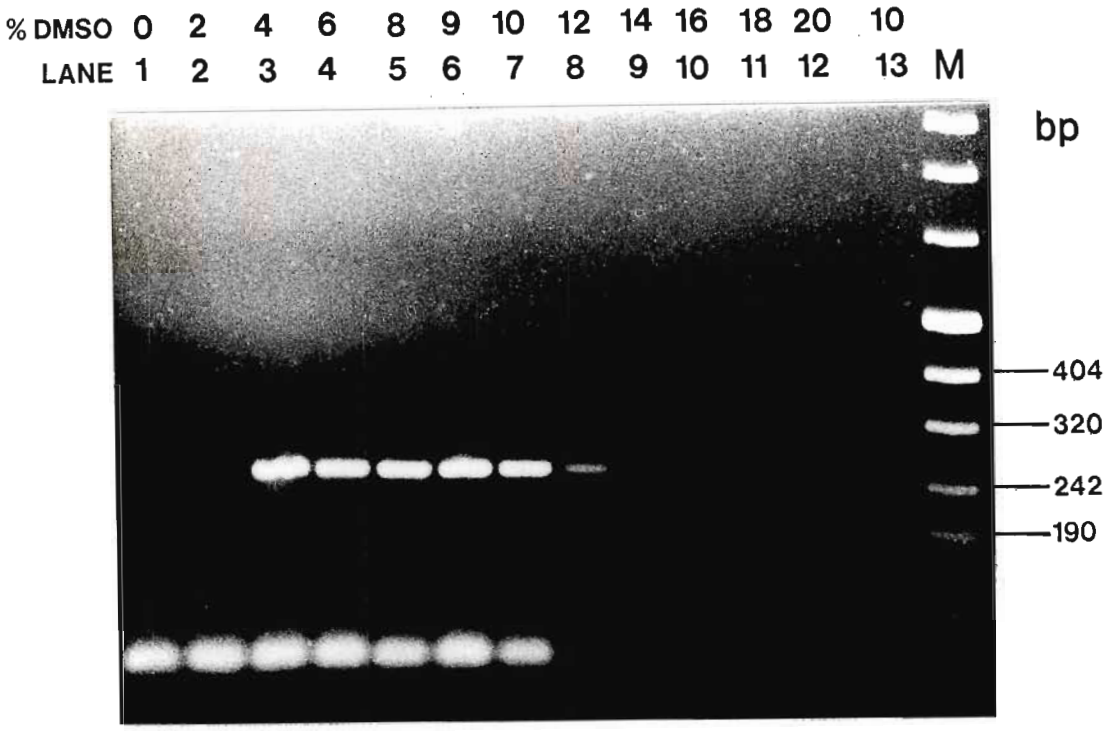


Plate 1: Agarose gel (2%) of the 244 bp apo E DNA amplified fragments at different DMSO concentrations. The concentration of DMSO (%) is indicated above the lanes. M = MW marker VIII. Lane 13 = no DNA. The 244 bp amplified product was detected in lanes 3-9 (DMSO concentration 4 to 14%). DMSO concentration less than 4% was insufficient for amplification. Amplification did not occur at a DMSO concentration higher than 14%.

3.4.2.2.2 PCR procedure

Each amplification reaction contained 8 µl DNA, *Taq* polymerase buffer [Tris HCl, 100 mmol/l; $MgCl_2$, 1.5 mmol/l; KCl, 500 mmol/l; gelatine, 1 mg/ml; pH 8.3], 1 µmol/l of each primer F4 and F6, 200 µmol/l of each dNTP; 10% (v/v) DMSO and 0.25 µl (1.25 units) of *Taq* DNA polymerase in a final volume of 50 µl. A "master mix" of the common reagents used in the PCR reaction was made up before each run. The pooled master mix (excluding the DNA) was aliquoted to reaction vials using a positive displacement pipette. Pooling the reagents in a master mix reduces tube to tube variability. Each run included a tube without added DNA to exclude contamination of reagents.

The sample was overlaid with 40µl of light mineral oil to prevent evaporation during thermal cycling. Silica gel was used in the heating block holes to ensure good thermal contact. Each reaction mix was subjected to the following temperature cycles in a Hybaid® DNA Thermal Cycler (Hybaid Ltd. UK). An initial denaturation step at 95°C for 5 min was followed by 30 cycles of
94°C for 1 min (denaturation)
63°C for 1 min (primer annealing) and
72°C for 2 min (extension).
A final 10 min extension step at 72°C completed the reaction.

3.4.2.2.3 Precautions taken during PCR procedure

The PCR is both very specific and sensitive and can detect DNA concentrations in the femtomole range or lower. This extreme sensitivity can lead to problems since contamination may result in false positives. Most occur as a result of "carry over" from a previous PCR run or contamination of reagents with DNA.

In order to avoid contamination, the following precautions were taken:

1) All stock reagents used were stored in aliquots.

2) Positive displacement pipettes were used.
3) PCR was performed at a site distant from areas in which DNA was extracted.
5) A separate pipette was used for dispensing DNA.
6) The reagent tubes were centrifuged before opening to minimise cross contamination of reagents.
7) Gloves were changed frequently, especially after handling amplified product.
8) Tubes were uncapped and closed carefully to avoid spillage and contamination of laboratory equipment and surfaces.
9) The DNA was always added last to the reaction tube.
10) Control tubes were included in each run in which the template (DNA) was omitted, to permit easy detection of contamination.

3.4.2.3 Detection of the PCR amplified products by agarose gel electrophoresis

The amplified DNA fragment can be visualised by a number of techniques (Chapter 2). Most reactions will result in sufficient DNA for visualisation by simple gel electrophoresis and staining with ethidium bromide. The method used in this study for the detection of the PCR product was electrophoresis through an agarose gel. At neutral pH, DNA is negatively charged and in the presence of an electrical current DNA loaded into a sample well at the cathode end of the gel moves through the gel towards the anode. The electrophoretic mobility of the DNA fragments in agarose is dependent on fragment size and fairly independent of base composition or sequence.

3.4.2.3.1 Procedure

Electrophoresis was carried out using the Hoefer Mighty Small® (Hoefer Scientific Instruments, San Francisco, California) horizontal apparatus. The running buffer was a Tris-borate (TBE) buffer made up as a 5X concentrated stock solution and diluted to a 0.5X working solution. The 5X TBE buffer was made up as follows: 0.4 mol/l Tris base; 0.4 mol/l boric acid; 1 mmol/l EDTA (pH 8.0). The loading buffer was 0.25% bromophenol blue in a 40% sucrose solution.

A 2% gel was cast by melting the agarose in TBE buffer (in a microwave oven) until a clear transparent solution was obtained. The solution was cooled quickly to approximately 70°C and ethidium bromide was added to a final concentration of 0.5 μg/ml. The melted solution was then poured into a mould and allowed to set. Each minigel was used to analyse seven samples plus a MW marker VIII (Boehringer Mannheim GmbH) which was run along-side the samples giving a ladder of fragments ranging from 1114 to 19 base pairs.

The gels submerged in running buffer were loaded with 7 μl of the PCR product mixed with 1.4 μl loading buffer and run at a constant voltage of 8 v/cm at room temperature until the dye front reached the bottom of the gel. After electrophoresis the gel was viewed and photographed as described below.

3.4.2.3.2. Staining and photography of gel

DNA itself is not fluorescent. When ethidium bromide binds to DNA in solution, the dye intercalates between the stacked bases of DNA and upon excitation under ultraviolet light, fluoresces pink/orange in the visible range. Ethidium bromide is best run in the gel as post-electrophoresis staining can result in blurring of DNA bands. Although the electrophoretic mobility of linear double stranded DNA is reduced by approximately 15% in the presence of the dye, the ability to examine the gel directly under ultraviolet light during or at the end of the run is often a great advantage. Ethidium bromide is a powerful mutagen and is moderately toxic and was handled with special care.

DNA fragments were visualised under UV illumination using a Camay transilluminator conversion kit at 366 nm. A polaroid CU-5 Land Camera and polaroid film type 667 was used (camera settings: speed t = ⅛; aperture f:11) to photograph the gels positioned on the UV transilluminator. The appropriate attachment lens was clamped in front of the camera objective and the height of the camera adjusted correspondingly to the

attached lens. The gels were photographed after a 5 second exposure time. A photograph of the 244 bp PCR amplified apo E gene product after agarose gel electrophoresis is shown in Plate 2 (Pg 54).

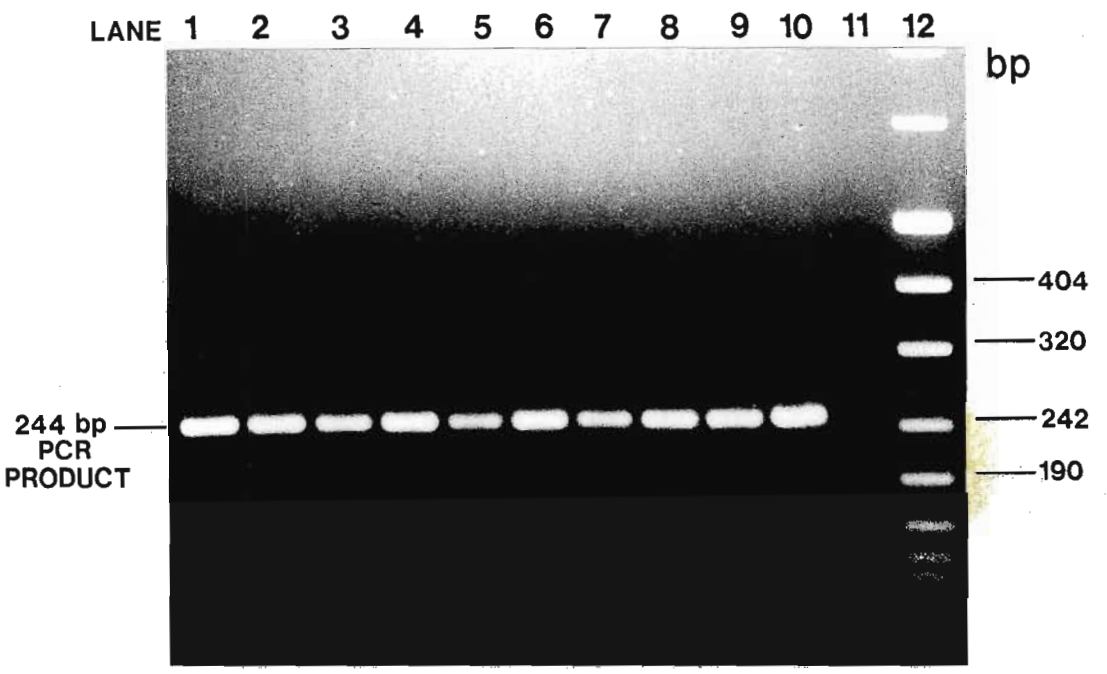


Plate 2: Agarose gel showing the amplified apo E PCR product using primers F4 and F6. Lane 11 = no DNA and lane 12 = MW marker VIII.

