Towards mapping QTLs influencing parameters of lipid metabolism in human twins

Marian Beekman

Towards mapping QTLs influencing parameters of lipid metabolism in human twins

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van de Rector Magnificus Dr. D.D. Breimer,
hoogleraar in de faculteit der Wiskunde en
Natuurwetenschappen en die der Geneeskunde,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 11 maart 2004
klokke 15:15 uur
door

Marian Beekman

geboren te Maarn in 1972

Promotiecommissie

Promotores:	Prof. dr. P.E. Slagboom
	Prof. dr. D.I. Boomsma (Vrije Universiteit, Amsterdam)
Copromotores:	Dr. B.T. Heijmans
	Dr. P. de Knijff
Referent:	Prof. dr. J. Kaprio (University of Helsinki, Finland)
Overige leden:	Prof. dr. R.R. Frants
	Prof. dr. E. de Geus (Vrije Universiteit, Amsterdam)

The studies presented in this thesis were performed at TNO Prevention and Health Leiden as well as the the Leiden University Medical Center. This work was financially supported by the *National Institutes of Health*.

ISBN 90-77595-14-7
Cover photo and design: Marco Veenstra
Printed by [Optima] Grafische Communicatie, Rotterdam

Contents

Chapter 1	General Introduction	7
Chapter 2	Heritabilities of apolipoprotein and lipid levels in three countries *Twin Research* 2002; 5: 87-97.	27
Chapter 3	A powerful and rapid approach to human genome scanning in small quantities of genomic DNA *Genetical Research* 2001; 77:129-134.	43
Chapter 4	Genome scan for loci influencing plasma levels of lipids and apolipoproteins	49
Chapter 5	Two-locus linkage analysis applied to putative quantitative trait loci for Lipoprotein(a) levels *Twin Research* 2003; 6: 322-324.	63
Chapter 6	Evidence for a QTL on chromosome 19 influencing LDL cholesterol levels in the general population *European Journal of Human Genetics* 2003; 11: 845-850.	67
Chapter 7	Combined association and linkage analysis applied to the *APOE* locus *Genetic Epidemiology* 2004: in press.	77
Chapter 8	Summary and General Discussion	89
Chapter 9	Samenvatting	97
	Abbreviations	105
	List of Publications	107
	Curriculum Vitae	109
	Appendix 1	111
	Appendix 2	113
	Appendix 3	115
	Appendix 4	121
	Appendix 5	127
	Appendix 6	131
	Appendix 7	141
	Nawoord	145

1 | General Introduction

Intermediate phenotypes of cardiovascular diseases

This thesis focuses on the unravelling of the genetic component of plasma levels of lipids and apolipoproteins in human as intermediate phenotypes of cardiovascular diseases (CVD). CVD are the leading causes of death in Western countries[1]. A primary cause of CVD is atherosclerosis, which is characterised by plaques (an accumulation of lipids, fibrotic tissue and cellular material) in the vessel wall. Growth of an atherosclerotic plaque will cause partial obstruction of the artery and rupture of plaques may trigger the formation of an artery-occluding thrombus, leading to cardiovascular events such as myocardial infarction, sudden cardiac death or ischaemic stroke.
Besides age, sex, cigarette smoking, alcohol consumption, physical inactivity and a positive family history as known risk factors for development of atherosclerosis, several other factors are involved in atherosclerotic plaque formation and rupture. For examples, processes including lipid metabolism, vasoconstriction and -dilation, inflammation, oxidation of lipoproteins, coagulation and insulin resistance. According to the response-to-injury hypothesis[2] the first step in atherosclerosis is dysfunction of the arterial endothelium due to accumulation of oxidised LDL particles as a consequence of elevated plasma levels of low-density lipoproteins (LDL) and reactive radicals. Since clinical endpoints of CVD are the result of the combination of these processes, the aetiology of CVD is complex. Changes in quantitative parameters, reflecting a single pathophysiologic process, for example an increase in lipid and apolipoprotein levels, can be considered as important and well-established intermediate phenotypes of CVD. Such intermediate phenotypes represent a single pathway of CVD development and may, therefore, be more accessible to genetic dissection than the clinical endpoints. To gain insight in the genetic component involved in CVD development, lipid and apolipoprotein levels are investigated as quantitative risk factors of CVD. The studies presented in this thesis are aimed at the identification of major genes involved in the regulation of lipid and apolipoprotein levels in human.

Outline of this thesis

The general introduction summarizes the knowledge of lipid and apolipoprotein levels, their contribution to CVD and the approaches available in the search for genes contributing to the variation in quantitative traits (Chapter 1). In the next chapters, the results of a heritability and a linkage study of plasma lipids and apolipoproteins in twin pairs are described. The subjects who took part in this study consist of adolescent and adult Dutch twin pairs, adult Swedish and Australian twin pairs, on whom previously data on plasma lipids and apolipoproteins were collected[3-6] The adolescent Dutch twin pair sample consists of 70 monozygotic (MZ) twin pairs and 90 dizygotic twin pairs, the adult Dutch twin pair sample of 91 MZ pairs and 117 DZ pairs, the Swedish twin pair sample of 115 MZ pairs and 187 DZ pairs, and the Australian twin pair sample of 711 MZ pairs and 651 DZ pairs. These twin data were used to estimate the heritabitities of variation in lipid and apolipoprotein levels and to investigate whether the influence of genes differs over populations (Chapter 2).

After establishing the role of genetic factors in lipid and apolipoprotein levels, a genome wide search was initiated to map chromosomal regions harbouring putative quantitative trait loci (QTL). The genome of DZ twin pairs was then scanned with evenly spaced, highly polymorphic short tandem repeats (STRs). Simulation studies were performed to determine the optimal spacing for STRs obtaining sufficient power to detect QTLs with a major effect. A new marker screening set was composed for this purpose (Chapter 3).

All DZ Dutch twin pairs were included in the linkage study and only the DZ Swedish twin pairs who were reared together (n=44). From the 651 DZ Australian twin pairs, 249 pairs were selected based on both twins being in the upper or lower 16% tail of the distribution of total cholesterol, apoB, or triglyceride levels, to investigate the most informative twin pairs[7]. Thus, the genome scan was performed in 493 DZ twin pairs (Table 1). By performing a genome scan, QTLs with a major effect on quantitative traits may be found. The power to find a QTL explaining 25% of the variance in lipid and apolipoprotein levels at a significant level of 0.05 in the adolescent Dutch twin pairs is 43%. In the adult Dutch twins 46%, in the Swedish twins 14% and in the Australian twin sample 59%. All samples combined showed 87% power to detect a such major QTL. The twin samples separately have not sufficient power for a linkage study. Separately, they may provide indications of replication of linkage, and combined, those four twin samples may have enough power to detect major loci involved in lipid and apolipoprotein levels.

Chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17, and 19 were analysed in an initial scan. These chromosomes harbour genes which are hypothesised to influence lipid and apolipoprotein levels (candidate genes). First, it was investigated whether two gene loci could be detected, which are known to have an effect on (apo)lipoprotein levels. These two genes are the *LPA* gene on chromosome 6q27, known to determine between 85 and 95% of the population variation in Lp(a) plasma levels[8], and the *APOE* gene on chromosome 19q13.32, known to determine between 10 and 20% of apoE plasma levels[9,10]. Second, these ten chromosomes were screened for the presence of loci, involved in the inter-individual variation in lipid or lipoprotein levels (Chapter 4).

Chapters 5-7 reflect four aspects of analysis necessary to optimise power for scans and subsequent fine-mappings aimed to localise QTLs: 1) simultaneous analysis of multiple chromosomal loci influencing the same trait, 2) simultaneous analysis of correlated phenotypes affected by the same QTL, 3) simultaneous analysis of linkage and association in identifying the

Table 1.
Overview of characteristics of the four twin samples.

	Adolescent Dutch twins	Adult Dutch twins	Swedish twins	Australian twins
Mean age (SD)	16.7 (2.0)	44.2 (6.5)	65.6 (8.5)	45.6 (11.3)
Phenotyped MZ twin pairs	70	92	115	711
Phenotyped DZ twins pairs	90	117	187	651
Genotyped DZ twin pairs	83	117	44	249
Available phenotypes:				
Total cholesterol	+	+	+	+
HDL cholesterol	+	+	+	+
Triglycerides	+	+	+	+
Lp(a)	+	+	+	+
ApoB	+	+	+	+
ApoE	+	+	-	+
ApoA1	+	+	+	+
ApoA2	+	-	-	+

gene variant responsible for the positive linkage result, 4) joint analysis of multiple study samples.

Simultaneous analysis of multiple chromosomal loci is exemplified by Lp(a) data. The presence of a putative second QTL influencing Lp(a) levels on chromosome 1, as was suggested by Broeckel et al[11], was investigated. To analyse the effects of both QTLs on chromosome 6 and 1 in one model, a two-locus model was employed (Chapter 5).

Simultaneous analysis of correlated phenotypes is exemplified by LDL cholesterol and apoB levels with markers on chromosome 19. Instead of searching for co-localisation of linkage peaks of correlated phenotypes, bivariate analyses were performed to investigate whether the putative QTL influencing LDL cholesterol levels has pleiotropic effects on apoB levels (Chapter 6).

Simultaneous analysis of linkage and association, which is crucial for fine-mapping steps, is exemplified by analysis of the *APOE* locus in its effect on apoE plasma levels. Although the APOEε2/ε3/ε4 polymorphism at the *APOE* gene is known to influence apoE levels, it is unclear whether additional genetic variation plays a role. Using the recently developed combined linkage-association approach[12,13], it can be established to what extent the tested polymorphism contributes to the linkage that was found. When the linkage can not be explained completely by the tested polymorphism, other variations in the chromosomal region should play a role (Chapter 7).

Finally, the results are summarised and discussed in the last chapter. A discussion of future perspectives for the identification of genes with a major effect on quantitative traits concludes this thesis (Chapter 8).

Lipid metabolism

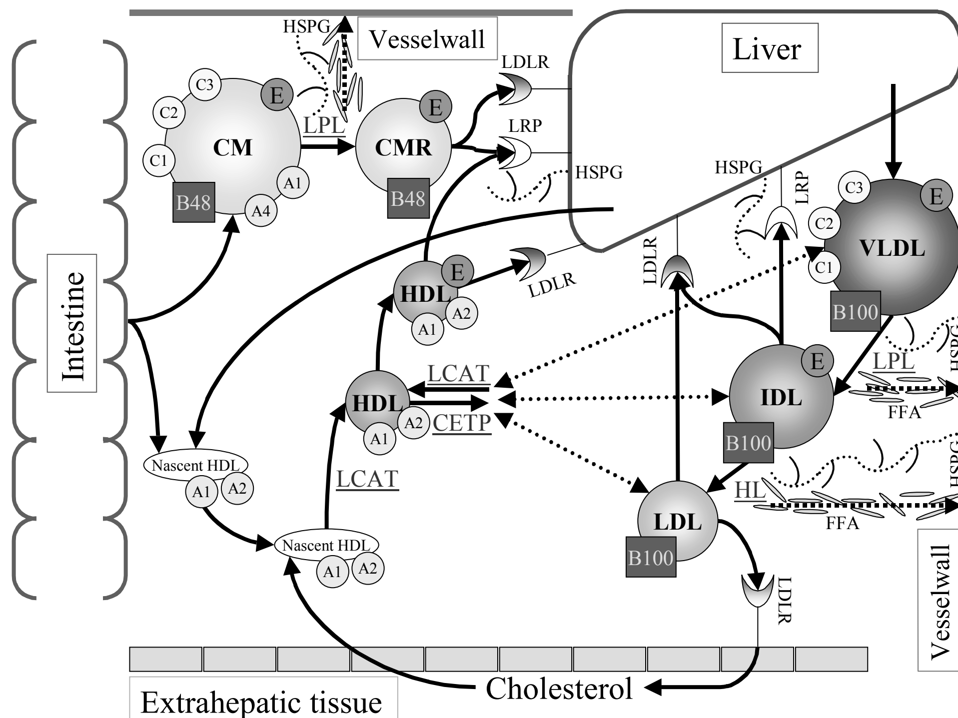
Human lipid metabolism hinges on the transport of cholesterol and triglycerides (TG), which are crucial in human cellular processes. TG are used in the form of free fatty acids (FFA) as energy source in peripheral tissues or it can be stored in adipose tissue. Cholesterol is a component of cell membranes and therefore essential for growth and viability of cells. Additionally, cholesterol

is a precursor of steroid hormones as well as vitamin D, and bile salts, derivatives of cholesterol, are essential for solubilisation of dietary lipids. Cholesterol and TG, the most common lipids of the human diet, are absorbed by the intestine and packed into water-soluble lipoproteins for their transport in the circulation. Lipoproteins are spherical macromolecular particles, of which the polar surface monolayer is composed of phospholipids, free cholesterol and several proteins called apolipoproteins (apo). The core contains hydrophobic lipids such as TG and cholesterol esters.

Lipoproteins can be divided into five classes, which differ in size, density, electrical charge and apolipoprotein composition. These lipoprotein classes are: 1. Chylomicrons, 2. Very low-density lipoproteins (VLDL), 3. Intermediate-density lipoproteins (IDL), 4. Low-density lipoproteins (LDL), 5. High-density lipoproteins (HDL). Each of these lipoprotein classes plays a specific role in the lipid metabolism, which can be subdivided in three different parts: 1. The exogenous pathway, 2. The endogenous pathway, 3. The reverse cholesterol pathway. These pathways[14-16] are depicted in figure 1.

Figure 1.

Lipid metabolism. The exogenous pathway involves the transport of dietary cholesterol and TG from the intestine in chylomicrons to the liver. The endogenous pathway involves the transport of cholesterol and TG packed into VLDL by the liver to other tissues via IDL and LDL. The reverse cholesterol pathway involves the transport of the excess of lipids in the extra-hepatic tissues back to the liver in HDL. Abbreviations: HSPG: Heparan sulphate proteoglycan, CM; chylomicron, CMR: chylomicron remnant, C1: apoC1, C2: apoC2, C3: apoC3, E: apoE, B48: apoB48, B100:apoB100, LDLR: low-density-lipoprotein receptor, LRP: LDLR related protein, VLDL: very low density lipoprotein, IDL: intermediate density lipoprotein, LDL: low density lipoprotein, HDL: High density lipoprotein, A1: apoA1, A2: apoA2, A4: apoA4, LPL: lipoprotein lipase, HL: hepatic lipase, LCAT: lecithin:cholesterol acyltransferase, CETP: cholesterol ester transfer protein, FFA: free fatty acids.



In the exogenous pathway, dietary cholesterol and TG are transported from the intestine in chylomicrons to the liver. In the endogenous pathway, cholesterol and TG are packed into VLDL by the liver and transported to other tissues via IDL and LDL. Since apoB is a major constituent of LDL particles, these plasma levels are highly correlated. In the reverse cholesterol pathway, the excess of lipids in the extra-hepatic tissues are transported back to the liver in HDL. Since apoA1 and apoA2 are major constituents of HDL particles, these plasma levels are highly correlated. An outline of these pathways and the proteins and enzymes involved is provided in Appendix 1.

Disturbances in lipid levels contribute to atherosclerosis

In a healthy arterial wall, there is balance between the uptake of cholesterol from LDL particles and the excretion of cholesterol by the vessel wall into the circulation. Development of atherosclerosis is initiated when this balance is shifted towards an increased uptake of LDL and/or decreased excretion of HDL. Subsequently, macrophages will accumulate the free cholesterol and cholesterol esters derived from LDL turning them into foam cells, that finally will form the core of an atherosclerotic plaque. The most important way for a macrophage to dispose of excess cholesterol is through active cholesterol efflux via membrane transporters with HDL serving as the major extracellular cholesterol acceptor. Hence, the balance of lipid and apolipoprotein levels in both the endogenous and the reverse cholesterol pathways play an important role in the pathogenesis of atherosclerosis.

Heritability of the plasma levels of lipids and apolipoproteins

In disentangling the genetics of lipid and apolipoprotein levels, it is relevant to know how large the genetic effect is on variation in these quantitative traits. The proportion of the population variation that is attributable to genetic variation, called heritability, can be estimated from family as well as twin data. To obtain a first estimate of trait heritability from twin data, the difference of the correlation (r) for the trait in monozygotic (MZ) twins (rMZ) and the correlation for the trait in dizygotic (DZ) twins (rDZ) is multiplied by 2, which can be written as 2(rMZ-rDZ). Differences between MZ twins are assumed to be entirely of environmental origin, since MZ twins have identical DNA, apart from somatic mutations, asymmetric transmission of mitochondria, skewed X inactivation or other genetic causes[17,18]. Differences between DZ twins can originate from environmental factors as well as genetic factors, since DZ twins share on average 50% of their segregating genes. Because the influence of the environment on traits is assumed to be equal in MZ and DZ twins, the difference in trait correlations of MZ and DZ twins reflects the genetic component of the trait. More sophisticated estimates of heritability and other parameters can be obtained by the use of maximum likelihood techniques, as can be implemented in software packages as Mx[19]. Employing a design including MZ and DZ twins reared together, allows decomposition of the trait variance into components of additive background genetic variance, shared environmental variance and unique environmental variance. When STR marker data are added to the data, the decomposition of the variance can be extended with a QTL variance component. This generalised methodology also allows formal tests for differences in parameter estimates between, for example, sexes or cohorts.
Several heritability studies on apolipoprotein and lipid levels reported heritability estimates between 40% and 80% for total cholesterol, LDL cholesterol, HDL cholesterol and triglyceride levels[20,21]. The estimates for the genetic contribution to levels of apolipoproteins is less well

established. Most studies lack data on apoE and apoA2 levels. Heritability estimates of apoB and apoA1 levels are also between 40% and 80%[22]. Sex, age and ethnicity do not seem to affect heritability of lipid and apolipoprotein levels[20,23]. In general, the heritability studies show that variation in lipid and apolipoprotein levels are to a large extent determined by genetic factors. How many and which gene loci are involved, however, remains largely elusive.