3.4.2.4 Restriction enzyme digestion

The principle of restriction enzyme digestion and application to apo E genotyping is outlined in Chapter 2. The restriction enzyme used in this protocol was *Cfo*I. This enzyme isolated from the bacterium *Clostridium formicoaceticum* is an isoschizomer of *Hha*I, that is, both enzymes recognise the same restriction site.

3.4.2.4.1 Procedure

After PCR amplification, 35 μl of each reaction mixture was incubated with 10 units of *Cfo*I (Boehringer Mannheim GmbH) for 3 h at 37°C. The enzyme was stored in a Tris-HCL buffer pH 7.5 containing NaCl 100 mmol/l; EDTA 1 mmol/l; 2-mercaptoethanol 10 mmol/l; glycerol 50% (v/v). It has optimal activity at 37°C in a Tris-HCl buffer at pH 7.5 containing Tris-HCl 10 mol/l; $MgCl_2$ 10 mmol/l; dithiothreitol 1 mmol/l. Both the storage and incubation buffers were supplied with the enzyme. The reaction was terminated by the addition of EDTA (0.2 mmol/l). The digestion process did not require purification of PCR products. Incomplete digestion was not encountered.

Final reaction mixture:

PCR product	35 μl
Buffer	5 μl
*Cfo*I enzyme	1 μl
Water	9 μl
Total volume	50 μl

Following restriction enzyme treatment the DNA fragments were separated by non-denaturing PAGE as described below. PAGE is preferred over agarose gel electrophoresis since the gels are much thinner and produce sharper and more discrete bands of DNA (Sambrook *et al.*, 1989). More importantly PAGE has much higher resolving power compared to agarose gels and enables clear distinction of even the

similarly sized fragments of 91, 81 and 72 base pairs, allowing unambiguous identification of the common apo E genotypes.

3.4.2.5 Polyacrylamide gel electrophoresis (PAGE) of restriction enzyme digests

Polyacrylamide gels are formed by the polymerization of acrylamide monomers into linear chains and the linking of these chains with *N,N'*-methylenebisacrylamide (bis). The polymerization is catalysed by ammonium persulfate and stabilised by TEMED (*N,N,N',N'*-tetramethylenediamine) (Rickwood and Hames 1985). When the bifunctional agent bis is included in the polymerisation reaction, the linear chains are cross-linked to form a gel, the porosity of which is determined by the length of the chains and the degree of cross-linking. The concentration of acrylamide and the ratio of acrylamide to bis determine the pore size of the resultant three dimensional network and hence its sieving effect on nucleic acids of different sizes.

3.4.2.5.1 Procedure

PAGE was carried out using the Hoefer Mighty Small® (Hoefer Scientific Instruments, San Francisco, California) vertical electrophoresis apparatus. A spacer thickness of 0.75 mm was found to be most suitable in terms of resolution and ease of handling. Gel-bond was not used as a backing since it was found to absorb UV light. One surface of each glass plate was treated with a silicone solution to prevent the gel from sticking to the glass. The bottom of the assembled gel moulds were sealed with agarose to prevent leakage of the acrylamide solution.

A 12% gel was found to yield optimal resolution in a series of experiments using polyacrylamide concentrations ranging from 8-20%. The amount of acrylamide used to make up one gel is given in Table III (p57). The 30% polyacrylamide stock solution was made up with 29 g of acrylamide and 1 g of N'N'methylenebisacrylamide dissolved in 100 ml water and the solution was filtered and stored at 4°C in a dark bottle. The TBE buffer was made up as described previously in section 3.4.2.3.2. The

10% ammonium persulphate was prepared fresh just before use. The 5.5 μl TEMED solution was added to the mixture prior to pouring the gel and mixed by swirling. Using a pasteur pipette the solution was transferred into the gel mould.

Table III: Volume of reagents used to cast a 12% polyacrylamide gel

REAGENTS	VOLUME
30% polyacrylamide (stock)	2 ml
TBE 5X	1 ml
distilled water	1.96 ml
10% ammonium persulfate	35 μl
TEMED	5.5 μl

The gel polymerised in approximately 30 min at room temperature. The wells were washed out thoroughly as soon as the comb was removed because small amounts of acrylamide solution trapped in the well polymerise producing irregularly shaped surfaces that give rise to distorted DNA bands. The gel was attached to the electrophoresis tank and the reservoirs of the tank were filled with 1X TBE buffer.

The digested reaction mixture (10 μl) containing loading buffer [Ficoll 30% and bromophenol blue 0.25% in 5X TBE buffer) was loaded into the wells using a Hamilton syringe. The DNA was electrophoresed for 3 h at 80 v at 4°C. Electrophoresis at a higher voltage caused "bowing" of the DNA bands.

3.4.2.5.2 Staining and photography of the gel

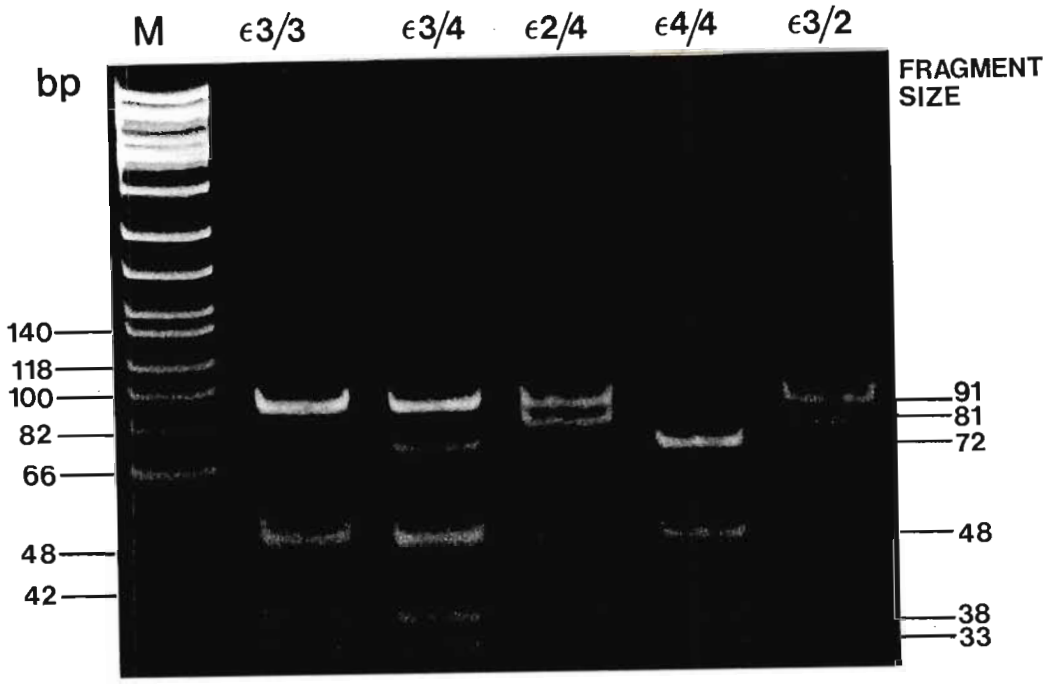
The gel was stained in a solution containing 0.5 μg/ml ethidium bromide in 1X TBE buffer for 30 min. The gel was viewed and photographed as described in section 3.4.2.3.2. above.

3.4.2.6. Analysis of *Cfo*I restricted fragments on photographed gel

Apo E restriction isotyping relies on the generation of allele specific fragments by cleavage at polymorphic *Cfo*I sites to distinguish $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$ sequences. The banding pattern after restriction enzyme digestion and separation of the common alleles on a polyacrylamide gel is illustrated in Fig. 6 (Pg 59), together with the restriction map for the three principal alleles $\epsilon 2$, $\epsilon 3$ and $\epsilon 4$ of the apo E gene and the fragment sizes from the polymorphic *Cfo*I sites. The DNA bands were estimated by comparison to molecular weight marker ϕx174 *Hinf*I (Promega) which ranged in length from 726 bp to 24 bp. The number and size of fragments obtained were used to characterise the apo E genotypes.

The gels were interpreted as shown in Table IV (p60). With the exception of a shared 38 bp fragment, each genotype possesses a unique combination of *Cfo*I fragment sizes. Samples from heterozygotic combinations contained both sets of fragments from each apo E allele. The $\epsilon 2/2$ genotype which was not present in this cohort contains 91 and 81 bp *Cfo*I fragments reflecting the absence of restriction sites at Cys_{112} and Cys_{158}.

(a)



(b)

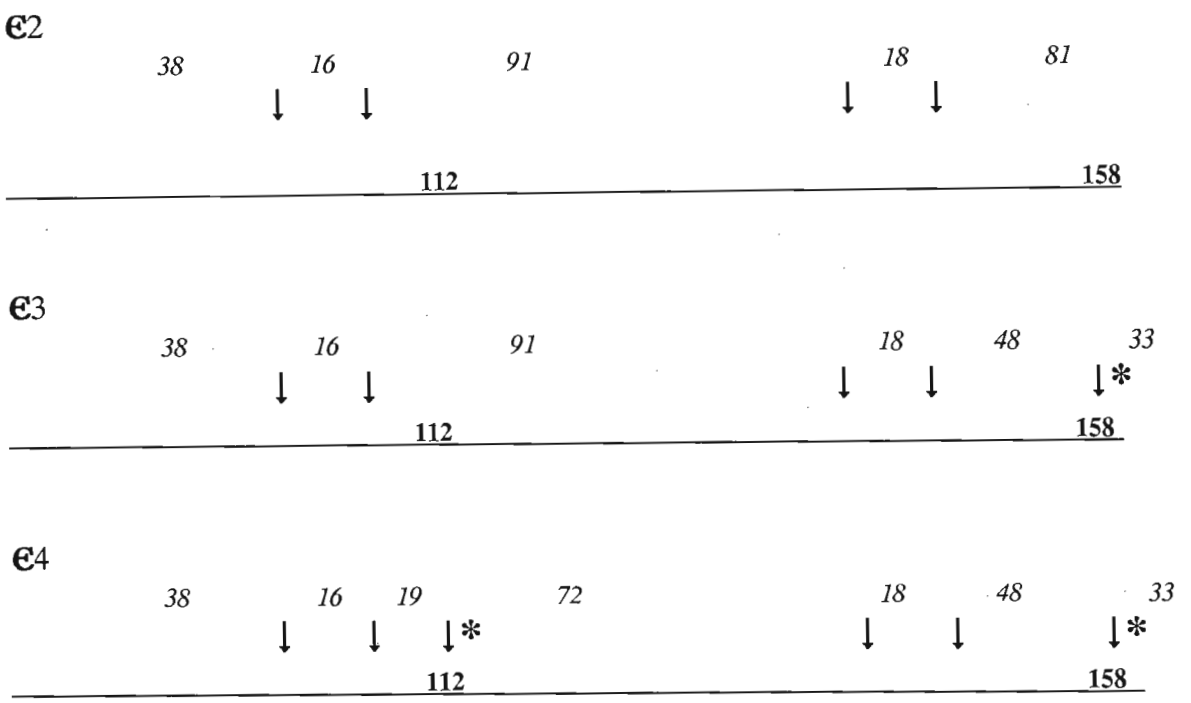


Fig. 6: **(a)** Polyacrylamide gel (12%) showing the of *Cfo*I restriction fragments from the five apo E genotypes encounted in the study sample. M = MW marker ϕx174 *Hinf*I. Note that the 48 bp band is more intense in the ϵ3/3, ϵ3/4 and ϵ4/4 samples compared to the ϵ4/2 and ϵ3/2 due the absence of fragment 48 bp in the ϵ2 allele.

(b) Restriction map of the common alleles ϵ2, ϵ3 and ϵ4 of the amplified 244 bp sequence from the apo E gene using primers F4 and F6 and restricted with restriction enzyme *Cfo*I. Codons 112 and 158 contain polymorphic nucleotides 3745 and 3883 respectively. Cutting sites for *Cfo*I are indicated by arrows, and variant sites with an asterisk. Numbers in italics indicate fragment sizes in base pairs. Fragment sizes below 33 bp are not visible on the gel.

Table IV: Interpretation of *Cfo*I digested fragments separated on polyacrylamide gels

Genotype	Amino acid		Polymorphic Nucleotide sequence		Fragments obtained after restriction enzyme digestion
	112	158	112	158	
ϵ3/3	Cys Cys	Arg Arg	GTGC GTGC	GCGC* GCGC*	91, 48, 38, 33
ϵ2/2#	Cys Cys	Cys Cys	GTGC GTGC	GTGC GTGC	91, 81, 38
ϵ4/4	Arg Arg	Arg Arg	GCGC* GCGC*	GCGC* GCGC*	72, 48, 38, 33, (19)
ϵ3/4	Cys Arg	Arg Arg	GTGC GCGC*	GCGC* GCGC*	91, 72, 48, 38, 33
ϵ3/2	Cys Cys	Arg Cys	GTGC GTGC	GCGC* GTGC	91, 81, 48, 38, 33
ϵ2/4	Cys Arg	Cys Arg	GTGC GCGC*	GTGC GCGC*	91, 81, 72, 48, 38, 33, (19)

* *Cfo*I recognises this sequences and cleaves it
\# Not shown on gel
Fragment size 19 indicated in parenthesis in not visible on the gel.

3.4.2.6.1 Examples of typical banding patterns encountered

Following optimisation of the method, the study group was genotyped as described above. The few ambiguous results were repeated. Samples of known apo E genotypes (received from M. Kotze, Department of Human Genetics, Tygerburg Hospital) were used as controls to confirm reliability of the method. Occasionally genotyping was carried out "blind" to verify credibility. Plates 3 (Pg 62), 4 (Pg 63), 5 (Pg 64) and 6 (Pg 65) are examples of polyacrylamide gel-separated products of apo E amplification and CfoI digestion of DNA representing the homozygotic and heterozygotic combination of the common apoE alleles in the study sample. Apo E genotypes ϵ3/3, ϵ2/4 and ϵ3/4 are shown in Plate 3 (Pg 62). The most common genotypes, ϵ3/3 and ϵ3/4 are readily distinguishable by the presence of a 72 bp fragment in the ϵ3/4 genotype. Sample labelled 'X' in lane 3 was analysed to be a ϵ2/4 genotype (by the presence of bands 91, 81, 72, 48, 38 and 33) and this was confirmed by a rerun in

Plates 4, lane 5. Referring to Plate 4 (Pg 63), genotypes $\epsilon 3/3$, $\epsilon 3/4$ $\epsilon 2/3$, $\epsilon 2/4$ and $\epsilon 4/4$ are shown. Lane 7 is the undigested 244 bp apo E PCR amplified product. The absence of the 91 bp fragment in lane 6 clearly distinguished the $\epsilon 4/4$ from the $\epsilon 3/3$ genotype and indeed all other apo E genotypes. Sample labelled 'Y' in Plate 4 was re-run on Plate 6 and confirmed to be a $\epsilon 3/2$ genotype.

Referring to Plate 5 (Pg 64), DNA samples labelled 'A', 'B' and 'C' with genotypes $\epsilon 3/3$ $\epsilon 4/4$ and $\epsilon 3/4$ were verified by the re-run shown in Plate 6. Plate 6 shows all five of the apo E genotypes encountered in the study group. Once again one could easily discriminate between the $\epsilon 3/3$ and $\epsilon 3/4$ genotypes made by the presence of a 72 bp fragment in the $\epsilon 3/4$ genotype can be seen in Plate 6 (Pg 65). The clear differentiation between the $\epsilon 3/4$ and the $\epsilon 3/2$ genotype (Plate 6, lane 3 and 6 respectively), is made possible by the MW marker fragment size 82 bp. The 81 bp fragment of the $\epsilon 3/2$ genotype co-migrates with the MW marker, fragment 82 bp, and the 72 bp fragment of the $\epsilon 3/4$ runs further away from the 82 bp fragment of the MW marker.

Five out of the six common genotypes were observed in the study group. The genotypes represent 2 homozygotic and three heterozygotic combinations of apo E alleles in the Indian population. The genotypes together with the lipid and lipoprotein profiles and anthropometric indices were recorded on a computer spreadsheet (Quattro Pro) for further analysis.

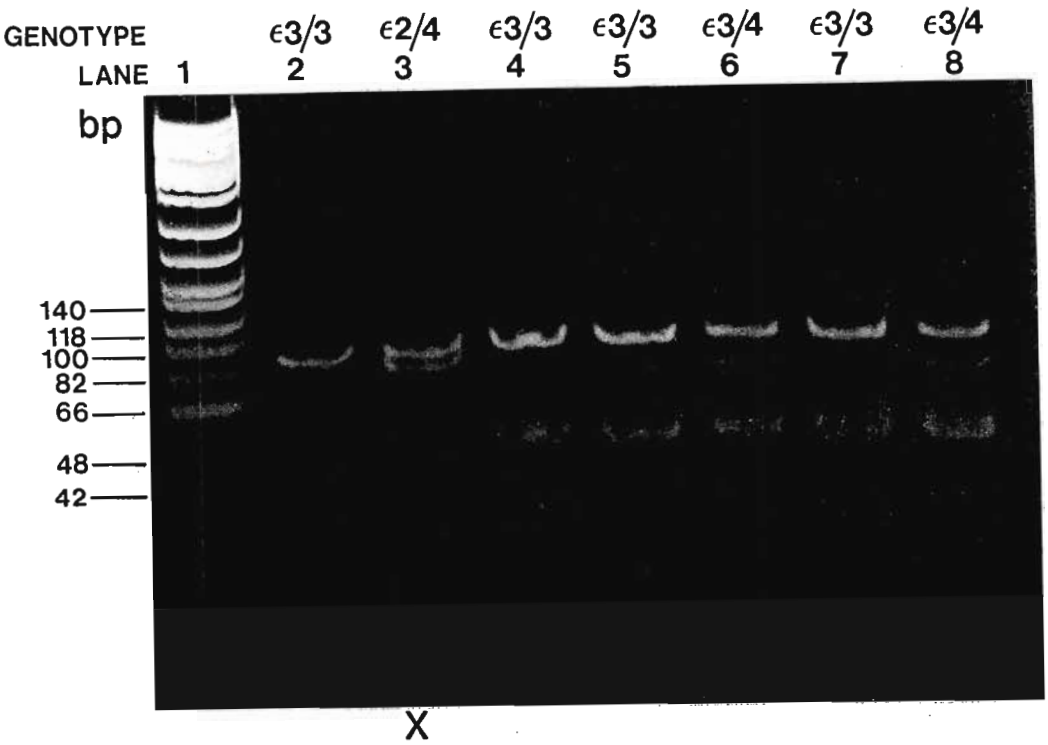


Plate 3: Polyacrylamide gel, of the *Cfo*I restricted fragments, stained with ethidium bromide and photographed under UV light. Lane 1 = MW marker ϕx174 *Hinf*I. Lanes 2-8 are different DNA samples whose apo E genotype is indicated above the gel. Sample labelled 'X' in lane 3 was genotyped as ϵ2/4 and confirmed in a rerun as shown in Plate 4.

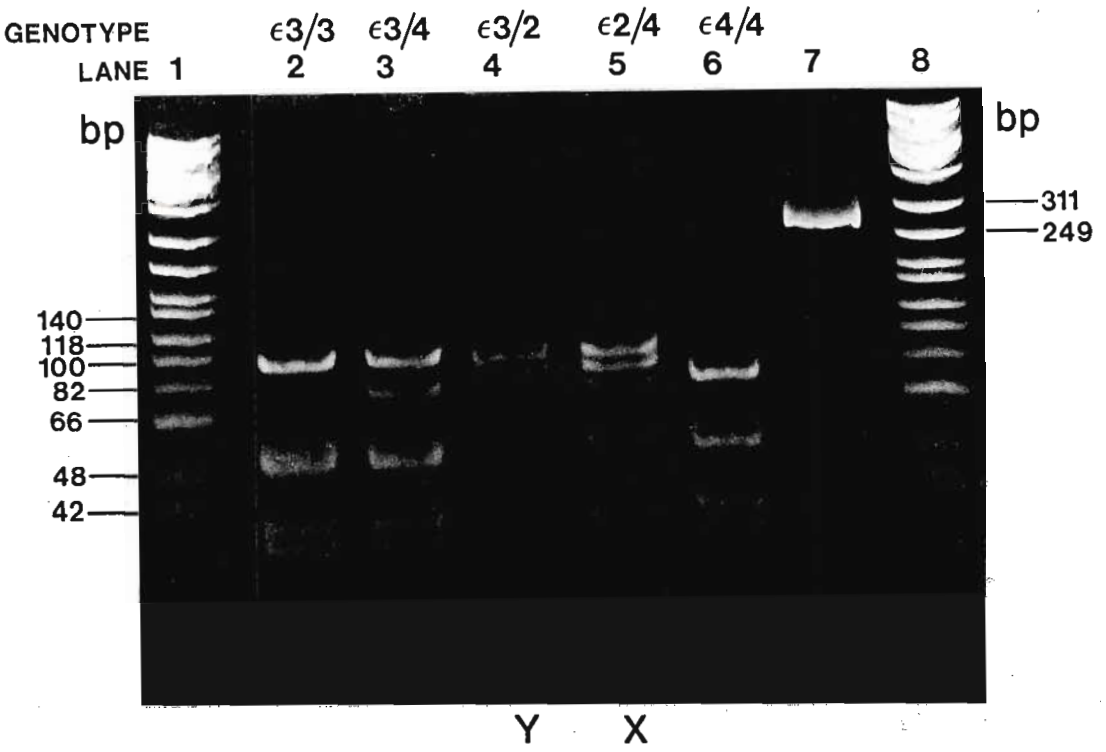


Plate 4: Polyacrylamide gel showing the *Cfo*I restricted fragments. Lane 1 and 8 = MW marker ϕx174 *Hinf*I. Lane 2-6 show restricted DNA samples whose apo E genotype is indicated above the gel. Sample labelled 'X' in the gel shown previously in Plate 3 was run in lane 5 and confirmed to be a ϵ2/4 genotype. Lane 7 is the apo E PCR product (244 bp) before digestion with *Cfo*I. Sample labelled 'Y' was rerun as shown in Plate 6.