Genetic dissection approaches

A number of genes encoding receptors, apolipoproteins, enzymes and transfer proteins in the lipid metabolism, have been identified in familial lipoprotein disorders (see Appendix 2), characterised by major increased or decreased levels of specific (apo)lipoproteins. Families with genetic lipoprotein disorders form approximately 10% of the general population[24]. This implies that mutations causing major changes in lipoprotein levels are relatively rare and consequently play a limited role in determination of lipid and apolipoprotein levels in the population at large. Either milder variants in the genes listed in Appendix 2 or other genes may be hypothesised to determine variance of lipid and apolipoprotein levels in the population at large. Essentially two approaches can be used in an attempt to identify these genes: testing hypothesised genes (candidate genes) by genetic association studies and a systematic testing of chromosomal positions (genome scan) by genetic linkage studies.

In classical genetic association studies the frequency of one or a set of gene variants is compared in groups of unrelated cases and unrelated controls. The variant to be studied (single nucleotide polymorphism (SNP) or a polymorphism based on repeat number variation) may be a neutral variation or a functional variant in promoter, exon, 5' or 3' untranslated region affecting gene transcription and/or translation. The cases and controls in lipid studies have mainly been CVD patients and controls, or subjects with high versus low lipid levels. An alternative, and frequently used method of association studies with quantitative traits, is comparison of mean trait levels per genotype group.

In genetic linkage studies one establishes co-inheritance of the phenotype and alleles at polymorphic loci in members of a family. This design allows an estimation of the chromosomal position of genes to be identified for a trait, by scanning the complete genome with 300 to 400 neutral polymorphic markers in sibling pairs or families. By conducting a genome scan, it may be possible to identify chromosomal regions not previously recognised as sources of inter-individual variation in disease or intermediate traits. Genome scans have been performed in families of CVD patients or patients with some lipoprotein disorder and in population based families of healthy subjects.

Candidate gene approach

Table 2 gives an indication of which genes have been tested for association with lipid and apolipoprotein levels and whether association has been found. Mainly, genes involved in familial lipoprotein disorders were investigated for putative effects on lipid and apolipoprotein levels in the general population. Additionally, variants in multiple genes encoding molecules involved in lipid metabolism were tested for association with lipid or apolipoprotein levels. Most of the polymorphisms examined show no significant association. However, polymorphisms in LDLR seems to have small effects on LDL cholesterol and apoB levels, and LPL variants seem to have a small effect on HDL cholesterol levels in the general population (5/5; 5 significant studies have large sample sizes and 5 non significant have small sample size). The success of replication of association studies seems limited due to the small effect sizes of the genes.

Table 2.
Overview of candidate gene studies with cardiovascular disease, lipid and apolipoprotein levels. Reported are the numbers of significant (sign) associations versus no significant (ns) findings. Associations are indicated with a grey background when significant associations occur as much or more often than non significant findings. ++ indicates that the effect of the gene on the phenotype has been established.

	Phenotype	TC	LDL-C	ApoB	HDL-C	ApoA1	ApoA2	TG	Lp(a)	ApoE	References
Gene	Locus	sign/ns	sign/ns	sign/ns	sign/ns	sign/ns	sign/ns	sign/ns	sign/ns	sign/ns	
LEPR	1p31	-	-	0/1	-	-	-	-	-	-	[34]
APOA2	1q21	-	-	-	-	-	-	1/0	-	-	[35]
APOB	2p24	1/3	1/2	1/3	0/4	0/2	-	0/2	-	-	[36-39]
MTP	4q24	3/6	2/8	0/7	0/8	0/1	-	1/7	-	-	[39-42]
FABP2	4q28	-	-	-	0/1	-	-	0/1	-	-	[43]
LPA	6q27	-	-	-	-	-	-	-	++	-	[8,44]
PON1	7q21.3	0/1	2/0	-	0/1	-	-	0/1	-	-	[36,45,46]
PON2	7q21.3	1/0	0/1	0/1	0/1	1/0	-	0/1	-	-	[46]
LPL	8p22	4/8	3/6	2/0	5/5	0/2	-	3/8	-	-	[36,45,47-54]
CYP7	8q11	1/3	2/2	0/4	1/3	1/3	-	1/3	-	-	[55,56]
VLDLR	9p24	0/1	0/1	1/1	0/1	-	-	0/1	-	-	[36,57]
APOA1	11q23	0/4	0/3	-	0/4	0/1	-	0/2	-	-	[37,58-60]
APOC3	11q23	1/10	1/9	1/4	4/8	0/4	-	3/7	-	-	[36-38,54,58,60,61]
LRP1	12p13.3	0/1	0/1	0/1	0/1	0/1	-	0/1	-	-	[36]
SR-B1	12q24.3	1/4	2/3	-	1/4	-	-	2/3	-	-	[62]
F7	13q12.3	0/2	0/2	0/2	0/2	0/2	-	0/2	-	-	[36,63,64]
LIPC	15q21	0/6	2/5	0/2	2/7	1/2	-	1/9	-	-	[36,39,47,49,52,54,65,66]
CETP	16q13	0/3	0/3	0/1	2/1	-	-	0/3	-	-	[39,47,54]
APOH	17q23	0/1	0/1	-	0/1	-	-	0/1	-	-	[60]
LDLR	19p13.2	3/5	8/2	3/0	1/7	0/1	0/1	0/8	0/1	-	[36,45,47,67-70]
APOE	19q13.2	++	++	++	-	-	-	-	-	++	[25]
APOC1	19q13.2	0/1	0/1	0/1	0/1	-	-	0/1	-	-	[38]
APOC2	19q13.2	1/0	0/1	-	0/1	-	-	1/0	-	-	[71]

Since the APOE gene is already very well investigated and reviewed, showing a clear influence on levels of total cholesterol, LDL cholesterol, apoB and apoE in the general population, only the generally known effects are indicated in table 2. APOE explains approximately 4-8% of the variance in levels of total cholesterol, LDL cholesterol and apoB, and approximately 10-20% of the variance in apoE plasma levels[25].
Gene variants with major effects on quantitative intermediate phenotypes of CVD still have to be identified, if they exist. About 900 genes of the ~8,000 known genes are known candidate genes for affecting lipid and apolipoprotein levels (Celera database, GO-database, 9-6-2003). However, since the total number of human genes approximates 30,000[26,27], unknown genes influencing lipid metabolism will be missed when solely relying on the candidate gene approach.

Systematic genome scan

Table 3 lists chromosomal regions which have shown positive linkage with lipid or apolipoprotein levels, resulting from a genome scan in families either selected for lipoprotein disorders or for CVD, or population based families. On each chromosome, linkage has been found with some lipid related trait. Replication of results is necessary to gain confidence that the chromosomal region harbours a gene of interest. In this respect, most promising seems chromosome 1 with replicated linkage around 40 cM with LDL cholesterol levels, around 80 cM with familial hypercholesterolaemia (FH) and LDL-C/HDL-C ratio and around 200 cM with HDL cholesterol levels. Putative TG-QTLs seem to be present on chromosome 3 and 7. On chromosome 11, two severe lipoprotein disorders are mapped and on chromosome 13 a putative HDL cholesterol QTL. Chromosome 19, harbouring LDLR and APOE, seems to be involved in LDL cholesterol levels, but the data are not as clear as could be expected from the association studies. Also, only one study shows linkage of a low HDL cholesterol phenotype with chromosome 8, harbouring the LPL gene which is involved in HDL cholesterol levels (Table 2). Overall, linkage results from a genome scan for complex diseases are rarely replicated[28]. This could be due to differences in sample size, in ethnicity of population samples, in age and sex of population samples, in phenotype assignment or in statistical methods. These differences between studies makes the comparison of linkage results difficult. Moreover, genome scans give broad estimates of the location of the disease-causing genes on the chromosomes and these estimations may be many centiMorgans (cM) from the true locus[29], so it may be difficult to determine which results constitute a replication and which results do not. Since often many genes can be considered as candidate gene in the region showing positive linkage, the in depth analysis of these positive linkage regions from scans require association analysis of candidate genes and linkage analysis to be merged.

Statistical analysis

Essential for this thesis, is the principle of linkage analysis of quantitative traits in siblings (DZ twins in our case). Sib pair analysis for quantitative traits is based on the co-inheritance of trait values in combination with allele sharing at a marker locus. When the variation in the trait has a genetic origin, it is expected that sib pairs, sharing alleles identical by descent (IBD) at a marker locus close to the responsible locus, will also share the phenotype. Sib pairs who share alleles IBD are expected to resemble each other more than siblings who do not share alleles IBD, or only share part of the marker alleles IBD. Each individual inherits two alleles of a marker; one from his father and one from his mother. By genotyping sib pairs and their parents, it can be

Table 3.
Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
1	16-49	D1S214/D1S228	LDL-C	2.43 (28cM)	Siblings with type 2diabetes	72
1	35-55	D1S199/D1S2843	LDL-C, ARH	3.07 (45 cM)	Syrian family (high LDL)	73
1	~40	D1S552	TC, LDL-C	P<0.01	Obese families	74
1	73-90	D1S233/D1S193	LDL-C/HDL-C	1.89 (81cM)	Siblings with type 2diabetes	72
1	~72-90	D1S2130/D1S1596	FH	6.8 (~75 cM)	FH family	75
1	~90	D1S405	TG, HDL-C	P<0.01	Obese families	74
1	~100	D1S1665	Elevated apoB	4.7 (102 cM)	Dutch FH families	34
1	170	D1S1679	Lp(a)	3.8 (170 cM)	MI families	11
1	190-210	D1S305/APOA2	HDL-C	1.45 (198cM)	Siblings with type 2diabetes	72
1		D1S104/D1S1677	FCH	5.93	Finnish FCH families	76
1		D1S104/D1S1677	TG	3.25	Finnish FCH families	76
1		D1S104/D1S1677	FCH	2.52	NHLBI FHS	77
1		D1S104/D1S1677	TG	1.86	NHLBI FHS	77
1	~195	D1S484	TC	P<0.01	Obese families	74
1	198		HDL-C	2.13	Caucasian HyperGEN	78
1	202	D1S518	HDL-C	1.45 (202 cM)	MI families	11
2	16		Total Cholesterol	2.19	Caucasian HyperGEN	78
2	22	D2S423	FCH	3.4	Finnish pooled low-HDL-C and FCH families	79
2	22	D2S423	TG	1.3	Finnish and Dutch FCH families	80

Table 3-continued.
Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
2	22	D2S423	HDL-C	1.2	Finnish and Dutch FCH families	80
2	22	D2S423	FCH	2.2	Finnish and Dutch FCH families	80
2	~50		TG	1.7 (~20 cM)	Pima Indians: isolate	81
2	47-58	D1S199	Age at onset CHD	2.62 (52 cM)	Indo-Mauritian families: isolate	82
2	125	D2S410	TG	Gwp-value= 0.035	South Dakota Hutterites: isolate	83
2	140	D2S1790	HDL-UC	2.28	San Antonio FHS Mexican American families	84
3	~20	D3S2397	TG	P<0.01	Obese families	74
3	~92		TG	1.77 (~92 cM)	Pima Indians: isolate	81
3	113-143	D3S1271	TG/HDL-C	2.10 (130 cM)	Indo-Mauritian families: isolate	82
3	140	D3S4529	Ln(TG/HDL-C)	1.8 (140 cM)	Framingham Study	85
3	~175		HDL-C	2.64 (~175 cM)	Pima Indians: isolate	81
3	244	D3S2436	LDL3-C	4.11 (244 cM)	San Antonio FHS Mexican American families	86
4	78	D4S3248	HDL-UC	2.61	San Antonio FHS Mexican American families	84
4	126	D4S2623	LDL3-C	4.11 (126 cM)	San Antonio FHS Mexican American families	86
4	169		ApoA2	2.35	Rochester FHS	87
5	2		LDL-C	1.89	NHLBI FHS	88
5	40	D5S1470	HDL-C	3.64	NHLBI FHS	89
5	48		HDL-C	2.74	Caucasian HyperGEN	78
5	74-83	D5S427	Total Cholesterol	2.09 (79 cM)	Siblings with type 2diabetes	72

Table 3-continued.

Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
5	79		ApoA2	2.13	Rochester FHS	87
5	186	D5S1456	HDL-UC	2.81	San Antonio FHS Mexican American families	84
5	191-201	D5S211/D5S408	LDL-C	1.97 (195 cM)	Siblings with type 2 diabetes	72
5	197		TC/HDL-C	1.57	Rochester FHS	87
6	64-81	D6S282	HDL-C	1.95 (72 cM)	Indo-Mauritian families: isolate	82
6	109		ApoC3	1.7	Rochester FHS	87
6	162	D6S1277	LDL3-C	2.92 (162 cM)	San Antonio FHS Mexican American families	86
6	~194	D6S297	LDL-C	P<0.01	Obese families	74
7	117	D7S653/D7S479	HDL-C	1.6	San Antonio FDS Mexican American families	90
7	~157	D7S530	TG	P<0.01	Obese families	74
7	155	D7S1805	Ln(TG/HDL-C), LN(TG)	2.5 (155 cM)	Framingham Study	85
8	68	D8S1477	HDL-UC	2.07	San Antonio FHS Mexican American families	84
8	119	D8S1132	Low HDL-C	4.7	Finnish pooled low-HDL-C and FCH families	79
8	150	D8S1128	HDL-UC	4.87	San Antonio FHS Mexican American families	84
9	35	D9S921	HDL-C	2.1	Finnish and Dutch FCH families	80
9	41	D9S925/D9S741	HDL-C	3.4	San Antonio FDS Mexican American families	90
9	64	D9S301/D9S175	HDL-C	2.4	San Antonio FHS Mexican American families	90
9	76	D9S1122	TG	1.9 (76 cM)	MI families	11
9	q34		BSCL	2.4	BSCL families (no adipose tissue. insulin resistance)	91

Table 3-continued.

Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
10	~105	D10S535	TG	P<0.01	Obese families	74
10	127		LDL-C	2.47	NHLBI FHS	88
10	130-140	D10S185	Age at onset CHD	2.03 (137 cM)	Indo-Mauritian families: isolate	82
10	~145	D10S670	Total cholesterol	P<0.01	Obese families	74
11	34		Total cholesterol	1.84	Rochester FHS	87
11	55-70	D11S4076/D11S480	BSCL	13.2	BSCL families (no adipose tissue. insulin resistance)	91
11	56		LDL-C	3.72	NHLBI FHS	88
11	~68	D11S1335	LDL-C	P<0.01	Obese families	74
11	~68	D11S1324/D11S2371	FCH	2.6	Dutch FCH families	92
11	121	D11S925	HDL-C	3.48	Early onset CVD families	93
12	96	PAH	HDL-UC	2.13	San Antonio FHS Mexican American families	84
12	128		ApoA1	2.02	Rochester FHS	87
12	~130	D12SPAH	Ln(TG)	P<0.01	Obese families	74
13	28	D13S1493	HDL-C	2.36	NHLBI FHS	89
13	33-49	D13S171/D13S263	HDL-C	1.93 (42 cM)	Siblings with type 2 diabetes	72
13	~70-85	D13s1267	ARH	3.08 (82 cM)	Syrian families (high LDL-C)	73
14	35	D14S608	LDL-C	1.5 (35 cM)	MI families	11
14	~130	D14S1426	MI	3.9 (130 cM)	MI families	11
15	18	GABRB3/D15S165	TG	3.9	San Antonio FDS Mexican American families	94

Table 3-continued.

Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
15	62	D15S643	HDL-UC	3.26	San Antonio FHS Mexican American families	84
15	80	D15S653	HDL-UC	2.54	San Antonio FHS Mexican American families	84
16	70	GATA67G11	Ln(TG)	1.5 (70 cM)	Framingham Study	85
16	95	D16S518	HDL-C	3.4	Finnish and Dutch FCH families	80
17	~65	D17S1291	TG	P<0.01	Obese families	74
17	89	D17S129	LDL-C	2.29 (89 cM)	MI families	11
17	~110	D17S937	FH	2.58 (~110 cM)	FH family	75
17	116		LDL-C	2.33	NHLBI FHS	88
17	126		TC/HDL-C	2.48	Rochester FHS	87
18	13		ApoA2	1.53	Rochester FHS	87
18	28	D18S843	Lp(a)	Gwp-value= 0.029	South Dakota Hutterites: isolate	83
18	28		ApoC2	1.55	Rochester FHS	87
18	55		ApoC3	1.51	Rochester FHS	87
19	0-14	D19S247/D19S209	TG/HLD-C	1.92 (11 cM)	Siblings with type 2 diabetes	72
19	~20	D19S1034	TC	3.89 (~20 cM)	Pima Indians: isolate	81
19	38		LDL1-C	2.26 (38 cM)	San Antonio FHS Mexican American families	86
19	52	D19S433	LDL-C	Gwp-value= 0.035	South Dakota Hutterites: isolate	83
19	62		LDL2-C	1.86 (62 cM)	San Antonio FHS Mexican American families	86
19	74-80	D19S178/APOC2	TG	3.16 (78 cM)	Siblings with type 2 diabetes	72

Table 3-continued.

Overview of genome scans for linkage with clinical endpoints or intermediate phenotypes of cardiovascular disease. ARH= autosomal recessive hypercholesterolaemia; Gwp-value= genome wide p-value; BSCL= Berardinelli-Seip syndrome; HDL-UC= HDL unesterified cholesterol; FHS= Family Heart Study; FDS= Family Diabetes Study.

Chromosome	cM from pter	Marker nearby	Phenotype	LOD score	Population	Reference
19	78		ApoE	4.2	Rochester FHS	87
20	29		Ln(TG)	2.77	African American HyperGEN	78
20	35	D20S604	Ln(TG/HDL-C)	1.7 (35 cM)	Framingham Study	85
20	~87	D20S476	TG	P<0.01	Obese families	74
21	29-48	D21S263	LDL-C/HDLC	1.95 (39 cM)	Indo-Mauritian families: isolate	82
21	45		LDL-C	2.74	NHLBI FHS	88
21	48		Total cholesterol	2.25	African American HyperGEN	78
22	~60	D22D1161	FH	1.98 (~60 cM)	FH family	75

established whether they share no alleles IBD (IBD=0), one allele IBD (IBD=1), or two alleles IBD (IBD=2) for a particular marker. When parents are not genotyped, as is the case for this thesis, the probability that sib pairs are IBD=0, IBD=1 or IBD=2 at a certain locus is estimated using the population frequencies of the marker alleles[30]. The sum of these probabilities equals 1.0. Under random Mendelian segregation, siblings have 25% chance to be IBD0, 50% chance to be IBD1 and 25% chance to be IBD2. Evidence for linkage of a locus with a trait is found when more than 25% of the concordant sibs are IBD2 and when more than 25% of the discordant sibs are IBD=0.
For analysing linkage of quantitative traits in sib pairs, several methods are available. Haseman and Elston[31] developed a method to detect QTLs that is based on regression analysis. This traditional HE-analysis is based on regressing the squared phenotypic differences of the twins on the proportion of alleles at a certain locus shared IBD, which is called π. The estimate of π is referred to as $\hat{\pi}$, which can be calculated as $\hat{\pi}$ =0.5*P(IBD=1)+P(IBD=2). The original HE method is robust, but not very powerful. Through the years, this approach has been improved[32] and generalised into a score test[33]
Variance components analysis, on the other hand, is a more powerful tool in quantitative linkage analysis. The total variance of a trait can be divided into variance which is attributable to shared factors, modelled as additive genetic background factors, unique environmental factors and a specific locus of which marker information is available (putative QTL). The maximum likelihood estimates for these variance components are calculated and the significance of the QTL effect is tested by comparing the model including the putative QTL effect, with the model without the QTL variance component. When there is one degree of freedom difference between the two models, the difference of the -2 log likelihoods of the two models is distributed as chi-squared statistic, which can be converted to a LOD score (10Logarithm of the odds ratio between likelihood of linkage and the likelihood of no linkage) by dividing it by 2*Ln(10). The advantage of variance components analysis is that both the between and within sib pair phenotypic variances can be considered, which allows to test for population stratification[12].
In this thesis, we studied adolescent and adult Dutch, Swedish and Australian twin pair samples to estimate the heritability of lipid and apolipoprotein levels using data on both the MZ and DZ twins in a variance component model. The DZ twins of these twin samples are genotyped with highly polymorphic markers to subsequently search for linkage of any chromosomal region with the highly heritable lipid and apolipoprotein levels using the variance components linkage analysis. To enlarge the power to detect a major QTL, the four twin samples are analysed simultaneously. Additional advanced analyses of genome scan data, like the simultaneous analysis of two loci influencing the same trait, the pleiotropic effects of a putative QTL and the combined analysis of linkage and association for fine mapping of the chromosomal region showing positive linkage are handled in this thesis.

References

1. http://www.who.int/en.
2. Glass CK, Witztum JL. Atherosclerosis. the road ahead. *Cell* 2001; 104: 503-516.
3. Boomsma DI, Kempen HJ, Gevers LJ, Havekes L, de Knijff P, Frants RR. Genetic analysis of sex and generation differences in plasma lipid, lipoprotein, and apolipoprotein levels in adolescent twins and their parents. *Genet Epidemiol* 1996; 13: 49-60.
4. Snieder H, Boomsma DI, van Doornen LJ, de Geus EJ. Heritability of respiratory sinus arrhythmia: dependency on task and respiration rate. *Psychophysiology* 1997; 34: 317-328.

5. Heath AC, Bucholz KK, Madden PA, Dinwiddie SH, Slutske WS *et al.* Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol Med* 1997; 27: 1381-1396.
6. Pedersen NL, McClearn GE, Plomin R, Nesselroade JR, Berg S, DeFaire U. The Swedish Adoption Twin Study of Aging: an update. *Acta Genet Med Gemellol (Roma)* 1991; 40: 7-20.
7. Risch NJ, Zhang H. Mapping quantitative trait loci with extreme discordant sib pairs: sampling considerations. *Am J Hum Genet* 1996; 58: 836-843.
8. Boerwinkle E, Leffert CC, Lin J, Lackner C, Chiesa G, Hobbs HH. Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations. *J Clin Invest* 1992; 90: 52-60.
9. Neale MC, de Knijff P, Havekes LM, Boomsma DI. ApoE polymorphism accounts for only part of the genetic variation in quantitative ApoE levels. *Genet Epidemiol* 2000; 18: 331-340.
10. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet* 1988; 42: 104-112.
11. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ *et al.* A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002; 30: 210-214.
12. Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 1999; 64: 259-267.
13. Cardon LR, Abecasis GR. Some properties of a variance components model for fine-mapping quantitative trait loci. *Behav Genet* 2000; 30: 235-243.
14. Gotto AM, Jr., Pownall HJ, Havel RJ. Introduction to the plasma lipoproteins. *Methods Enzymol* 1986; 128: 3-41.
15. Mahley RW, Innerarity TL, Rall SC, Jr., Weisgraber KH. Plasma lipoproteins: apolipoprotein structure and function. *J Lipid Res* 1984; 25: 1277-1294.
16. Breslow JL. Apolipoprotein genetic variation and human disease. *Physiol Rev* 1988; 68: 85-132.
17. Martin N, Boomsma D, Machin G. A twin-pronged attack on complex traits. *Nat Genet* 1997; 17: 387-392.
18. Petronis A. Human morbid genetics revisited: relevance of epigenetics. *Trends Genet* 2001; 17: 142-146.
19. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
20. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.
21. Hamsten A, Iselius L, Dahlen G, de Faire U. Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. *Atherosclerosis* 1986; 60: 199-208.
22. Lamon-Fava S, Jimenez D, Christian JC, Fabsitz RR, Reed T *et al.* The NHLBI Twin Study: heritability of apolipoprotein A-I, B, and low density lipoprotein subclasses and concordance for lipoprotein(a). *Atherosclerosis* 1991; 91: 97-106.
23. Snieder H, van Doornen LJ, Boomsma DI. Dissecting the genetic architecture of lipids, lipoproteins, and apolipoproteins: lessons from twin studies. *Arterioscler Thromb Vasc Biol* 1999; 19: 2826-2834.
24. Williams RR, Hopkins PN, Hunt SC, Wu LL, Hasstedt SJ *et al.* Population-based frequency of dyslipidemia syndromes in coronary-prone families in Utah. *Arch Intern Med* 1990; 150: 582-588.
25. Frikke-Schmidt R. Context-dependent and invariant associations between APOE genotype and levels of lipoproteins and risk of ischemic heart disease: a review. *Scand J Clin Lab Invest Suppl* 2000; 3-25.
26. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ *et al.* The sequence of the human genome. *Science* 2001; 291: 1304-1351.
27. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC *et al.* Initial sequencing and analysis of the human genome. *Nature* 2001; 409: 860-921.
28. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001; 69: 936-950.
29. Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS. Replication of linkage studies of complex traits: an examination of variation in location estimates. *Am J Hum Genet* 1999; 65: 876-884.
30. Broman KW. Estimation of allele frequencies with data on sibships. *Genet Epidemiol* 2001; 20: 307-315.
31. Haseman JK, Elston RC. The investigation of linkage between a quantitative trait and a marker locus. *Behav Genet* 1972; 2: 3-19.
32. Fulker DW, Cherny SS. An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet* 1996; 26: 527-532.
33. Putter H, Sandkuijl LA, van Houwelingen JC. Score test for detecting linkage to quantitative traits. *Genet Epidemiol* 2002; 22: 345-355.
34. Allayee H, Krass KL, Pajukanta P, Cantor RM, van der Kallen CJ *et al.* Locus for elevated apolipoprotein B levels on chromosome 1p31 in families with familial combined hyperlipidemia. *Circ Res* 2002; 90: 926-931.
35. Hong SH, Kang BY, Park WH, Kim JQ, Lee CC. Association between apolipoprotein A2 MspI polymorphism and hypertriglyceridemia in Koreans. *Hum Biol* 1998; 70: 41-46.
36. Hegele RA, Brunt JH, Connelly PW. Multiple genetic determinants of variation of plasma lipoproteins in Alberta Hutterites. *Arterioscler Thromb Vasc Biol* 1995; 15: 861-871.

37. Zaman MM, Ikemoto S, Yoshiike N, Date C, Yokoyama T, Tanaka H. Association of apolipoprotein genetic polymorphisms with plasma cholesterol in a Japanese rural population. The Shibata Study. *Arterioscler Thromb Vasc Biol* 1997; 17: 3495-3504.
38. Hubacek JA, Waterworth DM, Poledne R, Pitha J, Skodova Z *et al.* Genetic determination of plasma lipids and insulin in the Czech population. *Clin Biochem* 2001; 34: 113-118.
39. Watts GF, Riches FM, Humphries SE, Talmud PJ, van Bockxmeer FM. Genotypic associations of the hepatic secretion of VLDL apolipoprotein B-100 in obesity. *J Lipid Res* 2000; 41: 481-488.
40. Couture P, Otvos JD, Cupples LA, Wilson PW, Schaefer EJ, Ordovas JM. Absence of association between genetic variation in the promoter of the microsomal triglyceride transfer protein gene and plasma lipoproteins in the Framingham Offspring Study. *Atherosclerosis* 2000; 148: 337-343.
41. Ledmyr H, Karpe F, Lundahl B, McKinnon M, Skoglund-Andersson C, Ehrenborg E. Variants of the microsomal triglyceride transfer protein gene are associated with plasma cholesterol levels and body mass index. *J Lipid Res* 2002; 43: 51-58.
42. Herrmann SM, Poirier O, Nicaud V, Evans A, Ruidavets JB *et al.* Identification of two polymorphisms in the promoter of the microsomal triglyceride transfer protein (MTP) gene: lack of association with lipoprotein profiles. *J Lipid Res* 1998; 39: 2432-2435.
43. Erkkila AT, Lindi V, Lehto S, Pyorala K, Laakso M, Uusitupa MI. Variation in the fatty acid binding protein 2 gene is not associated with markers of metabolic syndrome in patients with coronary heart disease. *Nutr Metab Cardiovasc Dis* 2002; 12: 53-59.
44. Pati U, Pati N. Lipoprotein(a), atherosclerosis, and apolipoprotein(a) gene polymorphism. *Mol Genet Metab* 2000; 71: 87-92.
45. Boright AP, Connelly PW, Brunt JH, Morgan K, Hegele RA. Association and linkage of LDLR gene variation with variation in plasma low density lipoprotein cholesterol. *J Hum Genet* 1998; 43: 153-159.
46. Boright AP, Connelly PW, Brunt JH, Scherer SW, Tsui LC, Hegele RA. Genetic variation in paraoxonase-1 and paraoxonase-2 is associated with variation in plasma lipoproteins in Alberta Hutterites. *Atherosclerosis* 1998; 139: 131-136.
47. Knoblauch H, Bauerfeind A, Krahenbuhl C, Daury A, Rohde K *et al.* Common haplotypes in five genes influence genetic variance of LDL and HDL cholesterol in the general population. *Hum Mol Genet* 2002; 11: 1477-1485.
48. Mattu RK, Needham EW, Morgan R, Rees A, Hackshaw AK *et al.* DNA variants at the LPL gene locus associate with angiographically defined severity of atherosclerosis and serum lipoprotein levels in a Welsh population. *Arterioscler Thromb* 1994; 14: 1090-1097.
49. Heizmann C, Kirchgessner T, Kwiterovich PO, Ladias JA, Derby C *et al.* DNA polymorphism haplotypes of the human lipoprotein lipase gene: possible association with high density lipoprotein levels. *Hum Genet* 1991; 86: 578-584.
50. Reymer PW, Gagne E, Groenemeyer BE, Zhang H, Forsyth I *et al.* A lipoprotein lipase mutation (Asn291Ser) is associated with reduced HDL cholesterol levels in premature atherosclerosis. *Nat Genet* 1995; 10: 28-34.
51. Chamberlain JC, Thorn JA, Oka K, Galton DJ, Stocks J. DNA polymorphisms at the lipoprotein lipase gene: associations in normal and hypertriglyceridaemic subjects. *Atherosclerosis* 1989; 79: 85-91.
52. Thorn JA, Chamberlain JC, Alcolado JC, Oka K, Chan L *et al.* Lipoprotein and hepatic lipase gene variants in coronary atherosclerosis. *Atherosclerosis* 1990; 85: 55-60.
53. Peacock RE, Hamsten A, Nilsson-Ehle P, Humphries SE. Associations between lipoprotein lipase gene polymorphisms and plasma correlations of lipids, lipoproteins and lipase activities in young myocardial infarction survivors and age-matched healthy individuals from Sweden. *Atherosclerosis* 1992; 97: 171-185.
54. Talmud PJ, Berglund L, Hawe EM, Waterworth DM, Isasi CR *et al.* Age-related effects of genetic variation on lipid levels: The Columbia University BioMarkers Study. *Pediatrics* 2001; 108: E50.
55. Couture P, Otvos JD, Cupples LA, Wilson PW, Schaefer EJ, Ordovas JM. Association of the A-204C polymorphism in the cholesterol 7alpha-hydroxylase gene with variations in plasma low density lipoprotein cholesterol levels in the Framingham Offspring Study. *J Lipid Res* 1999; 40: 1883-1889.
56. Hegele RA, Wang J, Harris SB, Brunt JH, Young TK *et al.* Variable association between genetic variation in the CYP7 gene promoter and plasma lipoproteins in three Canadian populations. *Atherosclerosis* 2001; 154: 579-587.
57. Helbecque N, Dallongeville J, Codron V, Arveiler D, Ruidavets JB *et al.* The role of a triplet repeat sequence of the very low density lipoprotein receptor gene in plasma lipid and lipoprotein level variability in humans. *Arterioscler Thromb Vasc Biol* 1997; 17: 2759-2764.
58. Saha N, Tay JS, Low PS, Basair J, Hong S. Five restriction fragment length polymorphisms of the APOA1-C3 gene and their influence on lipids and apolipoproteins in healthy Chinese. *Hum Hered* 1995; 45: 303-310.
59. Petrovic D, Zorc M, Peterlin B. Effect of apolipoprotein E polymorphism and apolipoprotein A-1 gene promoter polymorphism on lipid parameters and premature coronary artery disease. 2000; 46: 181-185.
60. Kamboh MI, Bunker CH, Aston CE, Nestlerode CS, McAllister AE, Ukoli FA. Genetic association of five apolipoprotein polymorphisms with serum lipoprotein-lipid levels in African blacks. *Genet Epidemiol* 1999; 16: 205-222.

61. Russo GT, Meigs JB, Cupples LA, Demissie S, Otvos JD *et al.* Association of the Sst-I polymorphism at the APOC3 gene locus with variations in lipid levels, lipoprotein subclass profiles and coronary heart disease risk: the Framingham offspring study. *Atherosclerosis* 2001; 158: 173-181.
62. Acton S, Osgood D, Donoghue M, Corella D, Pocovi M *et al.* Association of polymorphisms at the SR-BI gene locus with plasma lipid levels and body mass index in a white population. *Arterioscler Thromb Vasc Biol* 1999; 19: 1734-1743.
63. Hegele RA, Breckenridge WC, Brunt JH, Connelly PW. Genetic variation in factor VII associated with variation in plasma lipoprotein(a) concentration. *Arterioscler Thromb Vasc Biol* 1997; 17: 1701-1706.
64. Lane A, Green F, Scarabin PY, Nicaud V, Bara L *et al.* Factor VII Arg/Gln353 polymorphism determines factor VII coagulant activity in patients with myocardial infarction (MI) and control subjects in Belfast and in France but is not a strong indicator of MI risk in the ECTIM study. *Atherosclerosis* 1996; 119: 119-127.
65. Zambon A, Deeb SS, Hokanson JE, Brown BG, Brunzell JD. Common variants in the promoter of the hepatic lipase gene are associated with lower levels of hepatic lipase activity, buoyant LDL, and higher HDL2 cholesterol. *Arterioscler Thromb Vasc Biol* 1998; 18: 1723-1729.
66. Guerra R, Wang J, Grundy SM, Cohen JC. A hepatic lipase (LIPC) allele associated with high plasma concentrations of high density lipoprotein cholesterol. *Proc Natl Acad Sci USA* 1997; 94: 4532-4537.
67. Haviland MB, Ferrell RE, Sing CF. Association between common alleles of the low-density lipoprotein receptor gene region and interindividual variation in plasma lipid and apolipoprotein levels in a population-based sample from Rochester, Minnesota. *Hum Genet* 1997; 99: 108-114.
68. Klausen IC, Hansen PS, Gerdes LU, Rudiger N, Gregersen N, Faergeman O. A PvuII polymorphism of the low density lipoprotein receptor gene is not associated with plasma concentrations of low density lipoproteins including LP(a). *Hum Genet* 1993; 91: 193-195.
69. Ahn YI, Kamboh MI, Aston CE, Ferrell RE, Hamman RF. Role of common genetic polymorphisms in the LDL receptor gene in affecting plasma cholesterol levels in the general population. *Arterioscler Thromb* 1994; 14: 663-670.
70. Humphries S, Coviello DA, Masturzo P, Balestreri R, Orecchini G, Bertolini S. Variation in the low density lipoprotein receptor gene is associated with differences in plasma low density lipoprotein cholesterol levels in young and old normal individuals from Italy. *Arterioscler Thromb* 1991; 11: 509-516.
71. Hong SH, Kang BY, Oh JH, Kim JQ, Lee CC. Genetic variations of the apo E-CI-CII cluster gene in Koreans. *Clin Biochem* 1997; 30: 215-219.
72. Elbein SC, Hasstedt SJ. Quantitative trait linkage analysis of lipid-related traits in familial type 2 diabetes: evidence for linkage of triglyceride levels to chromosome 19q. *Diabetes* 2002; 51: 528-535.
73. Al Kateb H, Bahring S, Hoffmann K, Strauch K, Busjahn A *et al.* Mutation in the ARH gene and a chromosome 13q locus influence cholesterol levels in a new form of digenic-recessive familial hypercholesterolemia. *Circ Res* 2002; 90: 951-958.
74. Reed DR, Nanthakumar E, North M, Bell C, Price RA. A genome-wide scan suggests a locus on chromosome 1q21-q23 contributes to normal variation in plasma cholesterol concentration. *J Mol Med* 2001; 79: 262-269.
75. Hunt SC, Hopkins PN, Bulka K, McDermott MT, Thorne TL *et al.* Genetic localization to chromosome 1p32 of the third locus for familial hypercholesterolemia in a Utah kindred. *Arterioscler Thromb Vasc Biol* 2000; 20: 1089-1093.
76. Pajukanta P, Nuotio I, Terwilliger JD, Porkka KV, Ylitalo K *et al.* Linkage of familial combined hyperlipidaemia to chromosome 1q21-q23. *Nat Genet* 1998; 18: 369-373.
77. Coon H, Myers RH, Borecki IB, Arnett DK, Hunt SC *et al.* Replication of linkage of familial combined hyperlipidemia to chromosome 1q with additional heterogeneous effect of apolipoprotein A-I/C-III/A-IV locus. The NHLBI Family Heart Study. *Arterioscler Thromb Vasc Biol* 2000; 20: 2275-2280.
78. Coon H, Leppert MF, Eckfeldt JH, Oberman A, Myers RH *et al.* Genome-wide linkage analysis of lipids in the Hypertension Genetic Epidemiology Network (HyperGEN) Blood Pressure Study. *Arterioscler Thromb Vasc Biol* 2001; 21: 1969-1976.
79. Soro A, Pajukanta P, Lilja HE, Ylitalo K, Hiekkalinna T *et al.* Genome scans provide evidence for low-HDL-C loci on chromosomes 8q23, 16q24.1-24.2, and 20q13.11 in Finnish families. *Am J Hum Genet* 2002; 70: 1333-1340.
80. Pajukanta P, Allayee H, Krass KL, Kuraishy A, Soro A *et al.* Combined analysis of genome scans of dutch and finnish families reveals a susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. *Am J Hum Genet* 2003; 72: 903-917.
81. Imperatore G, Knowler WC, Pettitt DJ, Kobes S, Fuller JH *et al.* A locus influencing total serum cholesterol on chromosome 19p: results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arterioscler Thromb Vasc Biol* 2000; 20: 2651-2656.
82. Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F *et al.* A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet* 2001; 10: 2751-2765.
83. Ober C, Abney M, McPeek MS. The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 2001; 69: 1068-1079.