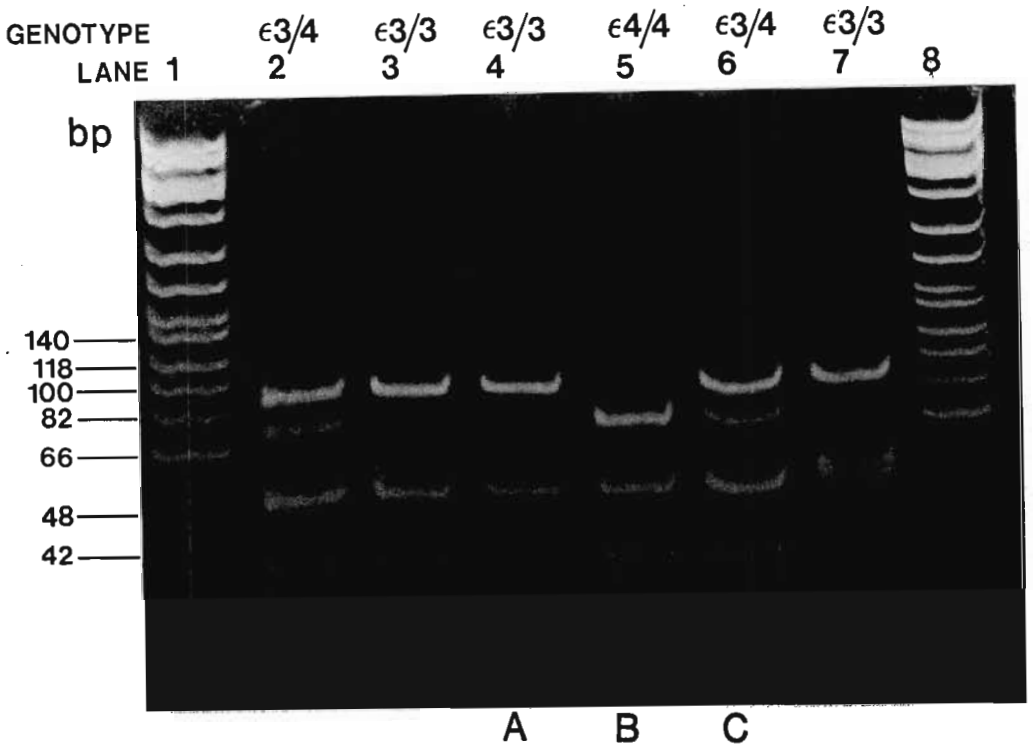


Plate 5: Polyacrylamide gel of *Cfo*I restricted fragments. Lane 1 and 8 = MW marker ϕx174 *Hinf*I. Lanes 2-7 are different DNA samples whose apo E genotype is indicated above the gel. Samples marked 'A', 'B' and 'C' were rerun in experiment shown in Plate 6 and confirmed to be apo E genotype ϵ3/3, ϵ4/4 and ϵ3/4, respectively.

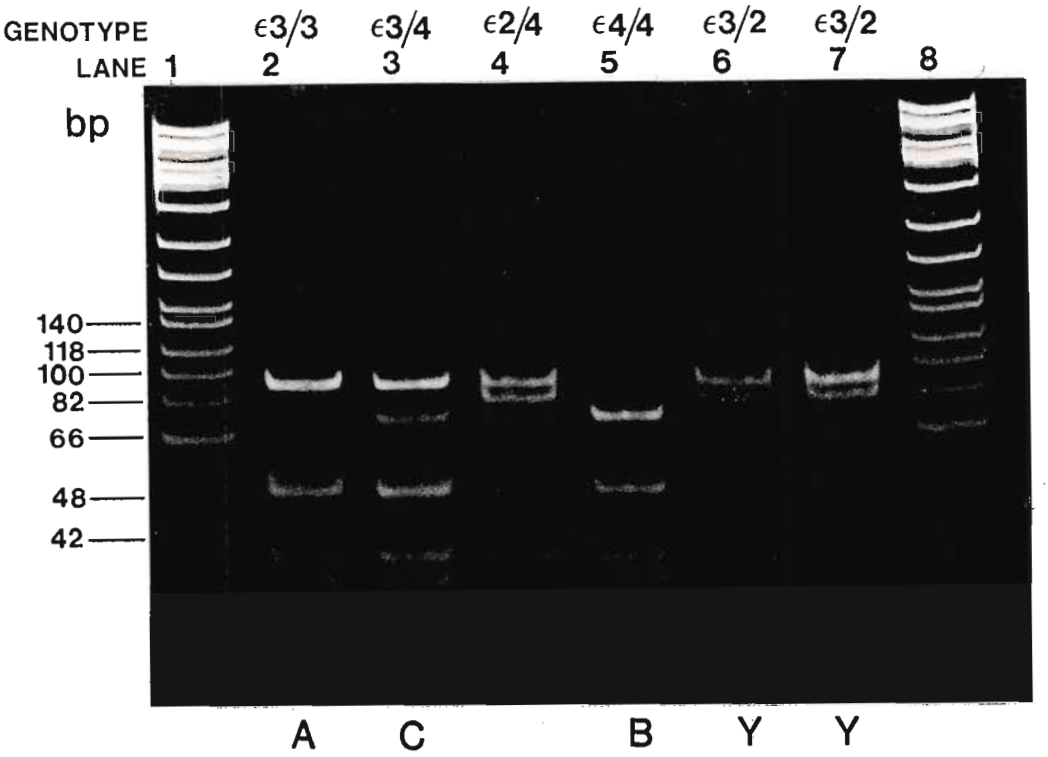


Plate 6: Polyacrylamide gel of *Cfo*I restricted fragments. Lane 1 and 8 = MW marker ϕx174 *Hinf*I. Lane 2-7 are different DNA samples whose apo E genotype is indicated above the gel. Samples marked 'A', 'B' and 'C' from Plate 5 were run in lanes 2, 5 and 3 respectively, and 'Y' from plate 4 was run in lanes 6 and 7. The genotypes of all these samples were confirmed.

3.4.2.7 Reagents

All reagents used for apo E genotyping were molecular biology grade and were purchased from the following companies: Sigma Chemical Co., BDH chemicals, Merck, Analar, Promega and Boehringer Mannheim. Where necessary the reagents were aliquoted on receipt and stored at -20°C.

3.5 STATISTICAL ANALYSES

Statistical analysis were performed using the computer package SAS (Statistical analysis system) version 6 (1990). The allele frequencies (denoted $f_{\epsilon 2}$, $f_{\epsilon 3}$, $f_{\epsilon 4}$) were estimated by the gene counting method. For example the frequency of $\epsilon 2$ was calculated as follows:

$$f_{\epsilon 2} = \frac{2(\epsilon 2/2) + \epsilon 3/2 + \epsilon 4/2}{\text{total number of alleles}}$$

Chi-square statistics was used to test for goodness of fit to the Hardy-Weinberg equilibrium. The variance in the frequency was given by the 95% confidence interval (CI).

Statistical analysis was also carried out to determine the influence of apo E genotype on plasma lipid levels. The basic technique was the use of analysis of variance (ANOVA) in which the lipid values were dependent variables and apo E genotype was the independent variable. The Student's t-test was used to determine significance and differences in lipid levels between the genotype classes. Since a number of other biological characters also influence lipid levels, covariate analysis was carried out to determine the influence of age, BMI, WHR and menopause. The analysis was done in a stepwise fashion to determine the contribution of each covariate as it entered the model. Least-square means were calculated to adjust for the effect of covariates. Details of the application of these statistical procedures will be given in the result section where relevant.

CHAPTER 4

RESULTS

LIST OF CONTENTS **PAGE**

In this chapter the frequency distribution of the common apo E genotypes and alleles in the study population is presented. A second series of analyses examined the influence of apo E polymorphism on various plasma and apolipoproteins (total cholesterol, HDL6-C, LDL-C, triglyceride, apo A-I, apo B) referred to as the "lipid profile".

4.1 APO E GENOTYPE DISTRIBUTION AND ALLELE FREQUENCIES IN THE STUDY GROUP

The apo E genotypes of 173 Indian nurses (clinical and other details in Chapter 3) were determined and included in the statistical analysis. The frequency of the different genotypes is presented in Table V (Pg 68). The $\epsilon 3$ is by far the most common allele and the $\epsilon 3/3$ is the most frequent genotype in this sample. The $\epsilon 2/2$ homozygote was not encountered in the study sample. The apo E allele frequencies in this sample were: apo $\epsilon 2$ = 0.012, apo $\epsilon 3$ = 0.876 and apo $\epsilon 4$ = 0.113 (Table VI, Pg 69).

TABLE V: Genotype counts and relative frequencies of the common apo E variants in the study cohort.

GENOTYPES	OBSERVED		95% CI
	Count	Frequency (%)	
$\epsilon 2/2$	0	0	0 - 1.1
$\epsilon 2/3$	3	1.7	0 - 3.6
$\epsilon 2/4$	1	0.6	0 - 1.8
$\epsilon 3/3$	135	78.0	71.8 - 84.2
$\epsilon 3/4$	30	17.3	11.6 - 23.0
$\epsilon 4/4$	4	2.3	0.2 - 4.5

TABLE VI: Apo E allele frequencies and 95% CI of the study population

Allele	Proportion	Frequency (%)	95% CI
ε2	0.0116	1.2	0.06 - 2.26
ε3	0.8757	87.6	84.08 - 91.06
ε4	0.1127	11.3	7.94 - 14.60

Table VII (Pg 69) compares the observed distribution of apo E genotypes in the population sample and the expected distribution assuming a Hardy-Weinberg equilibrium. The Hardy-Weinberg expectation of the phenotype distribution was calculated from the calculated allele frequencies (Pg 68). Alleles at the apo E locus are in equilibrium according to the Hardy-Weinberg law which states that genotypes are distributed in proportion to the frequencies of individual alleles in a population and remain constant from generation to generation (Thompson *et al.*, 1991), which was tested using a chi-square test.

Table VII: Apo E genotypes and gene frequencies in a sample of 173 South African Indian Females*

Genotype	Observed	Hardy-Weinberg expected
	%	%
ε2/2	0	0.01
ε2/3	1.73	2.03
ε2/4	0.58	0.26
ε3/3	78.03	76.69
ε3/4	17.34	19.74
ε4/4	2.31	1.27
Total	99.99	100

* The chi-square value to test the likihood that the observed apo E genotype distribution differs from the Hardy-Weinberg expected distribution by chance = 2.79; p (df=5) > 0.05

There was only one nurse with the apo E genotype $\epsilon 2/4$, three with $\epsilon 2/3$, and four with $\epsilon 4/4$ and no nurses with the $\epsilon 2/2$ genotype (Table V, Pg 68). Thus, the $\epsilon 3/3$ and $\epsilon 3/4$ genotypes together represent 96% of the study population. Of the 135 nurses with the $\epsilon 3/3$ genotype, 50 were postmenopausal and 85 were premenopausal. With respect to the $\epsilon 3/4$ genotype, 19 were postmenopausal while 11 were premenopausal.