84. Almasy L, Hixson JE, Rainwater DL, Cole S, Williams JT *et al.* Human pedigree-based quantitative-trait-locus mapping: localization of two genes influencing HDL-cholesterol metabolism. *Am J Hum Genet* 1999; 64: 1686-1693.
85. Shearman AM, Ordovas JM, Cupples LA, Schaefer EJ, Harmon MD *et al.* Evidence for a gene influencing the TG/HDL-C ratio on chromosome 7q32.3-qter: a genome-wide scan in the Framingham study. *Hum Mol Genet* 2000; 9: 1315-1320.
86. Rainwater DL, Almasy L, Blangero J, Cole SA, VandeBerg JL *et al.* A Genome Search Identifies Major Quantitative Trait Loci on Human Chromosomes 3 and 4 That Influence Cholesterol Concentrations in Small LDL Particles. *Arterioscler Thromb Vasc Biol* 1999; 19: 777-783.
87. Klos KL, Kardia SL, Ferrell RE, Turner ST, Boerwinkle E, Sing CF. Genome-wide linkage analysis reveals evidence of multiple regions that influence variation in plasma lipid and apolipoprotein levels associated with risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2001; 21: 971-978.
88. Coon H, Eckfeldt JH, Leppert MF, Myers RH, Arnett DK *et al.* A genome-wide screen reveals evidence for a locus on chromosome 11 influencing variation in LDL cholesterol in the NHLBI Family Heart Study. *Hum Genet* 2002; 111: 263-269.
89. Peacock JM, Arnett DK, Atwood LD, Myers RH, Coon H *et al.* Genome scan for quantitative trait loci linked to high-density lipoprotein cholesterol: The NHLBI Family Heart Study. *Arterioscler Thromb Vasc Biol* 2001; 21: 1823-1828.
90. Arya R, Duggirala R, Almasy L, Rainwater DL, Mahaney MC *et al.* Linkage of high-density lipoprotein-cholesterol concentrations to a locus on chromosome 9p in Mexican Americans. *Nat Genet* 2002; 30: 102-105.
91. Magre J, Delepine M, Khallouf E, Gedde-Dahl T, Jr., Van Maldergem L *et al.* Identification of the gene altered in Berardinelli-Seip congenital lipodystrophy on chromosome 11q13. *Nat Genet* 2001; 28: 365-370.
92. Aouizerat BE, Allayee H, Cantor RM, Davis RC, Lanning CD *et al.* A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J Hum Genet* 1999; 65: 397-412.
93. Kort EN, Ballinger DG, Ding W, Hunt SC, Bowen BR *et al.* Evidence of linkage of familial hypoalphalipoproteinemia to a novel locus on chromosome 11q23. *Am J Hum Genet* 2000; 66: 1845-1856.
94. Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL *et al.* A major susceptibility locus influencing plasma triglyceride concentrations is located on chromosome 15q in Mexican Americans. *Am J Hum Genet* 2000; 66: 1237-1245.

2 | Heritabilities of apolipoprotein and lipid levels in three countries

Marian Beekman, Bastiaan T. Heijmans, Nicholas G. Martin, Nancy L. Pedersen, John B. Whitfield, Ulf DeFaire, G. Caroline M. van Baal, Harold Snieder, George P. Vogler, P. Eline Slagboom, Dorret I. Boomsma.

Twin Research 2002; 5: 87-97

Abstract

This study investigated the influence of genes and environment on the variation of apolipoprotein and lipid levels, which are important intermediate phenotypes in the pathways toward cardiovascular disease. Heritability estimates are presented including for apolipoprotein E level and A2 levels, which has rarely been reported before. We studied twin samples from the Netherlands (two cohorts; n=160 pairs, aged 13-22 and n=204 pairs, aged 34-59), Australia (n=1362 pairs, aged 31-80) and Sweden (n=302 pairs, aged 42-81) which have large sample sizes, especially the Australian sample. The variation of apolipoprotein and lipid levels depended largely on the influences of additive genetic factors in each twin sample. There was no significant evidence for the influence of common environment. No sex differences in heritability estimates for any phenotype in any of the samples were observed. Heritabilities ranged from 0.48 to 0.87, with most heritabilities exceeding 0.60. The heritability estimates in the Dutch samples were significantly higher than in the Australian sample. The heritabilities for the Swedish were intermediate to the Dutch and the Australian samples and not significantly different from the heritabilities in these other two samples. Although sample specific effects are present, we have shown that genes play a major role in determining the variance of apolipoprotein and lipid levels in four independent twin samples from three different countries.

Introduction

Apolipoprotein and lipid profiles are important determinants of cardiovascular risk[1]. Knowledge of the environmental and genetic influences on these levels is relevant for an understanding of the process of cardiovascular disease. Genes described to date that promote atherosclerosis have mainly been identified in familial syndromes of lipid disorders, many of which are monogenic diseases. For example, specific defects in the genes encoding the low-density-lipoprotein-receptor and apolipoprotein B are found to be responsible for the most common forms of familial hypercholesterolemia[2-4]. Furthermore, mutations in the ABC1 gene, encoding the cholesterol efflux regulatory protein, induce familial HDL cholesterol deficiency[5]. However, while these specific mutations have a profound implication for persons with monogenic familial syndromes, they explain only a minor proportion of the population variation in serum apolipoprotein and lipid levels[6]. Hence, it is relevant to estimate the importance of genetic influences in the intermediate phenotypes in the general population. Given significant genetic influences, efforts can be undertaken to identify the genes influencing the variation of these intermediate phenotypes of cardiovascular disease in the population.

In general, the heritability, defined as the proportion of the population variation attributable to genetic variation, of total cholesterol, LDL, HDL and triglyceride levels is larger than 0.50. Genetic factors also explained more than 50% of the total variance in apolipoprotein A1 and B. No obvious age trend in heritability estimates could be detected and none of previously reported studies found much support for a considerable influence of common environment on the variance of the apolipoprotein and lipid levels[7,8]. Although a few studies report heritability estimates on apolipoprotein E and A2 levels[9-11], many of previously reported heritability studies mainly focus on lipid levels and lack heritability estimates on apolipoprotein E and A2 levels.

In this paper we report the variance components attributable to intermediate phenotypes of cardiovascular disease, including apolipoprotein E and A2 levels, compared in four twin samples with different mean age and geographical origin. We studied a adolescent Dutch sample (aged 13-22) of 160 twin pairs[12] and an adult Dutch sample (aged 34-59) of 204 twin pairs[13]. A Swedish twin sample (aged 42-81) provided data on 146 twin pairs reared apart and 156 twin pairs reared together[14]. Heritability estimates for some lipid and apolipoprotein levels have been reported before for the Dutch and Swedish samples. This paper includes for the first time on the heritability of apolipoprotein E level in the adult Dutch sample. The largest sample we studied was an Australian twin sample (aged 31-80) of 1362 twin pairs with data on apolipoprotein and lipid levels, including apolipoprotein E and A2 level. Heritabilities on apolipoprotein and lipid levels in these Australian twins has never been reported previously. Each complete sample was investigated independently using the same model in variance components analyses, testing whether genetics, common environment, unique environment and sex have influences on the variation of apolipoprotein and lipid levels.

Subjects and methods

Subjects

Since different subject and sampling methods have been used in the Netherlands, Australia and Sweden all within different time and age ranges, the subjects and methods used will be categorised by twin sample.

Adolescent Dutch twin pair sample

The adolescent Dutch twin sample is part of a larger study in which cardiovascular risk factors were determined in adolescent twin pairs and their parents[12,15]. The data reported here were collected between 1988 and 1992 in a subgroup of 160 pairs of twins between 13 and 22 years of age. Addresses of twins living in Amsterdam and neighbouring cities were obtained from City Council population registries. Twins still living with both their biological parents were contacted by letter. A family was included in the study if the twin and both parents were willing to comply. In addition, a small number of families who heard of the study from other twins also volunteered to participate. Three triplets were included by discarding the data of the second-born subject. None of the twin pairs reported taking lipid lowering medication.

Zygosity was determined by typing 11 blood group polymorphisms. Zygosity in the dizygotic pairs was confirmed using 103 microsatellite markers on 10 chromosomes typed as part of a genome wide search comprising 229 markers. In total, there were 35 monozygotic male twin pairs (MZM), 31 dizygotic male twin pairs (DZM), 35 monozygotic female twin pairs (MZF), 30 dizygotic female twin pairs (DZF) and 29 dizygotic twin pairs of opposite sex (DOS).

Adult Dutch twin pair sample

The adult Dutch twin sample is also part of the larger study[16]. The data reported here were collected between 1992 and 1996 in a group of 213 pairs of twins between 34 and 62 years of age. Twins were recruited by a variety of means, including advertisement in the media, advertisement in the information bulletin of the Dutch Twin Registry and solicitation through the Dutch Twin Club. In addition, a small number of twins who heard of the study from other twins volunteered to participate. One triplet was included by discarding the data of the second-born subject. Data of 9 twin pairs were excluded from the sample. In 8 of these twins one or both members of the pair used lipid lowering medication (HMG-CoA reductase inhibitors) and in one subject no blood could be obtained.

Zygosity was decided on the basis of the response to standard questions about physical similarity and the degree to which others confused them. In 76 same-sex twin pairs zygosity was also determined by DNA fingerprinting and in 98.7% of those twin pairs the zygosity was correctly classified. Zygosity of the dizygotic twin pairs was confirmed using 103 microsatellite markers. In total, there were 41 pairs of MZM, 37 pairs of DZM, 50 pairs of MZF, 40 pairs of DZF and 40 pairs of DOS.

Australian twin pair sample

The Australian twin sample is part of the Semi-Structured Assessment for the Genetics of Alcoholism study[17]. The data reported here were collected between 1993 and 1996 in a subgroup of 1403 pairs of twins between 28 and 92 years of age. Twins were recruited through the Australian NHMRC Twin Registry. Data of 41 twin pairs were excluded from the sample, since one or both members of the pair used lipid lowering medication.

Zygosity was decided on the basis of the response to standard questions about physical similarity and the degree to which others confused them. Pairs giving inconsistent responses were contacted for clarification. In 329 same-sex twin pairs zygosity was also confirmed by typing 11 highly polymorphic markers and in 98.5% of those twin pairs the zygosity was correctly classified. Of the pairs who were classified as dizygotic, zygosity of 263 pairs was confirmed using 103 microsatellite markers. In total, there were 194 pairs MZM, 107 pairs of DZM, 517 pairs of MZF, 272 pairs of DZF and 272 pairs of DOS.

Swedish twin pair sample

The Swedish twin sample is part of the larger Swedish Adoption/Twin Study of Aging sample[18,19]. The data reported here were collected between 1986 and 1988 in a subgroup of 302 pairs of twins with the same sex between 42 and 88 years of age. There are 146 twin pairs reared apart and 156 twin pairs reared together included in this study. All the older adults in this study reported not to take lipid lowering medication.

Zygosity was determined by typing serologic markers[18]. Of the dizygotic twins reared together, zygosity of 44 twin pairs was confirmed using 103 microsatellite markers. In total, there were 22 pairs of MZM reared apart, 31 pairs of DZM reared apart, 24 pairs of MZF reared apart and 69 pairs of DZF reared apart. In addition, there were 27 pairs of MZM reared together, 38 pairs of DZM reared together, 42 pairs of MZF reared together and 49 pairs of DZF reared together.

Methods

Adolescent Dutch twin sample

The subjects were requested to fast from 23:00 PM the preceding night. Blood was taken between 08:30 AM and 10:30 AM by venipuncture using Vacutainer tubes (Becton-Dickinson) containing sodium-EDTA. The tubes were placed on ice and centrifuged promptly (10 min, 3,000 rmp) at 4°C to separate plasma from the cells. Part of plasma was kept on 4°C for lipid determination within the next 5 days. The remainder was frozen using liquid nitrogen and stored at -20°C until processing.

Apolipoprotein E was quantified by enzyme-linked immunosorbent assays (ELISA)[20]. Apolipoprotein A2, A1 and B were quantified by radial immunodiffusion as described by Albers *et al.*[21] and Havekes *et al.*[22].

Cholesterol and triglycerides were measured using enzymatic methods (Boehringer Mannheim, FRG, CHOD_PAP kit number 236691 and GPO-PAP kit number 701904). High density lipoprotein (HDL) was measured after lipoproteins containing apolipoprotein B were precipitated with phosphotungstate-magnesium chloride[23].

Adult Dutch twin sample

The subjects were requested to fast from 23:00 PM the preceding night. Blood was taken at about 10:00 AM by venipuncture using Vacutainer tubes (Becton-Dickinson) containing sodium-EDTA. The tubes were placed on ice and centrifuged promptly (30 min, 2,000 g) at 4°C to separate plasma from the cells. Aliquots of plasma were frozen using liquid nitrogen and stored at -20°C until processing.

Apolipoprotein E was quantified by enzyme-linked immunosorbent assays (ELISA)[24]. Apolipoprotein A2 plasma levels were quantified by radial immunodiffusion as described by Albers *et al.*[21] and Zonderland *et al.*[25].

The MZ twin correlation coefficients for apoA2 level were about 0.25 as compared to >0.60 for all other phenotypes in this adult Dutch twin sample. This suggests a measurement problem and therefore, the apolipoprotein A2 levels in the adult Dutch twins were not analysed.

Apolipoprotein A1 and B were quantified with the method of Beckman using the Array Protein System (Beckman Instruments)[26]. The Beckman calibrator (standardised to the International Federation for Clinical Chemistry, which is traceable to the World Health Organisation International Reference Material for ApoA1 and ApoB no. 1883) was used as standard reference material. Monospecific goat-antihuman ApoA1 and ApoB antibodies were used (Beckman).

Total cholesterol, triglyceride and HDL cholesterol levels were determined in the same way as for the adolescent Dutch twin sample.

Australian twin sample

Blood was taken from the subjects throughout the day, depending on their availability. The time since the last meal was recorded. Aliquots of plasma were frozen in liquid nitrogen and stored at -70°C until analysis.

Plasma levels of apolipoprotein E, A2, A1, B and E were quantified on a Behring nephelometer with Behring reagents. Total cholesterol and triglyceride levels were measured on a Hitachi 747 analyser with Boehringer reagents by standard enzymatic methods[27]. HDL cholesterol was measured after lipoproteins containing apolipoprotein B were precipitated with dextran sulphate magnesium chloride[28].

Swedish twin sample

The subjects were requested to fast for 12 hours before blood was taken. Blood samples were frozen at -70°C, transported in dry ice, and thawed and analysed when received by the laboratory. Apolipoprotein A1 and B levels were quantified using commercial radioimmunoassay kits (RIA 100, Pharmacia Diagnostics, Uppsala, Sweden). Total cholesterol and triglyceride levels were measured with an enzymatic colorimetric assay (Boehringer Mannheim automated analysis for Hitachi systems 717, Diagnostica, Mannheim, Germany). HDL cholesterol was measured after lipoproteins containing apolipoprotein B were precipitated with phosphotungstate-magnesium chloride[29].

In all samples, concentrations of low density lipoprotein (LDL) cholesterol were estimated according to Friedewald[30], when triglyceride concentrations did not exceed 4.52 mmol/L[31]. LDL cholesterol level was considered missing, when the triglyceride concentration did exceed 4.52 mmol/L, which was the case for 1 individual of the adult Dutch sample, for 82 individuals of the Australian sample and for 11 individuals of the Swedish sample.

Statistical Analyses

To study the contribution of genetic and environmental factors to apolipoprotein and lipid levels variability, a structural modelling approach was used. Univariate models were fitted to the data by the method of maximum likelihood. A univariate model was tested in which the phenotypic variance is divided into an additive genetic, a common environment (including family environment) and a unique environment component. The additive genetic component reflects the heritability of a trait and the common environment reflects the shared-family environment. Since the Swedish twin sample consists of twins reared together and apart, an additional correlated environment component, which may reflect for example prenatal influence or post-rearing contact, was initially present in the Swedish model. However, correlated environment did not significantly influence the total variance of any of the phenotypes and is therefore not shown in this paper. Model fitting was carried out directly on the raw data. This allows us to model the covariance structure between family members simultaneously with any covariate effects on the means. In all analyses, means were adjusted for sex and age differences between pairs.

To study sex differences in genetic influences, a model with different parameter estimates for men and women for additive genetic and both environmental components was fitted to the data first, followed by a model in which parameter estimates for the various variance components were constrained to be equal across sexes. We also fitted a sex limitation scalar model, in which

the heritabilities are constrained to be equal across sexes, but total variances are allowed to be different; variance components for males are constrained to be equal to a scalar multiple of the female variance components. As a result, the standardised variance components such as heritability are equal across sexes, even though the non-standardised components differ[32].
The correlation between the additive genetic factors in dizygotic twins of opposite sex was free to be estimated between 0 and 0.5. When this correlation is estimated significantly smaller than 0.5, different genes in men and women influence the same phenotype, even though the proportion of the variance attributed to genetic factors may be the same.
By the principle of parsimony, the pattern of variances and covariances should be explained by as few parameters as possible. Therefore, the most extensive models were reduced by excluding respectively the genetic or common environmental component from the model for men and women separately. When the best fitting model was the same for men and women, common parameters were equated between men and women. Sub-models were compared to the most extensive model by hierarchic χ^2 tests. The difference in -2ln(likelihood) of the goodness-of-fit of each model is approximated by a chi-square distribution, with degrees of freedom equal to the difference in number of parameters estimated in the two models. A significant increase in the -2ln(likelihood) after a parameter has been excluded from the model indicates that the reduced model fits the data less well than the most extensive model.
Because the distribution of the values for triglyceride level was skewed in all samples, these values were transformed by natural logarithm. All samples were analysed separately. Model fitting was performed using Mx version 1.5[33]. Confidence intervals around standardised variance components (e.g. heritabilities) were also obtained from Mx.

Results

Descriptives

Table 1 shows means and standard deviations for levels of apolipoprotein E and A2, levels of other apolipoproteins and lipids in each sample for men and women separately. In all samples women had higher HDL cholesterol and apolipoprotein A1 levels than men. Adult Dutch and Australian men had higher mean values than adult Dutch and Australian women for LDL cholesterol, apolipoprotein B and triglyceride level. In adolescent Dutch and older Swedish subjects this sex difference seemed to be absent.

Twin correlation coefficients

Twin correlation coefficients for the traits for each twin sample are shown by zygosity groups for men and women in table 2a-d. Correlation coefficients for apolipoprotein E levels were available in the Dutch adolescent and adult twins and in the Australian sample. In monozygotic (MZ) twins the correlations for apolipoprotein E levels were higher than in dizygotic (DZ) twins, which is a strong indication that genetic factors influence these levels. Correlation coefficients for apolipoprotein A2 levels were only available in the adolescent Dutch and Australian twin sample, and these correlations also suggested the influence of genetic factors. This was also the case for the other apolipoprotein and lipid levels.
In the adolescent Dutch twin pairs (Table 2a), the correlation coefficients for total cholesterol and LDL cholesterol in DZ men, but not in women, are relatively low. However, it is unlikely that this low correlation reflects an error in the LDL cholesterol measurements of the adolescent

Tabel 1.
Descriptives of the twin samples. Means and standard deviations between brackets per sample and sex for levels of apolipoprotein E, A2, A1 and B and for levels of total, low-density-lipoprotein, high-density-lipoprotein cholesterol and the natural log of triglyceride levels.

Twin sample	Adolescent Dutch twins		Adult Dutch twins		Australian twins		Swedish twins	
Mean (SD)	*Men*	*Women*	*Men*	*Women*	*Men*	*Women*	*Men*	*Women*
Number of individuals	161	159	196	220	874	1850	234	350
Age (yr)	16.77 **(1.78)**	16.71 **(2.20)**	43.55 **(6.47)**	44.70 **(6.79)**	44.10 **(10.41)**	45.88 **(11.69)**	65.02 **(7.51)**	66.09 **(8.99)**
Apolipoprotein E (mg/dL)	6.03 **(2.06)**	7.58 **(2.42)**	2.50 **(1.08)**	2.57 **(0.97)**	4.28 **(1.89)**	3.87 **(1.62)**	-	-
Apolipoprotein A2 (g/L)	0.58 **(0.07)**	0.58 **(0.08)**	-	-	0.35 **(0.06)**	0.36 **(0.07)**	-	-
Apolipoprotein A1 (g/L)	1.34 **(0.15)**	1.44 **(0.23)**	1.50 **(0.35)**	1.93 **(0.38)**	1.32 **(0.22)**	1.51 **(0.29)**	1.27 **(0.24)**	1.48 **(0.28)**
Apolipoprotein B (g/L)	0.78 **(0.15)**	0.80 **(0.19)**	1.27 **(0.34)**	1.19 **(0.34)**	1.05 **(0.25)**	0.94 **(0.25)**	1.07 **(0.21)**	1.12 **(0.26)**
Total cholesterol (mmol/L)	4.05 **(0.65)**	4.47 **(0.86)**	5.39 **(1.03)**	5.46 **(1.04)**	5.71 **(1.05)**	5.60 **(1.08)**	6.35 **(1.10)**	7.20 **(1.46)**
Low density lipoprotein (mmol/L)	2.52 **(0.63)**	2.76 **(0.76)**	3.71 **(0.97)**	3.60 **(0.97)**	3.55 **(0.91)**	3.37 **(0.99)**	4.28 **(1.01)**	4.87 **(1.30)**
High density lipoprotein (mmol/L)	1.23 **(0.22)**	1.38 **(0.29)**	1.07 **(0.29)**	1.39 **(0.33)**	1.23 **(0.29)**	1.56 **(0.39)**	1.32 **(0.36)**	1.60 **(0.44)**
Ln(Triglycerides) (mmol/L)	-0.49 **(0.39)**	-0.41 **(0.36)**	0.16 **(0.52)**	-0.04 **(0.40)**	0.62 **(0.59)**	0.29 **(0.54)**	0.40 **(0.51)**	0.35 **(0.46)**

Table 2a.
Twin correlation coefficients in the adolescent Dutch twin sample.

Twin correlation	Adolescent Dutch twins				
Sex	*Men*		*Women*		*DOS*
Zygosity (number of pairs)	MZ (35)	DZ (31)	MZ (35)	DZ (30)	DOS (29)
ApoE	0.87	0.38	0.86	0.17	0.41
ApoA2	0.90	0.46	0.81	0.60	0.30
ApoA1	0.86	0.34	0.77	0.50	0.53
ApoB	0.83	0.55	0.82	0.60	0.71
Total cholesterol	0.90	0.11	0.81	0.63	0.41
LDL cholesterol	0.89	0.22	0.83	0.56	0.44
HDL cholesterol	0.73	0.48	0.81	0.47	0.57
Triglycerides	0.59	0.59	0.71	0.32	0.17

Table 2b.
Twin correlation coefficients in the adult Dutch twin sample.

Twin correlation	**Adult Dutch twins**				
Sex	*Men*		*Women*		*DOS*
Zygosity (number of pairs)	MZ (41)	DZ (37)	MZ (50)	DZ (40)	DOS (40)
ApoE	0.82	0.48	0.92	0.35	0.46
ApoA2	-	-	-	-	-
ApoA1	0.73	0.56	0.67	0.49	0.47
ApoB	0.80	0.67	0.78	0.30	0.35
Total cholesterol	0.75	0.55	0.79	0.41	0.46
LDL cholesterol	0.77	0.65	0.80	0.27	0.37
HDL cholesterol	0.61	0.34	0.70	0.37	0.17
Triglycerides	0.62	0.47	0.60	0.40	0.43

Table 2c.
Twin correlation coefficients in the Australian twin sample.

Twin correlation	**Australian twins**				
Sex	*Men*		*Women*		*DOS*
Zygosity (number of pairs)	MZ (194)	DZ (107)	MZ (517)	DZ (272)	DOS (272)
ApoE	0.51	0.17	0.65	0.33	0.28
ApoA2	0.54	0.17	0.44	0.24	0.23
ApoA1	0.34	0.22	0.56	0.22	0.18
ApoB	0.61	0.48	0.68	0.39	0.28
Total cholesterol	0.63	0.41	0.60	0.43	0.22
LDL cholesterol	0.61	0.39	0.63	0.44	0.22
HDL cholesterol	0.57	0.41	0.64	0.29	0.22
Triglycerides	0.52	0.40	0.55	0.41	0.25

Table 2d.
Twin correlation coefficients in the Swedish twin sample.

Twin correlation	**Swedish twins**				
Sex	*Men*		*Women*		*DOS*
Zygosity (number of pairs)	MZ (49)	DZ (69)	MZ (66)	DZ (118)	DOS
ApoE	-	-	-	-	-
ApoA2	-	-	-	-	-
ApoA1	0.70	0.15	0.57	0.31	-
ApoB	0.60	0.16	0.77	0.39	-
Total cholesterol	0.54	0.11	0.61	0.41	-
LDL cholesterol	0.53	0.20	0.56	0.42	-
HDL cholesterol	0.71	0.21	0.61	0.32	-
Triglycerides	0.42	0.15	0.63	0.34	-

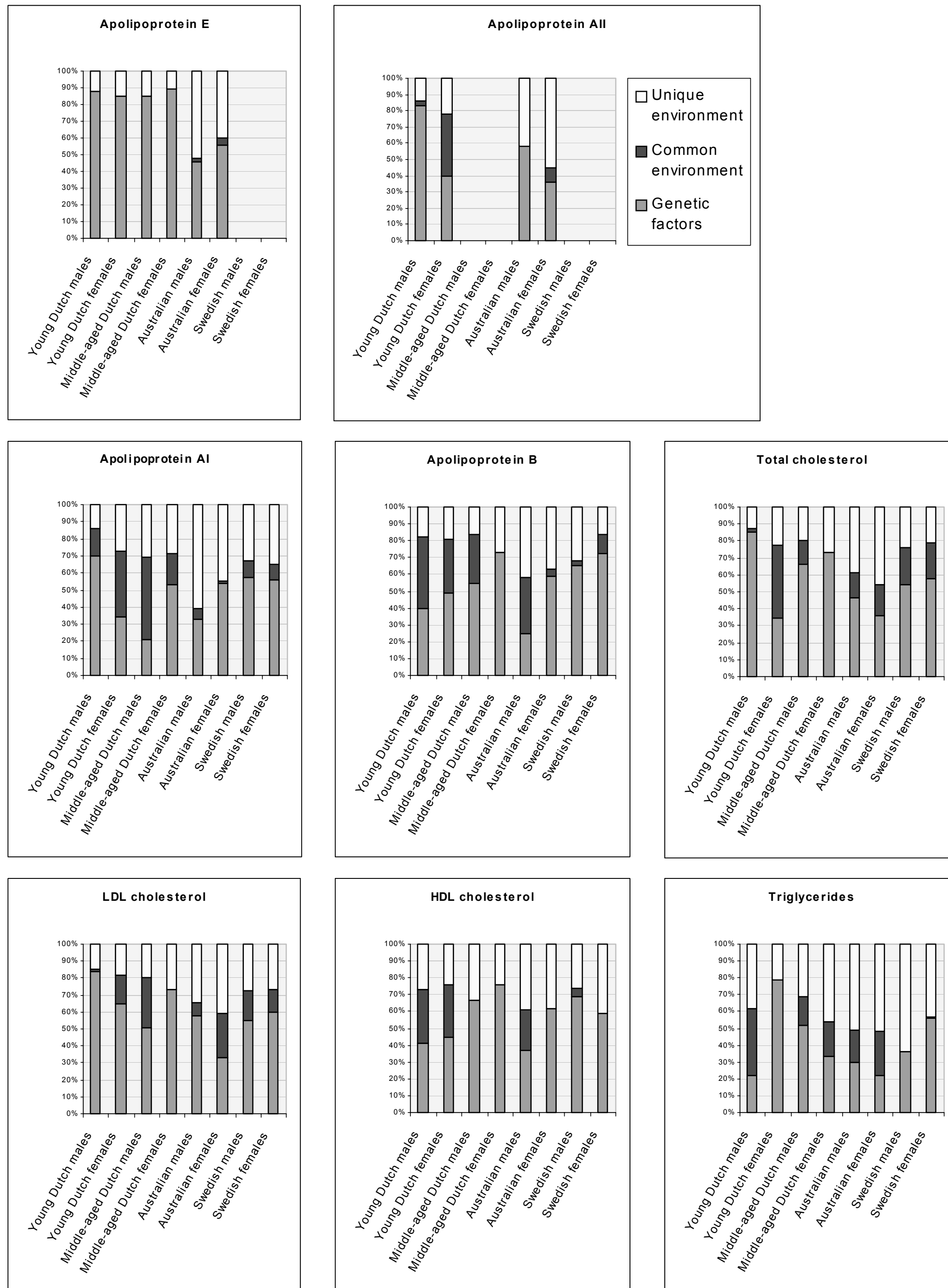
Dutch DZ men, since the phenotypic measurements were performed randomly over the adolescent Dutch sample.
The MZ correlations in the adult Dutch twins (Table 2b) are higher than the DZ correlations for all phenotypes, indicating that genetic factors influence these traits.
In the Australian sample (Table 2c), the correlation coefficients in DZ twins of opposite sex for apolipoprotein B, total cholesterol, LDL, HDL and triglyceride level seem to be lower than the correlation coefficients of DZ twins of the same sex. This may indicate that different genes affect these traits in men and women.
In the Swedish men (Table 2d), the DZ correlations are more than twice as low as the MZ correlations, which could indicate that dominant genetic factors might influence these traits.

Structural modelling

Figure 1 shows the proportion of the variances estimated under the conditions of the most extensive model, including additive genetic, common and unique environmental factors separate components for men and women in each sample. In the Australian sample, genetic factors explain 46-60% of the variance of apolipoprotein E levels and in the Dutch adolescent and adult samples 85-89%. Common environment explains at most 4% of the total variance of apolipoprotein E level.
Under the full model, genetic factors explain 36-40% of the total variance of apolipoprotein A2 level in women and 58-84% in men. Common environment may play a larger role in women than in men. The proportions of the variance explained by common environment are, respectively, 0.10-0.35 and 0.0-0.02, under the full model.
For apolipoprotein A1 and B levels, 21-70% of the variance is determined by genetic influences and 0-48% of the variance can be explained by common environment. Between 21-85% of the variance of lipid levels is determined by genetic factors and 0-43% by common environmental factors.
Overall, genetic factors seem to play a larger role than common environment in explaining inter-individual differences in levels of apolipoproteins and lipids.
After fitting the most extensive model, sub-models were compared to the full model by hierarchic χ^2 tests to find a model to explain the data that is as simple as possible. A model in which the variance was only explained by common and unique environment showed in all analyses a significant worse fit than the full model. A model in which the variance was only influenced by a genetic and a unique environmental factor did not show a significantly worse fit than the full model in any of the analyses. Therefore, common environment does not play a significant role in determining the levels of apolipoproteins and lipids in any of the samples. In table 3 the genetic variance versus the total variance is listed for men and women separately. Also the heritability estimates and the 95% confidence interval for all traits are listed for the most parsimonious models separately in each sample. All heritabilities were equal across sexes in every sample. In some analyses the scalar model was the best fitting model (Table 3, indicated with [b]), in which the proportion of the variance attributable to genetic factors are equal across sexes, even though the non-standardised estimates for the genetic effect differ.
Since all samples are collected at different time points and centres and since the measurements on these samples were all at different laboratories using slightly different methods, the variances differ over the samples. No distinction can be made between these factors of difference. Therefore, the 95% confidence intervals of the heritability give the information necessary for

Figure 1.

Proportions of variance that can be explained by influence of genes, common environment and unique environment. These proportions are estimated using the full model for each sample estimating parameters for males and females separately.



comparisons between the four samples. The heritability estimates for apolipoprotein E level range from 0.57 in the Australian sample to 0.87 in the adult Dutch sample. The heritability estimates in the Dutch samples are significantly higher than the estimate in the Australian sample. Also the heritability estimates for apolipoprotein A2 level in Australian sample (0.48) is significantly lower than in the adolescent Dutch sample (0.82).

For apolipoprotein A1 and B levels, the heritability ranges from 0.50 in the Australian twins to 0.82 in adolescent Dutch twins. The heritability estimates in the Swedish sample is intermediate to the estimates in the Dutch and Australian samples, which significantly differ from each other. Common environment seems not to play a significant role in the total variance of apolipoprotein A1 levels in all the samples. However, its role in the variance of apolipoprotein B levels is unclear. In the adolescent Dutch and Australian twins, models including the common environment nearly reached significance (Table 3; indicated with [a]). However, in the adult Dutch and Swedish samples, there is no suggestion that common environment might play a role in the inter-individual differences of apolipoprotein B levels.

The heritability estimates for lipid levels ranged from 0.48 for triglyceride level in the Swedish sample to 0.83 for LDL cholesterol level in adolescent Dutch twins. In the Australian and Swedish samples, common environment might play a role in the variance of total cholesterol levels, since the models including the common environmental factor nearly reached significance. However, in the Dutch samples no suggestion was found for the influence of common environment on the total cholesterol levels. Furthermore, the models including common environment for the Australian LDL cholesterol and triglyceride levels also nearly reached significance, although no indications for common environmental influences were observed in the other samples.

In general, the four twin samples show high heritability estimates for all apolipoprotein and lipid levels without sex differences. In none of the samples, the effect of common environment on the variances was significant. No significant difference has been observed between the heritability estimates of the two Dutch samples. The adolescent Dutch twins show significantly higher estimates for apolipoprotein A1 and LDL cholesterol levels than the Swedish twins do; while in the adult Dutch twins none of the heritability estimates significantly differ from the estimates in the Swedish twins. Roughly, there seems no difference between the Dutch and Swedish twins in heritability estimates, although the Swedish estimates are consistently lower.

The most striking differences in heritability are those between the Dutch and the Australian samples, that would have to be explained by sample specific differences. The heritability estimates in the Dutch twin cohorts are significantly higher than in the Australian twins, except for HDL cholesterol and triglyceride levels.

Analysing the Australian sample for total and LDL cholesterol, models in which the correlations between the additive genetic factors of DZ twins of opposite sex were fixed to 0.5 like in DZ twins of same sex, fitted significantly worse than the models in which these correlations were free to be estimated. The correlation coefficient between the additive genetic factors of DZ twins of opposite sex was then estimated to 0.26 (Table 3, indicated with [c]).

Hence, only half of the genes influencing total and LDL cholesterol levels are acting in both Australian men and women. Therefore, the other half of the genes influencing total and LDL cholesterol levels in the Australian sample is sex specific. However, this was not found in the two Dutch samples. The Dutch samples are much smaller than the Australian sample and therefore

Table 3
Genetic variance and total variance for men and women separately. Heritability estimates and the 95% confidence interval between brackets using the AE model that assumes only genetic and unique environmental influences. [a]: The heritability estimate might also include common environmental influences. [b]: The scalar model showed the best fit. [c]: The correlation between the additive genetic factors in DOS is 0.26 for both total and LDL cholesterol.