4.2 IMPACT OF APO E POLYMORPHISM ON THE LIPID PROFILE OF THE STUDY POPULATION

One of the problems encountered when considering the influence of apo E genotype was the small number of observations in the $\epsilon 2/2$, $\epsilon 2/3$, $\epsilon 2/4$ and $\epsilon 4/4$ genotype classes. Since the sample with the $\epsilon 3/4$ and $\epsilon 3/3$ genotypes represent 96% of the total apo E frequency in this study group these genotypes were the only ones considered in estimating the quantitative impact of apo E polymorphism on the lipid profile. Descriptive statistics was carried out on the lipid profile in these two genotype classes to give an indication of the mean and the dispersion from the mean of the sample. The means and standard deviations of the lipids are shown in Table VIII (Pg 71). All the lipid variables showed a Gaussian distribution except triglycerides which were positively skewed. Analysis of triglyceride measurements were performed after logarithmic transformation of the values. The lipid profile showed a significantly higher ($p < 0.05$) total cholesterol, LDL-C and triglyceride levels, in nurses with the $\epsilon 3/4$ genotype compared to the $\epsilon 3/3$ genotype.

Table VIII: The serum lipid profile as a function of apo E genotype ε3/3 and ε3/4.

Dependent variable	Genotype		p*
	ε3/3 (n=135)	ε3/4 (n=30)	
Total Cholesterol	5.18 (1.05)	5.87 (1.20)	0.0019
HDL-C	1.28 (0.39)	1.14 (0.04)	0.091
LDL-C	3.24 (0.97)	3.90 (1.0)	0.0013
Triglyceride	1.08 (3.66)	2.99 (4.01)	0.0002
Apo A_1	136.27 (37.33)	130.70 (32.67)	0.45
Apo B	83.99 (23.44)	90.67 (21.47)	0.15

* Probability of obtaining these differences among genotypes when the hypothesis of equal means is true calculated using the unpaired t-test.
Figures in parentheses are the standard deviation of the individual values.

A number of other factors are known to influence lipid levels such as age, menopause and indices of adiposity; hence, it was considered necessary to look at the effects of these risk factors in the study group in evaluating the effect of the apo E locus on the lipid profile. This was done using analysis of covariance to determine the effect of menopausal status, age, waist:hip ratio (WHR) and body mass index (BMI) on the distribution of each lipid measure. The two genotype classes did not differ significantly with regard to BMI and WHR (Table IX, Pg 72). However, with regard to age, the t-test showed the distribution of two genotype class to be significantly different ($p < 0.05$) in this study group, the postmenopausal group being relatively enriched in the ε3/4 genotype. While some variation of apo E genotype with age has been noted in previous studies (Eggertsen *et al.*, 1993), further work would be required to confirm this unexpected correlation of apo E allele frequency with age and its biological meaning.

Table IX: Biological characteristics according to apo E genotypes ε3/3 and ε3/4.

Variable	Genotype		p*
	ε3/3 (n=135)	ε3/4 (n=30)	
Age (years)	41.01 (6.72)	44.43 (5.86)	0.011
BMI (kg/m^2)	24.06 (4.27)	23.31 (3.06)	0.37
WHR (cm)	0.81 (0.08)	0.82 (0.07)	0.51

* Probability of obtaining these differences among genotypes when the hypothesis of equal means is true.
Figures in parentheses are the standard deviation of the individual values.

Analysis of covariance was carried out in a stepwise fashion to identify the significant covariates that affect the lipid profile (Table X, Pg 74) at the 5% level. In this analysis the coefficient of multiple determination (R^2) was used to calculate at each step, the proportion of the total variation in the dependent lipid variable explained by variation in all the independent variables or covariates. The level of significance for each covariate is indicated in parenthesis in Table X (Pg 74) after taking into account the effects of all the other covariates on the lipid profile. The analysis of covariance uses a generalised linear model procedure which also takes into account the unbalanced number of observations in the genotype classes. From Table X (Pg 74) the following is evident:

a) Apo E polymorphism had a statistically significant ($p < 0.05$) influence on total cholesterol, LDL-C and triglyceride levels.

b) The significant effect of menopausal status on the lipid profile was lost when age entered the analysis model. For example, the R^2 value for total cholesterol increased from 5.75 with the apo E genotype alone to 8.66 when menopause was entered. At this stage of the analysis the menopause was significant ($p = 0.024$), but when age was added to the model the significant level of menopause was $p = 0.34$ (not shown in

Table X, Pg 74). This implies that the effect of menopausal status is not statistically significant in this study group.

c) BMI made no significant contribution to the lipid profile.

d) When WHR entered the model the R^2 increased from 3.36 to 9.82 for HDL-C and 1.42 to 4.59 for apo A_1. HDL-C and apo A_1 are related lipid variables and were the only lipids significantly influenced by WHR.

Table X: Analysis of Covariance in a stepwise fashion showing the percentage variance, R^2, of the dependent variables as each covariate enters the model.

Dependent variables	Covariates				
	Apo E Genotype	Menopausal status	Age	BMI	WHR
Cholesterol	5.75 (0.016)	8.66 (0.40)	17.22 (0.0001)	17.22 (0.99)	18.02 (0.22)
HDL-C	1.74 (0.084)	2.05 (0.55)	2.11 (0.33)	3.36 0.10)	9.82 (0.0009)
LDL-C	6.20 (0.011)	8.31 (0.32)	15.68 (0.0004)	15.69 (0.89)	15.75 (0.73)
Triglyceride	8.31 (0.0034)	16.77 (0.39)	21.92 (0.0044)	22.11 (0.49)	22.84 (0.22)
Apo A-I	0.35 (0.68)	1.03 (0.83)	1.17 (0.69)	1.42 (0.59)	4.59 (0.023)
Apo B	1.24 (0.15)	1.26 (0.51)	1.73 (0.59)	2.34 (0.30)	2.64 (0.49)

The level of significance of the variation contributed by each covariate is given in parenthesis, taking into account the influence of the other covariates.

Two-way ANOVA was used to test the possibility of an interaction between apo E genotypes and menopausal status on the lipid profile. This analysis showed no interaction between apo E genotype and menopausal status on the lipid profile (Table XI, Pg 75), indicating that the effect of the apo E genotype on the nurses' lipid profile do not differ between the premenopausal and postmenopausal women. The two-way analysis of variance was extended to investigate the possibility of an interaction between apo E genotype and the other covariates (age, BMI, WHR) on the lipid levels (Table XI, Pg 75). Apo E genotype interacts with WHR for both HDL-C and triglyceride levels. An interaction between BMI and apo E genotype is shown for LDL-C levels. There was no interaction between apo E genotypes and covariates for apo A-I and apo B, which were minimally affected by the biological variables considered in this study.

Table XI: Two-way ANOVA looking for an interaction between apo E genotype and covariates such as menopause, age, BMI and WHR on the lipid profile

Lipid Variable	Covariate	Significance level (p)*		
		Covariate §	Apo E genotype ‡	Interaction
Total Cholesterol	Menopause	0.02	0.009	No
	Age	0.0001	0.018	No
	BMI	0.31	0.002	No
	WHR	0.67	0.002	No
HDL-C	Menopause	0.47	0.13	No
	Age	0.78	0.09	No
	BMI	0.21	0.07	No
	WHR	0.0001	0.02	Yes (p=0.014)
LDL-C	Menopause	0.06	0.005	No
	Age	0.001	0.009	No
	BMI	0.03	0.12	Yes (p=0.044)
	WHR	0.74	0.001	No
Triglyceride	Menopause	0.0001	0.002	No
	Age	0.0001	0.003	No
	BMI	0.02	0.0001	No
	WHR	0.0006	0.014	Yes (p=0.005)
Apo A-I	Menopause	0.12	0.07	No
	Age	0.07	0.06	No
	BMI	0.75	0.13	No
	WHR	0.41	0.12	No
Apo B	Menopause	0.89	0.16	No
	Age	0.82	0.15	No
	BMI	0.09	0.12	No
	WHR	0.05	0.17	No

* significant at $p < 0.05$

§ Effect of the covariate on the dependent variable after accounting for the influence of apo E genotype.

‡ Effect of apo E genotype on the dependent variable after accounting for the influence of the particular covariate

From the above statistical analyses each lipid variable was then adjusted to remove the effects of their respective set of significant covariates. Total cholesterol, LDL-C and

triglyceride were adjusted for age only, HDL-C and apo A-I for the influence of WHR. The mean lipid levels in the two genotype classes and the significant difference between classes after adjusting for these covariates is presented in Table XII (Pg 76).

Comparing the mean serum lipid profiles before adjustment (Table VIII. Pg 71) and after adjustment (Table XII, Pg 76) show that differences between total cholesterol, LDL-C, triglyceride between apo E genotypes $\epsilon 3/3$ and $\epsilon 3/4$ was still significant ($p < 0.05$). The $\epsilon 3/4$ heterozygotes had higher average total cholesterol, LDL-C and triglyceride levels than the $\epsilon 3/3$ homozygotes. It therefore follows that the $\epsilon 4$ allele is associated with higher total cholesterol, LDL-C and triglyceride levels compared to the $\epsilon 3$ allele in the study population. As will be discussed further in Chapter 5, the association between $\epsilon 3/4$ genotype and higher total cholesterol and LDL-C has also been described in with other populations. The association between the apo $\epsilon 3/4$ genotype and higher triglyceride levels is not a significant and consistent finding in other published studies.

Table XII: Mean adjusted# lipid profile for the genotype classes $\epsilon 3/3$ and $\epsilon 3/4$

Dependent Variable	Apo E Genotype		p*
	$\epsilon 3/3$ (n=135)	$\epsilon 3/4$ (n=30)	
Total cholesterol	5.21	5.71	0.019
HDL-C	1.27	1.15	0.12
LDL-C	3.28	3.77	0.012
Triglyceride‡	1.24	1.76	0.0006
Apo A-I	136.10	131.47	0.53
Apo B	83.99	90.67	0.15

Explained in previous text
* Probability of obtaining these differences among genotypes when the hypothesis of equal means is true.
‡ Triglyceride values were subjected to logarithmic transformation before analysis.

In summary apo E genotype clearly influenced total cholesterol, LDL-C and triglyceride levels in this cross-sectional analysis. The apo $\epsilon 3/4$ was associated with higher levels of total cholesterol, LDL-C and triglyceride compared to the $\epsilon 3/3$ genotype. There was no interaction between apo E genotype and menopausal status (or age) on the lipid profile. The greatest amount of variation in the lipid profile was seen in triglyceride levels (R^2 = 22.8%) and may be attributed mainly to apo E genotype and age. Still, 74.7% of the variation in triglyceride levels was not associated with these known risk factors.

CHAPTER 5

DISCUSSION

LIST OF CONTENTS **PAGE**

5.1 INTRODUCTION

A considerable body of literature is devoted to the association of apo E alleles or genotypes with particular features of the lipid profile and the implications this may have for the risk of developing CHD. This system has also been particularly fruitful in examining gene:gene or gene:environment interactions as exemplified by the range of expression of the $\epsilon 2/2$ genotype (see Chapter 2). In an effort to understand the role of genetic variation in determining the risk of CHD, this study investigated apo E polymorphism and its relationship to the lipid profile in a group of healthy middle-aged Indian women. Results are discussed together with the choice of method for determining apo E polymorphism as well as comparison of the apo E allele frequencies in this group to other published studies.