Twin sample	Adolescent Dutch twins			Adult Dutch twins			Australian twins			Swedish twins		
	Genetic/Total Variance		heritability (95% C.I.)	Genetic/Total Variance		heritability (95% C.I.)	Genetic/Total Variance		heritability (95% C.I.)	Genetic/Total Variance		heritability (95% C.I.)
	Men	*Women*		*Men*	*Women*		*Men*	*Women*		*Men*	*Women*	
ApoE	4.368/5.079	4.368/5.079	**0.86** **(0.80-0.91)**	0.884/1.016	0.884/1.016	**0.87** **(0.82-0.91)**	2.045/3.588	2.045/3.588	[b]**0.57** **(0.52-0.61)**	-	-	-
ApoA2	0.005/0.006	0.005/0.006	**0.82** **(0.79-0.82)**	-	-	-	0.002/0.003	0.002/0.003	[b]**0.48** **(0.42-0.53)**	-	-	-
ApoA1	0.036/0.045	0.017/0.021	[b]**0.80** **(0.71-0.86)**	0.084/0.117	0.084/0.117	**0.72** **(0.62-0.79)**	0.026/0.051	0.040/0.080	[b]**0.50** **(0.45-0.55)**	0.048/0.079	0.048/0.079	**0.61** **(0.49-0.70)**
ApoB	0.023/0.027	0.023/0.027	[a]**0.82** **(0.75-0.87)**	0.084/0.106	0.084/0.106	**0.79** **(0.71-0.85)**	0.036/0.057	0.036/0.057	[a]**0.63** **(0.58-0.66)**	0.048/0.065	0.048/0.065	**0.75** **(0.66-0.81)**
Total cholesterol	0.325/0.396	0.578/0.704	[b]**0.82** **(0.74-0.87)**	0.740/0.961	0.740/0.961	**0.77** **(0.68-0.83)**	0.608/1.067	0.593/1.040	[a, b, c]**0.57** **(0.52-0.61)**	1.145/1.709	1.145/1.709	[a]**0.67** **(0.56-0.76)**
LDL cholesterol	0.384/0.463	0.384/0.463	**0.83** **(0.76-0.88)**	0.656/0.852	0.656/0.852	**0.77** **(0.69-0.84)**	0.504/0.826	0.563/0.922	[a, b, c]**0.61** **(0.57-0.65)**	1.040/1.576	1.040/1.576	**0.66** **(0.55-0.75)**
HDL cholesterol	0.058/0.077	0.036/0.048	[b]**0.75** **(0.65-0.82)**	0.073/0.101	0.073/0.101	**0.72** **(0.60-0.80)**	0.053/0.085	0.096/0.155	[b]**0.62** **(0.57-0.66)**	0.116/0.181	0.116/0.181	**0.64** **(0.54-0.72)**
Triglycerides	0.090/0.127	0.109/0.153	**0.71** **(0.57-0.80)**	0.096/0.155	0.168/0.271	[b]**0.62** **(0.50-0.72)**	0.168/0.336	0.130/0.259	[a, b]**0.50** **(0.45-0.55)**	0.102/0.213	0.102/0.213	**0.48** **(0.35-0.59)**

have less power to detect sex-specific genetic effects. The DOS correlations in the Dutch samples, however, do not suggest such an effect. Since no twins of opposite sex were present in the Swedish twin pair sample, no conclusions can be drawn from this sample.
When the correlation in DZ twins was less than half of the correlation in MZ twins, models including a dominance genetic factor were examined. However, none of these models fitted significantly better than the model only including an additive genetic and unique environmental factor.

Discussion

In this paper, data from three different countries were analysed to obtain heritabilities for intermediate phenotypes of cardiovascular disease. We found that 48 to 87% of the total variances is attributable to genetic factors. The two Dutch samples showed the highest heritabilities ranging from 0.62 to 0.87. The Swedish heritabilities ranged from 0.48 to 0.75 and the Australian heritabilities ranged from 0.48 to 0.63. The differences in heritability estimates between the Dutch samples and the Australian sample were significant. In none of the samples, we found evidence for the influences of common environment on the variance of any apolipoprotein or lipid level. No sex differences in heritability estimates were observed in any of the twin samples.
The total variance of apolipoprotein E levels depended largely on genetic influences. We have previously reported heritability of apolipoprotein E levels in the adolescent Dutch twin sample[12]. These results are now supplied with heritability estimates in the adult Dutch twins of 0.87 and in the Australian twins of 0.57, which augments the role of genetic factors in determining the level of apolipoprotein E. Also a large part of the variance of apolipoprotein A2 levels can be explained by genetic factors, although there are sample specific differences in the heritability estimates. The heritability estimate in the adolescent Dutch twins was 0.82 and in the Australian twins 0.48. Our findings are in line with the few previously reported studies[9,34,35].
We found high heritabilities for apolipoprotein A1 levels in all samples ranging from 0.50 in the Australian twins to 0.80 in the adolescent Dutch twins. Heritability estimates for levels of apolipoprotein A1 previously gave contradicting results. Some studies found no evidence for additive genetic influences[9,35] and others found heritabilities of comparable magnitude[34] as we report in this paper for the four samples.
The total variance of apolipoprotein B levels in adolescent Dutch, adult Dutch, Australian and Swedish twins is for 63 to 82% attributable to additive genetic influences. This corresponds to previous findings reviewed by Snieder *et al.*[7]. Since the influences of common environment are almost significantly present in adolescent Dutch and adult Australian twins (Table 3, indicated with [a]), it may well be that these influences play a role in apolipoprotein B levels. However, Boomsma *et al.*[12] showed that when the adolescent Dutch twin sample was extended with parental data, the most parsimonious model on the apolipoprotein B levels was the model only including additive genetic and unique environmental factors. And since in the other twin samples no evidence was found for influences of common environment, we conclude that the common environment does not play a large role in the variation of apolipoprotein B levels. Furthermore, the heritability estimates for total cholesterol, LDL, HDL and triglyceride levels ranged from 0.48 to 0.87 without sex differences, which is in correspondence with previously reported heritabilities[7].

In the Australian sample, models for total and LDL cholesterol levels, in which the correlation between the additive genetic factors of DZ twins of opposite sex were fixed to 0.5, fitted significantly worse than the models in which this correlation was free to be estimated. Hence, only half of the genes influencing total and LDL cholesterol levels are acting in both Australian men and women. This could not be confirmed in the two Dutch samples, possibly due to lack of power. Since no twins of opposite sex were present in the Swedish twin pair sample, no conclusions can be made from this sample. From a biological point of view, it seems not very likely that only in the Australian sample different genetic influences act in men and women. It may be, however, that Australian women are exposed to higher levels of hormonal treatment than Dutch women are[36,37] that explains part of the variance in their LDL levels[38] and that part of the response to these hormones is genetically determined[39]. Since no other study reported such genetic sex differences, we are careful to draw a firm conclusion.
From our data can be concluded that none of the lipid traits studied in this paper is significantly influenced by common environment. In literature, contradicting results are found in relation to common environmental factors[40-43]. However, since our twin samples are quite large, especially the Australian sample including 1362 twin pairs, it seems unlikely that we have lack of power in this study to obtain significance.
We have shown that large parts of the variances of intermediate phenotypes of cardiovascular disease can be explained by genetic factors. A number of candidate genes are known to influence some of these traits. For example, it is known that the *APOE*ε2/ε3/ε4 polymorphism influences apolipoprotein E levels[44,45]. Additionally, the T allele of the –219GT *APOE* promoter polymorphism is independently associated with lower levels of apolipoprotein E[46]. However, only approximately 20% of the genetic variance can be explained by known polymorphisms at the *APOE* locus. None of the heritabilities reported can be totally explained by known polymorphisms in candidate genes.
In this study design, the influence of age on heritability is difficult to assess. Each sample was collected in a different time period and the apolipoprotein and lipid measurements were performed in different laboratories. It is impossible to distinguish between age and geographical origin of the samples. Age effects in the Dutch twin samples have been analysed by Snieder et al[13] using the data on parents of adolescent twins in addition to the data on twins themselves. It appeared that at different ages partly different genes influence the variance of levels of triglycerides, total, LDL, and HDL cholesterol. But age had no effect on the heritability of intermediate phenotypes of cardiovascular disease in the Dutch samples.
Assuming that age does not influence the heritabilities leads us to conclude that the significant differences between the Dutch and the Australian samples are caused by sample specific factors. Each sample is drawn from populations which are likely to differ in their genetic as well as environmental makeup. The larger influence of environmental factors in the Australian sample might be a genuine effect, explained either by a larger main effect of environment on lipid levels or by an interaction with genetic factors[39]. Furthermore, the measurements are performed with slightly different methods and in different laboratories, which could result in different measurement errors, which are included in the unique environmental factors. Increases in unique environmental estimates would lead to concomitant decreases in estimates of genetic factors.
In conclusion, we have shown that apolipoprotein E and A2 levels show high heritabilities in four independent samples, comparable with other intermediate phenotypes of cardiovascular disease. Common environment does not influence the variances and there are no sex differences in the

heritabilities of these intermediate phenotypes. Given the high heritabilities observed, these four twin cohorts would provide a unique sample for QTL mapping of the genes involved.

Acknowledgements

This study was supported by a grant from the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976).

References

1. Rifai N, Chapman JF, Silverman LM, Gwynnes JT. Review of serum lipids and apolipoproteins in risk-assessment of coronary heart disease. *Ann Clin Lab Sci* 1988; 18: 429-439.
2. Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat* 1992; 1: 445-466.
3. Ludwig EH, Hopkins PN, Allen A, Wu LL, Williams RR *et al.* Association of genetic variations in apolipoprotein B with hypercholesterolemia, coronary artery disease, and receptor binding of low density lipoproteins. *J Lipid Res* 1997; 38: 1361-1373.
4. Hansen PS. Familial defective apolipoprotein B-100. *Dan Med Bull* 1998; 45: 370-382.
5. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K *et al.* Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 1999; 22: 336-345.
6. Goldstein JL, Schrott HG, Hazzard WR, Bierman EL, Motulsky AG. Hyperlipidemia in coronary heart disease. II. Genetic analysis of lipid levels in 176 families and delineation of a new inherited disorder, combined hyperlipidemia. *J Clin Invest* 1973; 52: 1544-1568.
7. Snieder H, van Doornen LJ, Boomsma DI. Dissecting the genetic architecture of lipids, lipoproteins, and apolipoproteins: lessons from twin studies. *Arterioscler Thromb Vasc Biol* 1999; 19: 2826-2834.
8. Hamsten A, Iselius L, Dahlen G, de Faire U. Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. *Atherosclerosis* 1986; 60: 199-208.
9. Hasstedt SJ, Albers JJ, Cheung MC, Jorde LB, Wilson DE *et al.* The inheritance of high density lipoprotein cholesterol and apolipoproteins A-I and A-II. *Atherosclerosis* 1984; 51: 21-29.
10. Kuusi T, Kesaniemi YA, Vuoristo M, Miettinen TA, Koskenvuo M. Inheritance of high density lipoprotein and lipoprotein lipase and hepatic lipase activity. *Arteriosclerosis* 1987; 7: 421-425.
11. Allayee H, Castellani LW, Cantor RM, de Bruin TW, Lusis AJ. Biochemical and genetic association of plasma apolipoprotein A-II levels with familial combined hyperlipidemia. *Circ Res* 2003; 92: 1262-1267.
12. Boomsma DI, Kempen HJ, Gevers LJ, Havekes L, de Knijff P, Frants RR. Genetic analysis of sex and generation differences in plasma lipid, lipoprotein, and apolipoprotein levels in adolescent twins and their parents. *Genet Epidemiol* 1996; 13: 49-60.
13. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.
14. Heller DA, de Faire U, Pedersen NL, Dahlen G, McClearn GE. Genetic and environmental influences on serum lipid levels in twins. *N Engl J Med* 1993; 328: 1150-1156.
15. Boomsma DI, Kaptein A, Kempen HJ, Gevers LJ, Princen HM. Lipoprotein(a): relation to other risk factors and genetic heritability. Results from a Dutch parent-twin study. *Atherosclerosis* 1993; 99: 23-33.
16. Snieder H, Boomsma DI, van Doornen LJ, de Geus EJ. Heritability of respiratory sinus arrhythmia: dependency on task and respiration rate. *Psychophysiology* 1997; 34: 317-328.
17. Heath AC, Bucholz KK, Madden PA, Dinwiddie SH, Slutske WS *et al.* Genetic and environmental contributions to alcohol dependence risk in a national twin sample: consistency of findings in women and men. *Psychol Med* 1997; 27: 1381-1396.
18. Pedersen NL, Friberg L, Floderus-Myrhed B, McClearn GE, Plomin R. Swedish early separated twins: identification and characterization. *Acta Genet Med Gemellol (Roma)* 1984; 33: 243-250.
19. Pedersen NL, McClearn GE, Plomin R, Nesselroade JR, Berg S, DeFaire U. The Swedish Adoption Twin Study of Aging: an update. *Acta Genet Med Gemellol (Roma)* 1991; 40: 7-20.
20. Bury J, Vercaemst R, Rosseneu M, Belpaire F. Apolipoprotein E quantified by enzyme-linked immunosorbent assay. *Clin Chem* 1986; 32: 265-270.
21. Albers JJ, Wahl PW, Cabana VG, Hazzard WR, Hoover JJ. Quantitation of apolipoprotein A-I of human plasma high density lipoprotein. *Metabolism* 1976; 25: 633-644.
22. Havekes L, Hemmink J, de Wit E. Low-density-lipoprotein apoprotein B in plasma as measured by radial immunodiffusion and rocket immunoelectrophoresis. *Clin Chem* 1981; 27: 1829-1833.
23. Lopes-Virella MF, Stone P, Colwell JA. Cholesterol determination in high-density lipoproteins separated by three different methods. *Clin Chem* 1977; 23: 882-884.
24. Jong MC, van Dijk KW, Dahlmans VE, Van der BH, Kobayashi K *et al.* Reversal of hyperlipidaemia in apolipoprotein C1 transgenic mice by adenovirus-mediated gene delivery of the low-density-lipoprotein receptor, but not by the very-low-density-lipoprotein receptor. *Biochem J* 1999; 338: 281-287.

25. Zonderland ML, Erich WB, Peltenburg AL, Havekes L, Bernink MJ, Huisveld IA. Apolipoprotein and lipid profiles in young female athletes. *Int J Sports Med* 1984; 5: 78-82.
26. Maciejko JJ, Levinson SS, Markyvech L, Smith MP, Blevins RD. New assay of apolipoproteins A-I and B by rate nephelometry evaluated. *Clin Chem* 1987; 33: 2065-2069.
27. Bucolo G, David H. Quantitative determination of serum triglycerides by the use of enzymes. *Clin Chem* 1973; 19: 476-482.
28. Allen JK, Hensley WJ, Nicholls AV, Whitfield JB. An enzymic and centrifugal method for estimating high-density lipoprotein cholesterol. *Clin Chem* 1979; 25: 325-327.
29. Seigler L, Wu WT. Separation of serum high-density lipoprotein for cholesterol determination: ultracentrifugation vs precipitation with sodium phosphotungstate and magnesium chloride. *Clin Chem* 1981; 27: 838-841.
30. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; 18: 499-502.
31. Rifai N, Warnick GR, McNamara JR, Belcher JD, Grinstead GF, Frantz ID. Measurement of low-density-lipoprotein cholesterol in serum: a status report. *Clin Chem* 1992; 38: 150-160.
32. Neale MC, Cardon LR. Methodology for genetic studies of twins and families. Dordrecht: Kluwer Academic Publishers; 1992.
33. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
34. Berg K. Twin studies of coronary heart disease and its risk factors. *Acta Genet Med Gemellol (Roma.)* 1984; 33: 349-361.
35. Sistonen P, Ehnholm C. On the heritability of serum high density lipoprotein in twins. *Am J Hum Genet* 1980; 32: 1-7.
36. Shelley JM, Smith AM, Dudley E, Dennerstein L. Use of hormone replacement therapy by Melbourne women. *Aust J Public Health* 1995; 19: 387-392.
37. Barentsen R. The climacteric in The Netherlands: a review of Dutch studies on epidemiology, attitudes and use of hormone replacement therapy. *Eur J Obstet Gynecol Reprod Biol* 1996; 64 Suppl: S7-11.
38. Crook D, Seed M. Endocrine control of plasma lipoprotein metabolism: effects of gonadal steroids. *Baillieres Clin Endocrinol Metab* 1990; 4: 851-875.
39. Birley AJ, MacLennan R, Wahlqvist M, Gerns L, Pangan T, Martin NG. MN blood group affects response of serum LDL cholesterol level to a low fat diet. *Clin Genet* 1997; 51: 291-295.
40. Rice T, Vogler GP, Perry TS, Laskarzewski PM, Rao DC. Familial aggregation of lipids and lipoproteins in families ascertained through random and nonrandom probands in the Iowa Lipid Research Clinics family study. *Hum Hered* 1991; 41: 107-121.
41. Vogler GP, Wette R, Laskarzewski PM, Perry TS, Rice T *et al.* Heterogeneity in the biological and cultural determinants of high-density lipoprotein cholesterol in five North American populations: the Lipid Research Clinics Family Study. *Hum Hered* 1989; 39: 249-257.
42. Hunt SC, Hasstedt SJ, Kuida H, Stults BM, Hopkins PN, Williams RR. Genetic heritability and common environmental components of resting and stressed blood pressures, lipids, and body mass index in Utah pedigrees and twins. *Am J Epidemiol* 1989; 129: 625-638.
43. Ellison RC, Myers RH, Zhang Y, Djousse L, Knox S *et al.* Effects of similarities in lifestyle habits on familial aggregation of high density lipoprotein and low density lipoprotein cholesterol: the NHLBI Family Heart Study. *Am J Epidemiol* 1999; 150: 910-918.
44. Reilly SL, Ferrell RE, Kottke BA, Kamboh MI, Sing CF. The gender-specific apolipoprotein E genotype influence on the distribution of lipids and apolipoproteins in the population of Rochester, MN. I. Pleiotropic effects on means and variances. *Am J Hum Genet* 1991; 49: 1155-1166.
45. Neale MC, de Knijff P, Havekes LM, Boomsma DI. ApoE polymorphism accounts for only part of the genetic variation in quantitative ApoE levels. *Genet Epidemiol* 2000; 18: 331-340.
46. Lambert JC, Brousseau T, Defosse V, Evans A, Arveiler D *et al.* Independent association of an APOE gene promoter polymorphism with increased risk of myocardial infarction and decreased APOE plasma concentrations-the ECTIM study. *Hum Mol Genet* 2000; 9: 57-61.