5.2 RELIABILITY OF RESTRICTION ISOTYPING FOR APO E GENOTYPING

The advantages and limitations of the genotyping method, restriction isotyping, used in this study is discussed in Chapter 2. This method detects the allelic variation at nucleotides 3745 (amino acid 112) and 3883 (amino acid 158) of the apo E gene (Chapter 2). To do this it depends upon the creation or abolition of a *Cfo*I restriction site arising from substitutions at these polymorphic nucleotides. Rare apo E variants have been reported which do not affect the common polymorphic sites (amino acid 112 and 158). Most of the reported instances have been associated with type III HLP phenotype with only three exceptions (see Chapter 2, Table II, Pg 22). The electrophoretic isoform has ranged from an E7 to E1. Such variants would not be identified by the method used in this study. According to Wenham *et al.*(1991) these variants are so rare that such misclassification is insignificant in a study of this nature, especially in comparison to the misclassification encountered with commonly used protein electrophoretic analysis.

The restriction isotyping method has proved to be both convenient and reliable: it

obviates the need for radio-isotopically labelled probes; only small sample volumes are required; identification of genotypes is sensitive and specific (with the limitations already noted). Moreover, it can be performed on DNA isolated by the boiling method without the need for purification. The method is well established in practice and its reliability is verified in Chapter 3. No ambiguous results that could not be resolved on repetition were encountered in this investigation. These properties have made the restriction isotyping procedure suitable for large-scale epidemiologic and clinical studies.

5.3 APO E ALLELE FREQUENCY

5.3.1 Epidemiological considerations

The geographic distribution of the study population was Durban and surrounding areas. The majority of Indians living in South Africa are concentrated in these areas according to the Central States Statistics Department Pretoria (1991). The study population ancestry may be traced to both North and South of India. The present study is a reasonably sized population and is larger than 14 of the 50 other populations in which apo E frequencies have been reported (Hallman *et al.*, 1991; Gerdes *et al.*, 1992; Baileul *et al.*, 1993; Eggertsen *et al.*, 1993; Kamboh *et al.*, 1993; James *et al.*, 1993). Subjects of both Hindu and Muslim religious persuasions were included (ratio 85:15).

The study population was constrained by sex (female), age (35-55 years), profession (nursing) and health requirements. In these respects it does not meet the criteria for a random cross-section of the Indian community. It has been reported that the exclusion of hyperlipidemic individuals, as they were in this study, may result in skewed estimates of allele frequencies (Davignon *et al.*, 1988). The common effect of such criteria is a higher $\epsilon 3$ allele frequency in the study population than the background population (Gerdes *et al.*, 1992). Thus, the allele frequencies obtained in this study are not the optimum estimate of the overall distribution of the apo E allele frequencies of the South African Indian community as a whole. Bearing this limitation in mind, the

size and reasonably representative nature of the group studied allows for some cautious generalisations to be discussed below.

5.3.2 Interpopulation heterogeneity

At least 50 population groups have been studied which have shown a significant interpopulation variation in apo E allele frequency (Hallman *et al.*, 1991; Gerdes *et al.*, 1992; Baileul *et al.*, 1993; Eggertsen *et al.*, 1993; Kamboh *et al.*, 1993; James *et al.*, 1993). The $\epsilon 4$ allele occurs at a lower frequency in the Japanese (0.067) and Chinese (0.064) and a much higher frequency in the New Guineans (0.368), Blacks (0.297) and the Fins (0.227) (Gerdes *et al.*, 1992). The $\epsilon 2$ allele has similarly shown variation with a lower frequency in the Mayans (0.000), Blacks (0.027) and Japanese (0.023) populations compared to the New Guineans (0,146), Germans (0.141) and Caucasians in USA (0.130) (Gerdes *et al.,* 1992). Studies on the same populations have sometimes yielded differences in apo E allele frequencies, such as the Caucasians, and Chinese (Davignon *et al.*, 1988; Gerdes *et al.*, 1992) and the French (Bailleul *et al.*, 1993), which may reflect genetic microheterogeneity or design inadequacies. Taking all these factors into consideration a rough classification is given in Table XIII (Pg 82).

The apo E allele frequency profile of the study sample was compared to 46 other populations around the world that were grouped by cluster analysis according to the frequency of their mutant alleles (Gerdes *et al.*, 1992). The study population was found to cluster with populations from Italy (Padova), Spain (Barcelona), Japan (Asahikawa, Tokyo, Kuamoto, Hiroshima and Nagasaki), as well as, Mexicans (Texas, USA), Mayans (Yucatan, Mexico) and interestingly Indians from Singapore. This cluster is characterised by a relatively low frequency of $\epsilon 2$ allele and it is improbable that a more representative sample of the South African Indian population would materially alter its ranking amongst populations with this characteristic.

Table XIII: Heterogeneity amongst the population frequency of the common mutant apoE alleles (adapted from Gerdes *et al.*, 1992)

Mutant allele	Study Population	Frequency
High ϵ4 (apo E4)	New Guineans	0.368
	Nigeria (Benin)	0.297
	Sudan (Khartoum)	0.291
	Finland (Helsinki)	0.227
	Denmark (Aarhus)	0.174
Low ϵ4 (apo E4)	Chinese (Montreal)	0.064
	Japan (Kyushu and Fukuoka)	0.067
	China (Beijing)	0.070
	Japan (Hiroshima and Nagasaki)	0.086
	Italy (Padova)	0.098
High ϵ2 (apo E2)	New Guineans	0.146
	Germany (Marburg)	0.141
	USA (Massachusetts)	0.130
	Nancy (France)	0.130
	China (Beijing)	0.124
Low ϵ2 (apo E2)	Mayana (Yucatan, Mexico)	0.000
	(Indian - present study)	**0.012**
	Japan (Hiroshima and Nagasaki)	0.023
	USA Blacks (3 regions)	0.034
	Finland (Helsinki)	0.041
	USA (Pittsburgh)	0.059
	Germany (Marburg and Giessen)	0.077
	Paris (France)	0.079
	Chinese (Singapore)	0.097

Of the 50 published population studies on apo E frequency, around the world (Hallman *et al.*, 1991; Gerdes *et al.*, 1992; Baileul *et al.*, 1993; Eggertsen *et al.*, 1993; Kamboh *et al.*, 1993; James *et al.*, 1993), this study population has the third highest $\epsilon 3$ allele frequency (0.876) [highest being the Mayans (n=135) from Yucatan Mexico (0.911), followed by the Japanese (n=110) from Hiroshima and Nagasaki (0.891)]. With regard to the $\epsilon 4$ allele, there are 15 other population groups with a lower frequency of the $\epsilon 4$ allele compared to the study population. Apart from the Mayans from Yucatan and Mexico (Kamboh *et al.*, 1991) where the reported $\epsilon 2$ allele was zero (n=135) the present study population has the lowest $\epsilon 2$ allele frequency of the 49 other populations studied around the world. Hence, the apo E allele frequency distribution in the South African Indian population may be skewed towards a low $\epsilon 2$ frequency and a high frequency of $\epsilon 3$ allele. The study population clearly differs from the Caucasian populations in the northern regions of Europe, in Canada and in the USA in whom a much higher frequency of the $\epsilon 4$ allele is observed, and the Chinese, Finns, Blacks and New Guineans; where the Chinese are associated with a much higher $\epsilon 2$ allele and the Finns, Blacks and New Guineans are associated with a higher frequency of the $\epsilon 4$ allele.

The clinical significance of apo E allele frequencies is presented in Chapter 2. To summarise, most population epidemiological studies are compatible with a lower risk for CHD in individuals possessing the $\epsilon 2$ allele, mediated, at least in part, through low total cholesterol and LDL-C values. This may be counter-balanced to some extent, by the development of hyperlipidaemia in some individuals with the $\epsilon 2/2$ genotype. The low prevalence of the $\epsilon 2$ allele in the South African Indian population may contribute to the high incidence of CHD reported in this group.

5.4 APO E GENOTYPE AND OTHER DETERMINANTS OF THE LIPID PROFILE IN THE STUDY POPULATION

The chief features of this study were that the $\epsilon 3/4$ genotype was strongly associated with

an increase in triglyceride levels in this group with smaller effects on cholesterol levels. The other major determinants of the lipid profile were age, on total cholesterol, LDL-C and triglyceride levels, and WHR which was negatively correlated to HDL and apo A-I. Age showed the greatest influence on the lipid profile in this study population. It is now well accepted that aging is accompanied by a progressive increase of cholesterol, triglyceride and LDL-C, and a tendency for HDL-C to decrease (Moreno and Manson, 1993).

In contrast to most studies that demonstrate serum lipids and lipoproteins to be significantly altered as a consequence of the menopause (Moreno and Manson, 1993), in the present study group menopausal status alone had no significant effect on the lipid profile. This suggests that the changes in the lipid profile in this study group may be attributed to factors other than the menopausal status. No interaction was found between the apo E genotype and menopausal status indicating that the effect of apo E genotype on the lipid profile does not differ between the premenopausal and postmenopausal women. However, apo E genotype did interact with WHR for triglyceride and HDL-C, and with BMI for LDL-C (Chapter 4, Table XI, Pg 75).

The relationship of the apo E alleles to phenotype was reviewed in Chapter 2. In this study due to sample size limitations only the $\epsilon 3/3$ and the $\epsilon 3/4$ genotypes could be compared. Most studies have shown a predominant impact of the $\epsilon 3/4$ genotype on cholesterol levels (Cumming *et al.*, 1984; Davignon *et al.*, 1988; Hallman *et al.*, 1991; Eichner *et al.*, 1993); a significant and consistent relationship between the apo E phenotype and the plasma triglyceride was not demonstrated. A more recent meta-analysis of pooled data arrived at the conclusion that $\epsilon 4$ has a propensity to raise plasma triglyceride levels (Dallongeville *et al.*, 1992). In support of these results, previous reports have indicated that the $\epsilon 4$ allele may be implicated in the pathogenesis of type V hyperlipoproteinaemia, where a combination of genetic and acquired factors

often combine tendencies for the overproduction and delayed clearance of triglyceride-rich lipids (Kuusi *et al.*, 1988).

It is significant that migrant Indian populations have higher triglyceride levels than the host population in many parts of the world (McKeigue *et al*, 1989). This suggests the possibility that the apo E genotype may interact with another genetic trait to elevate triglyceride levels in the Indian community. Apo E4, has a greater preference for association with triglyceride-rich lipoproteins, compared with apo E3 resulting in delayed lipolysis observed in $\epsilon 4/3$ and a faster clearance of VLDL remnants in subjects carrying the $\epsilon 4/4$ genotype (Rall and Mahley, 1992). This could result in significantly lower triglyceride levels in the $\epsilon 4/4$ subset compared to persons bearing the $\epsilon 3/4$ genotype. This comparison was not possible in the present study because of the small number of observations with the $\epsilon 4/4$ genotype.