3 | A powerful and rapid approach to human genome scanning in small quantities of genomic DNA

Marian Beekman, Nico Lakenberg, Stacey S. Cherny, Peter de Knijff, C. Cornelis Kluft, Gert-Jan B. van Ommen, George P. Vogler, Rune R. Frants, Dorret I. Boomsma, P. Eline Slagboom

Genetical Research 2001; 77:129-134

Abstract

Dense maps of short-tandem-repeat polymorphisms (STRPs) have allowed genome-wide searches for genes involved in a great variety of diseases with genetic influences, including common complex diseases. Generally for this purpose, marker sets with a 10 cM spacing are genotyped in hundreds of individuals. We have performed power simulations to estimate the maximum possible inter-marker distance that still allows for sufficient power. In this paper we further report on modifications of previously published protocols, resulting in a powerful screening set containing 229 STRPs with an average spacing of 18.3 cM. A complete genome scan using our protocol requires only 80 multiplex PCR reactions which are all carried out using one set of conditions and which do not contain overlapping marker allele sizes. The multiplex PCR reactions are grouped by sets of chromosomes, which enables on line statistical analysis of a set of chromosomes, as sets of chromosomes are being genotyped. A genome scan following this modified protocol can be performed using a maximum amount of 2.5 μg of genomic DNA per individual, isolated from either blood or from mouth swabs.

Introduction

The localisation of genes in genome-wide searches is performed by the use of saturated maps of short-tandem-repeat polymorphisms (STRPs) or microsatellite loci. Recent applications of STRP maps are genome scans for genes involved in common multifactorial diseases such as type II diabetes[1-3], osteoarthritis[4] or for QTL loci determining quantitative variables associated with disease risk[5].

Genome scans thus far comprise genotypings of on average 400 microsatellite marker loci at 10 cM spacing in hundreds of individuals and are highly time consuming. To increase the efficiency, a balanced choice of markers can be made based on the maximum possible spacing to limit the number of genotypings at a minimum risk of type II errors. To minimise genotyping errors, tri- and tetranucleotide repeat polymorphisms can more accurately be scored than CA-repeat polymorphisms, because of greatly reduced strand slippage during amplification and concomitant simplification of banding patterns. Additionally, CA-repeat loci have a mutation rate 1.5 to 2 times higher than tetranucleotide repeat loci[6]. Multiplex PCR reactions, amplifying several STRPs simultaneously, should be developed in which markers in the same reaction do not overlap.

We have modified existing protocols of Human Screening Set 8[7], developed by the Marshfield Medical Research Foundation, for genome scanning in which minimal amounts of genomic DNA, including DNA collected by the non-invasive mouth swab procedure[8] is required.

Material en methods

Power simulation

Power simulations were based on 500 unselected sib pairs and a trait heritability of 75%. Simulations were carried out for markers with a heterozygosity of 75%. The QTL effect accounted for 25% of the phenotypic variance and the QTL was always located in-between two loci. Data were analysed using a maximum-likelihood variance components approach modelling the full 2x2 sibship covariance structure[9], with pi-hats estimated from IBD distributions obtained from running Mapmaker/SIBS[10].

Marker set

The Human Screening set 8[7] was used as a basis for composing the alternative screening set. The primer sequences and the inter-marker distances are available at the Marshfield Medical Research Foundation website (http://research.marshfieldclinic.org/genetics).

DNA isolation

Genomic DNA was isolated from whole blood and from mouth swabs using a chloroform:isoamylalcohol (24:1) extraction method (see below). Mouth swabs were taken from subjects according to our previously published method[8], but since the chloroform:isoamylalcohol (24:1) extraction method appeared to be as successful as the phenol/chloroform extraction we used the former. The mouth swabs were collected from participants by mail. Mouth swabs were taken and stored by the participants in STE buffer (100mM NaCl, 10 mM EDTA, 10 mM Tris) including 0.1 mg/ml proteinase K and 0.5% SDS. At the arrival of the swabs by mail in the TNO laboratories, the proteinase K concentration was increased up to 0.2 mg/ml and the swabs were incubated at 65°C for two hours. After centrifugation of the samples as described[8], 0.2 volume of 8 M KAc was added to the lysate and mixed well but gently. The mixture was kept on ice for 15 minutes. One volume of chloroform:isoamylalcohol (24:1) was added and placed in a top-over

for half an hour. Subsequently samples were centrifuged and isopropanol was added to the supernatant for precipitation of genomic DNA. DNA of 1911 individuals was obtained using this mouth swab procedure and the average yield per cotton bud was 2.0 ± 1.4 µg genomic DNA on average. Depending on the subjects taking their mouth swabs, the DNA yield is quite variable. For DNA extraction from whole blood lysis was performed as regular, followed by the same extraction steps as for mouth swabs.

PCR conditions

The PCR was performed in 96 wells V-microtiter plates (Biozyme) in a total volume of 10 µl. The reaction mix contained 10 ng genomic DNA template, either isolated from whole blood or mouth swabs, PCR buffer as supplied with the enzyme by the manufacturer (Amersham Pharmacia Biotech), 0.073 µM of each primer, of which the forward primer was labeled with Cy5 (Amersham Pharmacia Biotech), 200 µM of each dATP, dTTP, dGTP, dCTP (Amersham Pharmacia Biotech), 0.2 units recombinant Taq DNA polymerase (Amersham Pharmacia Biotech). To make the multiplex PCR reaction successful, it was only necessary in some cases to adapt the final primer concentration. The PCR reactions were performed in PTC-100 machines (MJ Research, Inc). The program consisted of an initial denaturation for 1 minute at 94°C, followed by 27 cycles of 30 seconds denaturing at 94°C, 75 seconds annealing at 55°C and 15 seconds extension at 72°C, which were concluded with a final extension of 6 minutes at 72°C.

Electrophoresis

The electrophoresis and fragment separation were performed using short gel systems of the automated laser fluorescent DNA sequence analyser ALF*express* (Amersham Pharmacia Biotech). When Sequagel-6 (National Diagnostics) was used, the gel could be loaded twice with PCR products and when High Resolution ReproGel (Amersham Pharmacia Biotech), which polymerises during 10 minutes of exposure to UV light, was used, the gel could be loaded at least three times in subsequent order, still resulting in sharp peak patterns.

Genotyping

The allele analysis was performed using the Fragment Analyser software (Amersham Pharmacia Biotech). The genotypes of the parents of eight CEPH families (102, 884, 1331, 1332, 1347, 1362, 1413, and 1416) were used for composing the allelic standards. Each genotype was reviewed manually by two individuals to confirm the accuracy of allele calling.

Results

Simulations showed that genome scans with a spacing of 20 cM have enough power to detect QTL effects, which account for 25% of the phenotypic variance of a trait with a heritability of 75%, in a population of 500 sibpairs. The 25 cM Human Screening Set 8a has thus too little power in searches for such QTLs and the 10 cM Human Screening Set 8[7] seems too elaborate. Therefore, we composed an alternative set of 229 markers based on the Human Screening Set 8 with an average spacing of 18.3 ± 3.9 cM (mean ± SD). The average heterozygosity of these markers is 0.77 ± 0.06 (mean ± SD) and 86.5% are tri- and tetranucleotide repeat polymorphisms. The Marshfield Medical Research Foundation described multiplex PCR combinations for their 10 cM spaced screening set (http://research.marshfieldclinic.org/genetics). We redesigned the composition of markers in each multiplex PCR and the corresponding conditions. The resulting 80 multiplex PCR reactions (Appendix 3) are grouped by markers randomised over sets of 3 to 5

Table 1.
Composition of the screening sets of chromosomes.

Set of chromosomes	Multiplex reaction numbers
1, 8, 11, 19	1-23, 42, 43, 67
6, 7, 16	6, 9, 10, 13, 14, 18, 20, 24-31
2, 15, 17	7, 8, 10, 13, 18, 32-43
5, 9, 14	12, 16, 19, 40, 41, 44-54, 78
3, 4, 10	11, 15, 29, 40, 51, 55-64, 78
12, 13, 18	6, 12, 52-54, 62, 65-72
20, 21, 22, X, Y	15, 17, 21, 52, 54, 62, 72-80

chromosomes (Table 1) to enable statistical analysis of the search data before finishing the whole genome scan at large. In addition, a margin of at least 15 base pairs between allele sizes of different markers in a multiplex was chosen to avoid overlap.

A protocol was developed in which only 10 ng genomic DNA per multiplex PCR reaction is used. Hence, a complete genome scan requires at the most 2.5 µg of genomic DNA. Our protocol is successful using DNA isolated from whole blood as well as from mouth swabs. After evaluation of 90,000 genotypes, which were performed in 16 months using two ALF*express* (Amersham Pharmacia Biotech) systems, the average genotypic error rate was less than 1% and the average missing data rate was 8% using this protocol.

Discussion

We performed power simulations and adapted existing protocols in order to limit the number of genotypings in a genome scan and still retain sufficient power in statistical analysis. This resulted in a set of 229 markers, measured in 80 multiplex PCR reactions, with an average inter-marker distance of 18.3 cM. The high percentage of tri- and tetranucleotide repeat markers reduces the percentage of genotypic errors. The markers in each multiplex PCR reaction, which all require the same PCR conditions, allow 15 base pairs spacing of alleles to avoid overlap. The genotypings were performed using an ALF*express* system (Amersham Pharmacia Biotech), but these multiplex PCR reactions can easily be used in other genotyping equipment.

Multiplex PCR reactions were composed from marker combinations grouped by sets of several chromosomes. Statistical analysis can therefore be performed per set of chromosomes, while other sets are being genotyped. Using our protocol, fine-mapping of positive chromosome regions could be started in an earlier stage.

Due to the adapted PCR conditions smaller than usual amounts of DNA are required. For a whole genome scan using our protocol at the most 2.5 µg genomic DNA is required, whereas for a 10 cM spaced genome scan of the Marshfield Medical Research Foundation (http://research.marshfieldclinic.org/genetics), a 6-fold increment of genomic DNA is required.

Finally, our genotyping procedure is successful on genomic DNA isolated from blood as well as from mouth swabs collected by mail. The use of DNA isolated from mouth swabs is especially suitable for studies involved in geographically scattered subjects and for studies in which it is too elaborate to obtain blood from participants.

In conclusion, we have increased the efficiency of genome scanning and developed a protocol to facilitate scanning in small quantities of genomic DNA.

Acknowledgements

This study was supported by a grant from the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976). Stacey S. Cherny is supported in part by NIH grant EY-12562 and a Programme Project grant from the Medical Research Council of Great Britain. We furthermore acknowledge the Marshfield group for providing all the information on the Human Screening Set 8/8a on the Internet.

References

1. Mahtani MM, Widen E, Lehto M, Thomas J, McCarthy M *et al.* Mapping of a gene for type 2 diabetes associated with an insulin secretion defect by a genome scan in Finnish families. *Nat Genet* 1996; 14: 90-94.
2. Ghosh S, Hauser ER, Magnuson VL, Valle T, Ally DS *et al.* A large sample of Finnish diabetic sib-pairs reveals no evidence for a non-insulin-dependent diabetes mellitus susceptibility locus at 2qter. *J Clin Invest* 1998; 102: 704-709.
3. Ghosh S, Watanabe RM, Hauser ER, Valle T, Magnuson VL *et al.* Type 2 diabetes: evidence for linkage on chromosome 20 in 716 Finnish affected sib pairs. *Proc Natl Acad Sci USA* 1999; 96: 2198-2203.
4. Chapman K, Mustafa Z, Irven C, Carr AJ, Clipsham K *et al.* Osteoarthritis-Susceptibility Locus on Chromosome 11q, Detected by Linkage. *Am J Hum Genet* 1999; 65: 167-174.
5. Xu X, Rogus JJ, Terwedow HA, Yang J, Wang Z *et al.* An extreme-sib-pair genome scan for genes regulating blood pressure. *Am J Hum Genet* 1999; 64: 1694-1701.
6. Chakraborty R, Kimmel M, Stivers DN, Davison LJ, Deka R. Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc Natl Acad Sci USA* 1997; 94: 1041-1046.
7. Yuan B, Vaske D, Weber JL, Beck J, Sheffield VC. Improved set of short-tandem-repeat polymorphisms for screening the human genome. *Am J Hum Genet* 1997; 60: 459-460.
8. Meulenbelt I, Droog S, Trommelen GJ, Boomsma DI, Slagboom PE. High-yield noninvasive human genomic DNA isolation method for genetic studies in geographically dispersed families and populations. *Am J Hum Genet* 1995; 57: 1252-1254.
9. Fulker DW, Cherny SS. An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet* 1996; 26: 527-532.
10. Kruglyak L, Lander ES. A nonparametric approach for mapping quantitative trait loci. *Genetics* 1995; 139: 1421-1428.

4 | Genome scan for loci influencing plasma levels of lipids and apolipoproteins

Abstract

The genetic basis of cardiovascular disease (CVD) with its complex etiology is still largely elusive. Plasma levels of lipids and apolipoproteins are among the major quantitative risk factors for CVD and are well-established intermediate traits that may be more accessible to genetic dissection than clinical CVD endpoints. To identify quantitative trait loci influencing lipid and apolipoprotein levels in the general population, we performed a genome scan with an inter-marker distance of 18 cM in four independent samples including adolescent Dutch twins and adult Dutch, Swedish and Australian twins totalling 493 dizygotic twin pairs. Chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17, and 19 were analysed for linkage (the other chromosomes are currently scanned). These chromosomes harbour candidate genes for lipid and apolipoprotein levels, including two known quantitative trait loci: the *LPA* locus on chromosome 6q27 is known to explain approximately 90% of the population variation in Lp(a) plasma levels[1] and the *APOE* locus on chromosome 19q13.32 explains approximately 15% of apoE plasma levels[2,3]. These two loci were tested for linkage to investigate whether QTLs with major as well as moderate effects can be detected using our genome scanning protocol. In this paper, we report 4 significant linkage results, i.e. maximum LOD scores (MLS) higher than 3.6 and 7 suggestive linkage results, i.e. MLS between 2.2 and 3.6. Since we have unique samples to test for replication of linkage results, we also report linkage results with MLS higher than 2.2 in one twin sample and MLS higher than 1.4 in an other twin samples within 20 cM for the same trait. Potentially interesting loci, include putative Lp(a) QTLs on chromosome 1 and 2, a putative apoB QTL on chromosome 11, and a putative QTL on chromosome 19 influencing LDL cholesterol levels.

Introduction

Dissecting the genetic basis of cardiovascular disease (CVD) is complicated by the etiologic heterogeneity underlying seemingly indistinguishable clinical endpoints. Major quantitative risk factors for CVD such as lipid and apolipoprotein levels may be more accessible to genetic dissection[4]. Despite many candidate gene studies, quantitative trait loci (QTLs) with a major effect on lipid an apolipoprotein levels still have to be identified[5] if they exist at all. Multiple genome scans were conducted to map QTLs in lipid metabolism. Thus far, these scans have mainly been performed in genetically isolated populations and in patient populations. However, linkage results are hardly replicated[6]. We investigated four independent samples of outbred Caucasian subjects who were collected without regard to any disease or trait. These include adolescent Dutch twins and adult Dutch, Swedish and Australian twins; a total of 493 dizygotic twin pairs participated in the study. In these pairs, we performed a genome scan of 18 cM spacing, which simulations suggested to be powerful enough to detect QTLs with a major effect in this data set[7]. Linkage results for chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17, and 19 are reported while other chromosomes are currently being scanned.

Chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17, and 19 harbour many candidate genes encoding lipid and apolipoprotein related proteins. The most obvious candidate genes on these chromosomes are listed in table 1. It is investigated whether we are able to detect QTLs with major as well as moderate effects on traits using our genome scanning protocol. Therefore, the *LPA* gene locus on chromosome 6q27, which is known to explain approximately 90% of the population variation in Lp(a) plasma levels[1], is tested for linkage and also, the *APOE* gene locus on chromosome 19q13.32, which is known to explain approximately 15% of variance in apoE plasma levels[2,8,9]. Subsequently, variance components linkage analyses of the 10 chromosomes were conducted in all twin samples separately, to find putative QTLs involved in lipid and apolipoprotein levels.

Table 1.
List of most obvious candidate gene loci on the chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17, and 19. This overview is based on information in the NCBI database and is updated until May 2003.