In most studies, the $\epsilon 4$ allele has been proposed as a risk factor for CHD. Higher cholesterol levels are thought to mediate this risk. The present study indicated that subjects carrying the $\epsilon 3/4$ genotype, in addition to elevated plasma cholesterol concentrations, have significantly higher triglyceride levels than subjects carrying the $\epsilon 3/3$ genotype. Hence, the serum lipid profile is significantly altered as a consequence of the apo E polymorphism, resulting in a more atherogenic lipid profile in individuals with the $\epsilon 3/4$ genotype. Together with the low frequency of the $\epsilon 2$ allele which has cardioprotective properties, this may contribute to a higher risk factor for CHD. Hence, these early results in the South African Indian population requires further epidemiological study to more clearly establish the relationship of apo E genotype to CHD risk in this vulnerable population.

CHAPTER 6

CONCLUSION

This study of 173 adult healthy Indian women confirmed the previous finding of a low $\epsilon 2$ frequency in an Indian population resident in Singapore. The data presented in this study provides strongly suggestive evidence that the frequency of the $\epsilon 2$ allele in the South African Indian female population is lower than most other populations around the world. The under-representation of the $\epsilon 2$ allele may contribute to the high risk of CHD. On the other hand the low prevalence of the $\epsilon 2$ allele may protect the South African Indian population against type III HLP precipitated by diabetes and obesity which are common in this community.

The study also clearly demonstrates that the major determinants of variation in triglyceride levels were apo E genotype and age. Age and apo E genotype also contribute to the variation in total cholesterol and LDL-C, but in this instance, age was by far the more significant. Waist:hip ratio contributed significantly to HDL-C and less clearly to apo A-I. The variables (apo E genotype, age and WHR) contributed minimally to the variance in apo A-I (4.6%) and apo B (2.6%), but strongly to triglyceride (22%), total cholesterol (18%) and to LDL-C (15%). These variables (particularly WHR) contributed 9.8% of the variance in HDL-C levels.

Menopausal status and BMI were not important predictors of plasma lipids in this study. Thus, any negative influence of the menopause *per se* on CHD risk in this Indian population does not appear to be mediated through the lipid profile. Apo E genotype and age, however, are significant determinants of potentially atherogenic changes in the plasma lipid profile. In particular the observation that the $\epsilon 3/4$ genotype is associated with a more atherogenic lipid profile supports the contention that it predisposes to CHD.

REFERENCES

Angelin B and Rudling M. Molecular aspects of human lipid metabolism. *Eur J Clin Nutrit*, 1992; **46:** 153-160.

Assmann G, Funke H and Jabs HU. Analytical procedures for the detection and characterization of apolipoprotein E mutants. *Am Heart J*, 1987; **113:** 598-603.

Assmann G, von Eckardstein A, Funke H, Rust S, and Sandkamp M. Analysis of genetically determined structural polymorphisms in apolipoproteins E, B and (a). *Curr Opin Lipidol*, 1991; **2:** 367-375.

Bailleul S, Couderc R, Landais A, Lefevre G, Raichvarg D and Etienne J. Direct phenotyping of human apolipoprotein E in plasma: Application to population frequency distribution in Paris (France). *Hum Hered*, 1993; **43:** 159-165.

Beardsley T. Unraveling Alzheimer's. *Sci Am*, 1993; **269:** 13.

Blair GE and Zajdel MEB. The polymerase chain reaction - already an established technique in biochemistry. *Biochem Educ*, 1992; **20:** 87-91.

Brasaemle DL, Cornely-Moss K and Bensadoun A. Hepatic lipase treatment of chylomicron remnants increases exposure of apolipoprotein E. *J Lipid Res*, 1993; **34:** 455-463.

Breslow JL. The genetic basis of lipoprotein disorders. Introduction and overview. *J Intern Med*, 1992; **231:** 627-631.

Brown TA. Manipulation of purified DNA. In *Gene Cloning*, Thetford Press Ltd, Great Britain, 1986, pp 43-66

Brown MS, Herz J, Kowal RC and Goldstein JL. The low-density lipoprotein receptor-related protein: double agent or decoy? *Curr Opin Lipidol*, 1991; **2:** 65-72

Bush T, Cowan L, Heiss G, Chambliss L, and Wallace R. Ovarian function and lipid/lipoprotein levels: Results from the Lipid Research Clinics Program. *Am J Epidemiol*, 1984; **120:** 489.

Bush TL, Fried LP and Connor EB. Cholesterol, lipoproteins, and coronary heart disease in women. *Clin Chem;* 1988; **34:** B60-B70.

Cruickshank JK. Diabetes: contrasts between peoples of black (West African), Indian and White European origin. In *Ethnic factors in health and disease.* Ed Cruickshank JK and Beevers DG. Butterworth and Co. Ltd, 1989; pp 289-304.

Cumming AM, and Robertson FW. Polymorphism at the apoprotein E locus in relation to risk of coronary disease. *Clin Genet*, 1984; **25:** 310-313.

Dallongeville J, Lussier-Cacan S and Davignon J. Modulation of plasma triglyceride levels by apo E phenotype: a meta-analysis. *J Lipid Res*, 1992; **33:** 447-454.

Davignon J, Gregg RE and Sing CF. Apolipoprotein E polymorphism and atherosclerosis, *Arteriosclerosis* 1988; **8:** 1-21.

Des HK, McPherson J, Bruns GA, Karathanasis SK and JL Breslow. Isolation, characterization and mapping to chromosome 19 of the human apolipoprotein E gene. *J Biol Chem*, 1985; **260:** 6240-6247

Durrington PN. Lipids and their metabolism. In *Hyperlipidaemia: Diagnosis and management,* Butterworth and Co. Ltd, London, 1989; pp 4-23.

Editorial; Coronary heart disease in Indians overseas, *Lancet*, 1986; **1:** 1307-1308.

Eggertsen G, Tegelman R, Ericsson S, Angelin B, Berglund L. Apolipoprotein E polymorphism in a healthy Swedish population: Variation of allele frequency with age and relation to serum lipid concentrations. *Clin Chem*, 1993; **39:** 2125-2129.

Eichner JE, Kuller LH, Orcharg TJ, Grandits GA, McCallum LM, Ferrell RE and JD Neaton. Relation of apolipoprotein E phenotype to myocardial infarction and mortality from coronary artery disease. *Am J Cardiol* 1993; **71:** 160-165.

Emi M, Wu LL, Robertson MA, Myers RL, Hegele RA, Williams RR, White R and Lalouel JM. Genotyping and sequence analysis of apolipoprotein E isoforms. *Genomics* 1988; **3:** 373-379

Eto M, Watanabe K and Ishii K. A racial difference in apolipoprotein E allele frequencies between the Japanese and Caucasian populations. *Clin Genet*, 1986, **34:** 258-264.

Fazio S, Horie Y, Weisgraber KH, Havekes LM and Rall SC. Preferential association of apolipoprotein E Leiden with very low density lipoproteins of human plasma. *J Lipid Res*, 1993; **34:** 447-453.

Filichkin SA and Gelvin SB. Effect of dimethyl sulfoxide concentration on specificity of primer matching in PCR. *BioTechniques*, 1992; **12:** 829-830.

Funke H, Rust S and Assmann G. Detection of apolipoprotein E variants by an oligonucleotide melting procedure. *Clin Chem,* 1986; **32**: 1285-1289.

Gelfand DH and White TJ. Thermostable DNA Polymerases. In PCR Protocols: A guide to methods and applications. ed. Innis M.A., Gelfand, D.H., Sninky, J.J., and T.J.White, Academic press. San Diego, California. 1990 p 129-140.

Gerdes LU, Klausen CIB, Sihm I and Faergeman O. Apolipoprotein E polymorphism in a Danish population compared to findings in 45 other study populations around the world. *Genet Epidemiol*, 1992; **9:** 155-167.

Glomset JA. The plasma lecithin:cholesterol acyltransferase reaction. *J Lipid Res,* 1968; **9:** 155-167.

Hallman MD, Boerwinkle E, Saha N, Sandholzer C, Menzel HJ, Csazar A and Utermann G. The apolipoprotein E polymorphism: A comparison of allele frequencies and effects in nine populations. *Am J Hum Genet*, 1991; **49:** 338-349.

Havel RJ and Kane JP. Primary dsybetalipoproteinemia: Predominance of a specific apoprotein species in triglyceride-rich lipoproteins. *Proc Natl Acad Sci*, 1973, **70:** 2015-2019.

Havel RJ, Kotite L, Kane JP, Tun P and Bersot T. Atypical familial dyslipoproteinaemia associated with apolipoprotein phenotype E3/3. *J Clin Invest*, 1983; **72:** 379-387.

Hixson JE, and Vernier DT. Restriction isotyping of human apolipoprotein E by gene amplification and cleavage with *Hha*I. *J lipid Res* 1990; **31:** 545-548.

Houlston RS, Wenham PR and Humphries SE. Detection of apolipoprotein E polymorphisms using PCR/ASO probes and Southern transfer: application for routine use. *Clin Chim Acta*, 1990; **189:** 153-158.

Innis MA and Gelfand DH; PCR Protocols. A guide to methods and applications; edited by Innis MA, Gelfand DH, Sninsky JJ and White TJ. Academic Press, New York; 1990.

James RW, Boemi M, Giansanti R, Fumelli P and Pometta D. Underexpression of the apolipoprotein E4 isoform in an Italian population. *Arterioscler Thromb,* 1993; **13:** 1456-1459.

Jensen J, Nilas L and Christiansen C. Influence of menopause on serum lipids and lipoproteins. *Maturitas*, 1990; **12:** 321-331.

Kamboh MI, Aston CE, Ferrell RE and Hamman RF. Impact of apolipoprotein E polymorphism in determining interindividual variation in total cholesterol and low density lipoprotein cholesterol in Hispanics and non-Hispanic whites. *Atherosclerosis*, 1993; **98:** 201-211.

Kannel WB, Hjortland MC, McNamara PM and Gordon T. Menopause and risk of cardiovascular disease. The Framingham study. *Ann Intern Med,* 1976; **85:** 447-452.

Kontula K, Aalto-Setala K, Kuusi T, Hamalainen L and Syvanen AC. Apolipoprotein E polymorphism determined by restriction enzyme analysis of DNA amplified by polymerase chain reaction: Convenient alternative to phenotyping by isoelectric focusing. *Clin Chem* 1990; **36:** 2087-2092.

Kroon PA and Powell EE. Liver, lipoproteins and disease: I. Biochemistry of lipoprotein metabolism. *J Gastroenterol and Hepatol,* 1992; **7:** 214-224.

Kuller LH, Gutai JP, Meilahn E, Matthews KA and Plantinga P. Relationship of endogenous sex steroid hormones to lipids and apolipoproteins in postmenopausal women. *Arteriosclerosis,* 1990; **10:** 1058-1066.

Kuusi T, Taskinen MR, Solakivi T and Kauppinen-Makelin R. Role of apolipoproteins E and C in type V hyperlipoproteinemia. *J Lipid Res,* 1988; **29:** 293-298.

Kuusi I, Nieminen SM, Ehnholm C, Yki-Jarvinen H, Valle M, Nikkila EA and Taskinen MR. Apoprotein E polymorphism and coronary artery disease. *Arteriosclerosis,* 1989; **9:** 237-241.

Lohse P, Brewer III B, Meng MS, LaRosa J and Brewer Jr HB. Familial apolipoprotein (apo) E deficiency: identification of a new kindred with a premature stop codon in the apo E gene. *Arteriosclerosis,* 1989; **9:** 706a (abstract).

Lohse P, Mann WA, Stein EA and Brewer Jr HB. Apolipoprotein E-$4_{Philadelphia}$ (GLU^{13}→Lys, Arg^{145}→Cys). Homozygosity for two rare point mutations in the apolipoprotein E gene combined with several type III hyperlipoproteinaemia. *J Biol Chem,* 1991; **266:** 10479--10484.

Longmire Jl, Albright KL, Lewis AK, Meincke LJ and Hildebrand CE. A rapid and simple method for the isolation of high molecular weight cellular and chromosome-specific DNA in solution without the use of organic solvents. *Nucleic Acids Res,* 1987; **15:** 859.

Mahley RW and Innerarity TL. Lipoprotein receptors and cholesterol homeostasis. *Biochim Biophys Acta*, 1983; **737:** 197-222.

Mahley RW. Apolipoprotein E: Cholesterol transport protein with expanding role in Cell Biology. *Science*, 1988; **240:** 622- 630.

Main BF, Jones PJH, MacGillivary RTA and Banfield DK. Apolipoprotein E genotyping using the polymerase chain reaction and allele-specific oligonucleotide primers. *J lipid Res,* 1991; **32:** 183-187.

Mann WA, Gregg RE, Sprecher DL and Brewer Jr HB. Apolipoprotein E-$1_{Harrisburg}$: a new variant of apolipoprotein E dominantly associated with type III hyperlipoproteinaemia. *Biochim Biophys Acta*, 1989; **1005:** 239-244.

Manttari M, Koskinen P, Ehnholm C, Huttunen JK and Manninen V. Apolipoprotein E polymorphism influences the serum cholesterol response to dietary intervention. *Metabolism*, 1991; **40:** 217-221.

McDowell IFW, Wisdom GB and Trimble ER. Apolipoprotein E phenotype determined by agarose gel electrofocusing and immunoblotting. *Clin Chem* 1989; **35:** 2070-273.

McKeigue PM, Miller GJ and Marmot MG. Coronary Heart Disease in South Asians overseas: A review. *J Clin Epidemiol*, 1989; **42:** 597-609.

Mercier B, Gaucher C, Feugas O and Mazurier C. Direct PCR from whole blood, without DNA extraction. *Nucleic Acids Res,* 1990; **18:** 5908.

Moreno GT and Manson JE. Cholesterol and coronary heart disease in women: an overview of primary and secondary prevention. *Coronary Artery Disease*, 1993; **4:** 580-587.

Mullis KB and Faloona FA. Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. *Methods Enzymol*, 1987; **155:** 335-350.

Omar MAK, Seedat MA, Dyer RB, Rajput MC, Motala AA and Joubert SM. The prevalence of diabetes mellitus in a large group of South Africans. *S Afr Med J*, 1985; **67:** 924-926.

Paik YK, Chang DJ, Reardon CA, Davies GE, Mahley RW and Taylor JM. Nucleotide sequence and structure of the human apolipoprotein E gene. *Proc Natl Acad Sci USA*, 1985; **82:** 3445-3449.

Payne MN, Green E, Walker MJ, Murray RG and Jones AF. Apolipoprotein $\epsilon 4$ and coronary artery disease. *Lancet*, 1992; **340:** 1350.

Poirier J, Davignon J, Bouthillier D, Kogan S, Bertrand P and Gauthier S. Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet*, 1993; **342:** 697-699.

Rall Jr SC, Weisgraber KH and Mahley RW. Human apolipoprotein E. The complete amino acid sequence. *J Biol Chem*, 1982; **257:** 4171-4178.

Rall SC and Mahley RW. Role of apolipoprotein E genetic variants in lipoprotein disorders. *J Internal Med*, 1992; **231:** 653-659.

Sambrook J, Fritsch EF, Maniatis T. Polyacrylamide gel electrophoresis. In *Molecular cloning - a laboratory manual.* 2nd edition. Ed Nolan C. Cold Spring Harbor Laboratory Press. 1989; pp 6.36-6.62.

Schaefer EJ, Gregg RE, Ghiselli G, Forte TM, Ordovas JM, Zech LA and Brewer HB. Familial apolipoprotein E deficiency. *J Clin Invest*, 1986; **78:** 1206-1219.

Scharf S J, Horn GT and Erlich HA. Direct cloning and sequence analysis of enzymatically amplified genomic sequences. *Science*, 1986; **233:** 1076-1078.

Seedat YK, Mayet FGH, Khan S, Somers SR and Joubert G. Risk factors for coronary heart disease in the Indians of Durban. *S Afr Med J*, 1990; **78:** 447-454.

Sewdarsen M, Vythilingum S, Jialal I, Desai RK and Becker P. Abnormalities in sex hormones are a risk factor for premature manifestation of coronary artery disease in South African Indian men. *Atherosclerosis*, 1990; **83:** 111-117.

Sewdarsen M, Vythilingum S, Jialal I and Becker P. Lipid and lipoprotein abnormalities in South African Indian men with myocardial infarction. *Cardiology*, 1991; **78:** 348-356

Shore VG and Shore B. Heterogeneity of human plasma very low density lipoproteins. Separation of species differing in protein components. *Biochemistry*, 1973; **12:** 502-507.

Siedel J, Hagele EO, Ziegenhorn J and Wahlefeld AW. Reagent for the enzymatic determination of serum total cholesterol with improved lipolytic efficiency. *Clin Chem*, 1983; **29:** 1075-1080.

Smeets HJM, Poddighe J, Stuyt PMJ, Stalenhoef AFH, Ropers HH and Wieringa B. Identification of apolipoprotein E polymorphism by using synthetic oligonucleotides. *J Lipid Res*, 1988; **29:** 1231-1237.

Smit M, van der Kooij-Meijs E, Frants RR, Havekes L and Klasen EC. Apolipoprotein gene cluster on chromosome 19. Definite localization of the APOC2 gene and the polymorphic *Hpa*I site associated with type III hyperlipoproteinemia. *Hum Genet*, 1988; **78:** 90-93.

Smit M, de Knijff P and van der Kooij-Meijs E *et al.* Genetic heterogeneity in familial dysbetalipoproteinemia. The E2(lys 146→gln) variant results in a dominant mode of inheritance. *J Lipid Res*, 1990; **31:** 45-53.

Snowden C, Houston RS, Arif MH, Laker MF, Humphries SE and Alberti KGMM. Disparity between apolipoprotein E phenotypes and genotypes (as determined by polymerase chain reaction and oligonucleotide probes) in patients with non-insulin-dependent diabetes mellitus. *Clin Chim Acta*, 1991; **196:** 49-58.

Steinmetz A. Phenotyping of human apolipoprotein E from whole blood by immunoblotting. *J Lipid Res*, 1987; **28**: 1364-70.

Stevenson JC, Crook D and Godsland IF. Influence of age and menopause on serum lipids and lipoproteins in healthy women. *Atherosclerosis*, 1993; **98:** 83-90.

Talmud PJ. Editorial comment: Genetics and molecular biology, *Curr Opin Lipidol*, 1992; **3:** II-4.

Tajima S, Yamamura T, Menju M and Yamamoto A. Analysis of apolipoprotein E7 (apolipoprotein E-Suita) gene from a patient with hyperlipoproteinaemia. *J Biochem*, 1989; **105:** 249-253.

Thompson MW, McInnes RR and Willard HF. Genes in populations. In *Genetics in Medicine*. Ed Wonsiewicz MJ. W.B. Saunders Co USA, 1991; pp 143-165.

Tikkanen MJ, Huttunen JK, Ehnholm C, Pietinen P. Apolipoprotein E4 homozygosity predisposes to serum cholesterol elevation during high fat diet. *Arteriosclerosis*, 1990; **10:** 285-288.

Utermann G, Jaeschke M and Menzel J. Familial hyperlipoproteinemia type III: deficiency of a specific apolipoprotein (apo E-III) in the very-low-density lipoproteins. *FEBS Lett*, 1975; **56:** 352-355.

Utermann G, Hees M and Steinmetz A. Polymorphism of apolipoprotein E and occurrence of dysbetalipoproteinaemia in man. *Nature*, 1977; **269:** 604-607.

Utermann G. Apolipoprotein E polymorphism in health and disease. *Am Heart J*, 1987; **113:** 433-439.

Wardell MR, Brennan SO, Janus ED, Fraser R and Carrell RW. Apolipoprotein E2-Christchurch (136Arg→Ser). New variant of human apolipoprotein E in a patient with type III hyperlipoproteinaemia. *J Clin Invest*, 1987; **80:** 483-490.

Wardell MR, Weisgraber KH, Havekes LM and Rall Jr SC. Apolipoprotein E3-Leiden contains a seven amino acid insertion that is a tandem repeat of residues 121 to 127. *J Biol Chem*, 1989; **264:** 21205-21210.

Warnick RG, Benderson J, Albers JJ, *et al.* Dextran sulfate-Mg^{2+} precipitation procedure for quantitation of high density lipoprotein cholesterol. *Clin Chem*, 1982; **28:** 1379-1388.

Wenham PR, Newton CR and Price WH. Analysis of apolipoprotein E genotypes by the Amplification Refactory Mutation System. *Clin Chem*, 1991; **37:** 241-24.

Wenham PR, Price WH, Blundell G. Apolipoprotein E genotyping by one -stage PCR. *Lancet*, 1991; **337:** 1158-1159.

Wenham PR, Sedky A and Spooner RJ. Apolipoprotein E phenotyping: a word of caution. *Ann Clin Biochem*, 1991; **28:** 599-605.

Wenham PR, McDowell IFW and Hodges VM. Rare apolipoprotein E variant identified in a patient with type III hyperlipidaemia. *Proceedings ACB National Meeting*, 1993 April 19-23; Birmingham UK, Association of Clinical Biochemists 1993.

Schneider WJ and Nimpf J. Lipoprotein receptors: old relatives and new arrivals. *Curr Opin Lipidol*, 1993; **4:** 205-209.

Wyndham CH. Trends with time of cardiovascular mortality rates in the populations of the RSA for the period 1968-1977. *S Afr Med J*, 1982; **61:** 987-993.

Xu CF, Boerwinkle E, Trikkanen MJ, Huttunen JK, Humphries SE and Talmud. Genetic variation at the apolipoprotein gene loci contribute to response of plasma lipids to dietary change. *Genet Epidemiol*, 1990; **7:** 261-275.

Zannis VI and Breslow JL. Human VLDL apo E isoprotein polymorphism is explained by genetic variation and post-translation modification. *Biochemistry*, 1981; **20:** 1033-1041.

Zannis VI, Beslow JL and Utermann G. Proposed nomenclature of apoE isoproteins, apoE genotypes, and phenotypes. *J Lipid Res*, 1982; **23:** 911-914.

PERSONAL COMMUNICATION

Williamson B. St Mary's Hospital Medical School, London.