Chromosome	Gene	Locus	Position in cM from pter (Marshfield)
1	ARH	1p36-35	55
	LRP8	1p34	70
	HCHOLA3	1p32	76
	ANGPLT3	1p33-22	96
	LMNA	1q21.2	76
	APOA1BP	1q22	96
	APOA2	1q21-23	162
	SOAT1	1q25	162
2	LPIN1	2p25.1	29
	APOB	2p24-23	41
	PLB	2p23.3	49
	CYP1B1	2p21	58
	ABCG5	2p21	66
	ABCG8	2p21	66
	FABP1	2p11	111
	LRP1B	2q21.2	150
	ABC11	2q24	173
	LRP2	2q24-31	173

Chromosome	Gene	Locus	Position in cM from pter (Marshfield)
6	AGPAT1	6p21.3	45
	ACAT2	6pq	165
	LPA	6q27	165
7	NPY	7p15.1	37
	CD36	7q11.2	94
	ABCB1	7q21.1	98
	ABCB4	7q21.1	98
	PON1	7q21	107
	PON2	7q21	107
	PON3	7q21	107
8	MSR1	8p22	33
	LPL	8p22	40
	CYP7A1	8q11-12	76
	DGAT1	8q-qter	166
11	LRP4	11p11.2-12	58
	NR1H3	11p11.2	58
	LRP7	11q13.4	68
	DGAT2	11q13.3	78
	ACAT1	11q22.3-23.1	100
	APOA1/C3/A4 gene cluster	11q23	111
15	CHR39B	15q15-21	48
	LIPC	15q21-23	51
	CYP11A	15q23-24	73
16	APOEL1	16q11-24	58
	APOB48R	16q11-24	58
	CETP	16q13	77
	LYPLA3	16q22.1	87
	LCAT	16q22	87
17	SREBR1	17p11	44
	APOH	17q23-qter	86
19	INSR	19p13.3-13.2	24
	LDLR	19p13.2-13.1	35
	LRP3	19q13.12	57
	CEBPA	19q13.12	57
	LISCH7	19q13.12	57
	CYP2A7	19q13.2	64
	CYP2B7	19q13.2	64
	LIPE	19q13.2	66
	APOE/C1/C4/C2 gene cluster	19q32	69

Subjects and methods

Subjects

In this study we had access to data from four twin pair samples of 83 dizygotic (DZ) adolescent Dutch twins (aged 13-22) and 117 DZ adult Dutch twins (aged 34-62) (Project for cardiovascular risk factors), 44 DZ Swedish twin pairs (aged 42-81) (Swedish Adoption/Twin Study of Aging) and 249 DZ Australian twin pairs (aged 32-80) (Semi-Structured Assessment for the Genetics of Alcoholism study) twins. The latter were selected from 651 DZ twin pairs on the basis of both twins being in the upper or lower 16% tail of the distribution of total cholesterol, apoB, or

triglyceride levels, to investigate the most informative sib pairs[10]. The recruitment of the twins and the measurements of lipid and apolipoprotein levels in the different samples are described elsewhere[11]. The relationships of the twins were tested with more than 100 STR genotypings of the 10 chromosomes scanned using the Graphical Representation of Relationships (GRR) software package[12]. The relationship of 3 DZ twin pairs appeared to be MZ and 10 DZ pairs appeared to be unrelated. Except for one sample mix-up that could be corrected, these twin pairs were excluded from linkage analyses, leaving 493 DZ twin pairs with confirmed relationships. Total cholesterol, HDL cholesterol, apolipoprotein B, apolipoprotein A1, apolipoprotein A2, triglycerides, Lp(a) and apolipoprotein E levels were assessed in plasma using standard procedures[11]. Concentrations of low density lipoprotein (LDL) cholesterol were calculated according to Friedewald[13]. LDL cholesterol level was considered missing, when the triglyceride concentration did exceed 4.52 mmol/L, which was the case for 1 individual of the middle-aged Dutch sample, for 82 individuals of the Australian sample and for 11 individuals of the Swedish sample[14].

Power

Power simulations were based on 500 unselected sib pairs and a trait heritability of 75%. Simulations were carried out for markers with a heterozygosity of 75%. The QTL effect accounted for 25% of the phenotypic variance and the QTL was always located in-between two loci[7]. Data were analysed using a maximum-likelihood variance components approach modelling the full 2x2 sibship covariance structure[15] in the software package MX 1.50d[16], with IBD probabilities obtained from running Mapmaker/SIBS[17]. A 20 cM spacing resulted in a power of 80% to obtain a LOD score of 0.8.

Genotyping

Short tandem repeats (STRs) with an average inter-marker distance of 18.3 centiMorgans (cM) on chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17 and 19 were genotyped as described elsewhere[7]. The average heterozygosity for these markers was estimated at 0.77 and the Marshfield genetic map was used (http://research.marshfieldclinic.org/genetics/). For genotyping, the Cy5-labeled PCR products were electrophoretically separated on an automated-fluorescence DNA sequencer, ALF*express* (Amersham Pharmacia Biotech). Analysis and assignment of the marker alleles were performed with Fragment Analyser 1.02 (Amersham Pharmacia Biotech). To reduce genotyping errors, one known genotype was present on each gel, 5% of the genotypings were repeated and two independent individuals performed the allele calling. SIBMED[18] was used to identify unlikely double recombinants, the occurrence of which may be due to genotyping errors. After running SIBMED and checking the raw genotyping data of the 10 chromosomes in the 493 DZ twin pairs, approximately 6 possible genotyping errors per chromosome had to be checked. On average 1.5 genotypings per chromosome appeared to be erroneous (0.1‰ of all genotypings). Dependent on the error, these genotypes were changed into the right genotype or were set to missing. The cleaned genotyping data were used for linkage analysis.

Statistical analysis

Multipoint linkage analyses were carried out using the variance components approach as implemented in Genehunter 2.1[19] for each twin sample separately. Since age and sex contribute to lipid and apolipoprotein levels, these factors were used as covariates in the linkage analyses. Allele frequencies were estimated separately for the twin samples using marker data for all

individuals[20]. Plasma levels of triglycerides, Lp(a) and apoE showed a skewed distribution and these values were therefore transformed by natural logarithm prior to analysis.
To investigate the heterogeneity in QTL effects among the twin samples on the test loci *LPA* and *APOE*, simultaneous analyses were carried out. The simultaneous linkage analyses were performed by a variance components approach as can be implemented in the software Mx[16], which allows the inclusion of also data in MZ twins. The total variance was then modelled as A+E+Q, where A represents additive genetic background factors, E non-shared environmental factors and Q the QTL effect. The covariance equals A+Q for MZ twins and ½A+ $\hat{\pi}$Q for DZ twins, where $\hat{\pi}$ is estimated as ½P(IBD=1)+P(IBD=2). Each variance component was estimated separately for each twin sample. First, heterogeneity between the effects of the QTL in the different twin samples was tested by comparing the model estimating all parameters for the samples separately, with the model in which the proportions of the QTL effects were constrained to be equal. When the latter fits not significantly worse, no heterogeneity in the QTL effects is assumed. Although the proportions of the QTL effects were equated over the different twin samples, the additive genetic and unique environmental variance components were estimated separately per twin sample. To test for linkage, the fit of the AE model was compared to the fit of the AEQ model.
The distribution of multipoint identity-by-descent (IBD) sharing probabilities was estimated using Genehunter 2.1[19] with twin sample specific allele frequencies. Since Lp(a) levels do not show a normal distribution even after transformation, the $\hat{\pi}$ approach was used, which seem to be less sensible for non-normality than the weighted likelihood approach. The latter makes use of the full distribution of IBD-probabilities[21] and takes, therefore, all genetic information into account. For the analysis of apoE levels the weighted likelihood approach was used with adjustments for age and sex, which contribute to apoE levels. Age and sex are no risk factors for Lp(a) levels and therefore, no adjustments are made in the simultaneous analysis.

Results

In table 2, the characteristics of phenotypes measured in the adolescent Dutch, adult Dutch,

Table 2.
Characteristics of adolescent Dutch and adult Dutch, Swedish and Australian dizygotic twin samples.

	Adolescent		Adult					
Phenotype	**Netherlands**		**Netherlands**		**Sweden**		**Australia**	
	(n=166)		(n=234)		(n=88)		(n=498)	
men, %	49.4		48.7		59.1		35.7	
Age, years – mean (range)	17	(13-22)	44	(34-59)	65	(42-81)	44	(31-80)
Body Mass Index, kg/m^2 – mean (SD)	20.28	(2.21)	24.64	(3.06)	25.21	(3.07)	25.44	(5.19)
LDL-C, mmol/L – mean (SD)	2.56	(0.65)	3.63	(0.99)	4.26	(1.07)	3.37	(1.07)
ApoB, g/L – mean (SD)	0.79	(0.17)	1.21	(0.35)	1.04	(0.21)	0.97	(0.28)
Total cholesterol, mmol/L – mean (SD)	4.15	(0.70)	5.39	(1.06)	6.50	(1.11)	5.64	(1.20)
HDL-C, mmol/L – mean (SD)	1.28	(0.26)	1.22	(0.38)	1.47	(0.34)	1.46	(0.41)
ApoA1, g/L – mean (SD)	1.38	(0.20)	1.67	(0.43)	1.39	(0.30)	1.44	(0.27)
ApoA2, g/L – mean (SD)	0.58	(0.08)	-	-	-	-	0.36	(0.06)
Triglycerides, mmol/L – mean (SD)	0.68	(0.30)	1.22	(0.71)	1.62	(0.87)	1.90	(1.69)
Lipoprotein (a), mg/L – mean (SD)	14.2	(21.2)	18.4	(26.6)	175.7	(158.1)	214.8	(267.3)
ApoE, mg/dL – mean (SD)	6.59	(2.39)	2.54	(1.02)	-	-	4.18	(1.88)

Swedish and Australian twin samples are summarised. The histograms of the phenotype distributions are shown in Appendix 4, for each twin pair sample separately.
Using the phenotypic information of both MZ and DZ twin pairs, heritabilities of the lipid and apolipoprotein levels were estimated[11], which are summarised in table 3. Maximum LOD score was estimated twice per twin sample (Table 4) using the Genetic Power Calculator[22]. As input parameters we used once the lowest, and once highest estimated trait heritability among the different twin samples (Table 3), a QTL effect of half the heritability and a recombination fraction of 0.
According to the Lander&Kruglyak criteria[23], in sib pair analyses maximum LOD scores (MLS) higher than 3.6 can be assigned as significant linkage, which is expected to occur at random once in 20 genome scans. Since the linkage analyses in the separate twin samples have limited power (Table 4), this criterion may be too stringent for detecting QTLs in the separate twin samples. Therefore, suggestive linkage results (MLS between 2.2 and 3.6), which are expected to occur one time at random in a genome scan, are also reported in this paper. However, to be credible, linkage results must be replicated and the four different twin provide a unique opportunity to look for replication. According to Lander & Kruglyak, for replication a point-wise p-value of 0.01 is needed for a 20 cM interval-wide significance level of 0.05[23]. Since this point-wise p-value corresponds with a LOD score of 1.4, the loci with MLS higher than 2.2 in one twin sample and 1.4 in a second twin samples within approximately 20 cM are also reported in this paper.
Previous simulations showed that genome scans with a spacing of 20 cM have the power to detect QTL effects, which account for 25% of the phenotypic variance of a trait with a heritability of 75%, in a population of 500 sib pairs[7]. It is, therefore, expected that only major loci influencing lipid and apolipoprotein levels will be detected in our genome scan. It was investigated whether linkage could be detected at the *LPA* locus, known to determine between 85 and 95% of the variation in Lp(a) levels. The MLS detected on chromosome 6q were 2.7, 4.4, 2.3, and 5.1 in adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively (figure 1, left panel). When all 493 dizygotic twin pairs were analysed simultaneously, constraining the proportion of the QTL effects to be equal over the different samples, no evidence for

Table 3.
Heritability of the phenotypes measured in the adolescent Dutch and adult Dutch, Swedish and Australian dizygotic twin samples.

	Adolescent		**Adult**					
Heritability	**Netherlands**		**Netherlands**		**Sweden**		**Australia**	
Age, years – mean (range)	**17**	(13-22)	**44**	(34-59)	**65**	(42-81)	**44**	(31-80)
LDL-C – 95% CI	**0.83**	0.76-0.88	**0.77**	0.69-0.84	**0.66**	0.55-0.75	**0.61**	0.57-0.65
ApoB – 95% CI	**0.82**	0.75-0.87	**0.79**	0.71-0.85	**0.75**	0.66-0.81	**0.63**	0.58-0.66
Total cholesterol – 95% CI	**0.82**	0.74-0.87	**0.77**	0.68-0.83	**0.67**	0.56-0.76	**0.57**	0.52-0.61
HDL-C – 95% CI	**0.75**	0.65-0.82	**0.72**	0.60-0.80	**0.64**	0.54-0.72	**0.62**	0.57-0.66
ApoA1 – 95% CI	**0.80**	0.71-0.86	**0.72**	0.62-0.79	**0.61**	0.49-0.70	**0.50**	0.45-0.55
ApoA2 – 95% CI	**0.82**	0.79-0.82	**-**	-	**-**	-	**0.48**	0.42-0.53
Triglycerides – 95% CI	**0.71**	0.57-0.80	**0.62**	0.50-0.72	**0.48**	0.35-0.59	**0.50**	0.45-0.55
Lp(a) – 95% CI	**0.95**	0.92-0.97	**0.90**	0.86-0.93	**0.95**	0.93-0.97	**0.89**	0.79-0.90
ApoE – 95% CI	**0.86**	0.80-0.91	**0.87**	0.82-0.91	**-**	-	**0.57**	0.52-0.61

heterogeneity was present (p=0.21). The MLS on chromosome 6q for Lp(a) levels increased to 9.8 at 162 cM pter (figure 1, right panel), which is in close proximity of the *LPA* gene at 165 cM from pter. This QTL on chromosome 6q27 explains 83% of the total variance in Lp(a) levels. Since we detected in two twin samples MLS higher than 3.6 (significant linkage) for Lp(a) levels and in the other two higher than 2.2 (suggestive linkage), the 18 cM spaced marker set[7] is indeed powerful enough to detect significant linkage at loci with a large effect on the variation of a trait.

Secondly, it was investigated whether linkage could be detected at a locus with a moderate effect; the *APOE* gene locus, which is known to determine approximately 15% of the variation in apoE plasma levels. In the Swedish sample, levels of apoE were not available. The MLS on chromosome 19q were 1.0, 0.6, and 0.0 in adolescent Dutch, adult Dutch and Australian twins, respectively (figure 2, left panel). In the simultaneous analysis of the three twin samples, no evidence for heterogeneity at the *APOE* locus was present (p=0.33) and the MLS was 1.0 at 70 cM from pter (figure 2, right panel). Since these linkage results meet the criteria nor for significant neither for suggestive linkage, and also not for replicated linkage, we are not able to detect loci with a moderate effect in our genome scan protocol.

Subsequently, the first set of 10 chromosomes was analysed for linkage for each twin sample separately. Appendix 5 shows tables with maximum LOD scores per chromosome per trait for each twin sample separately and linkage graphs are shown in Appendix 6. In table 5, significant linkage results (MLS higher than 3.6)[23] are listed. Among the significant loci is the *LPA* locus on chromosome 6, and two other QTLs, also influencing Lp(a) levels. At nine loci, we found suggestive linkage (Table 6). Again two loci, on chromosome 1 and 2, are suggestively linked to Lp(a) levels. Two loci, on chromosome 1 and 19, are suggestively linked to LDL cholesterol levels in the Australian and adult Dutch twin samples, respectively. Two loci, on chromosome 7 and 11, are suggestively linked to apoB levels in the Australian twin sample and a locus on chromosome 15 is suggestively linked to apoA1 levels in the adolescent Dutch twin sample. Three loci are suggestively linked to apoA1 levels.

In table 7, the linkage results are listed, which are replicated in a second twin sample with a MLS higher than 1.4 within 20 cM. On chromosome 1, the Swedish and the Australian sample show MLS of 2.6 and 1.8 respectively for Lp(a) levels. In the same chromosomal area, also the adult

Table 4.

Maximum expected LOD score estimated using the Genetic Power Calculator[22] (http://statgen.iop.kcl.ac.uk/gpc/). The maximum LOD score was estimated twice per twin sample. Once using the lowest, and once using the highest estimated trait heritability among the different twin samples (Table 3), a QTL effect of half the heritability and a recombination fraction of 0.

		Adolescent	Adult		
Maximum expected LOD score	**Heritability** (lowest-highest)	**Netherlands** n=83 pairs	**Netherlands** n=117 pairs	**Sweden** n=44 pairs	**Australia** n=249 pairs
LDL-C	61-83	0.4-1.7	0.6-2.4	0.2-0.9	1.3-5.2
ApoB	63-82	0.5-1.6	0.7-2.2	0.3-0.8	1.5-4.7
Total cholesterol	57-82	0.3-1.6	0.5-2.2	0.2-0.8	1.0-4.7
HDL-C	62-75	0.4-1.0	0.6-1.4	0.2-0.5	1.3-3.0
ApoA1	50-80	0.2-1.4	0.3-1.9	0.1-0.7	0.7-4.1
ApoA2	48-82	0.2-1.4	-	-	0.6-4.1
Triglycerides	48-71	0.2-0.8	0.3-1.1	0.1-0.4	0.6-2.3
Lp(a)	89-95	2.8-5.3	3.9-7.5	1.5-2.8	8.4-15.9
ApoE	57-86	0.3-2.1	0.5-3.0	-	1.0-6.4

Figure 1.
The linkage result of chromosome 6 with Lp(a) levels. Left panel shows linkage results in adolescent Ducth (▬▬), adult Dutch (▬ · · ▬ · ·), Swedish (· · · ·) and Australian (▬ ▬ ▬) twins. Right panel shows simultaneous linkage analysis of chromosome 6 with Lp(a) levels in the four twin samples (▬▬).

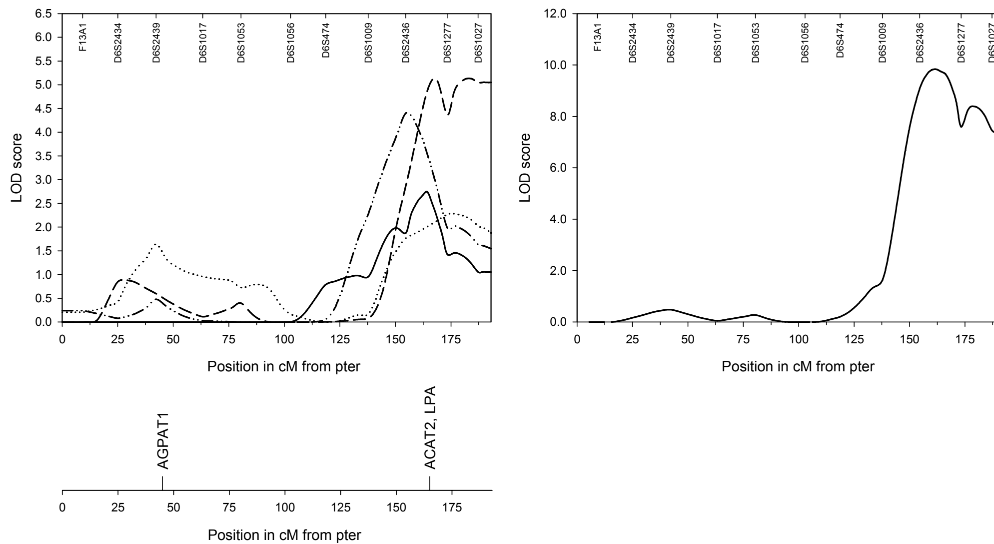


Figure 3.
Linkage result of chromosome 19 with apoE levels. Left panel shows linkage analysis in adolescent Ducth (▬▬), adult Dutch (▬ · · ▬ · ·) and Australian (▬ ▬ ▬) twins. Right panel shows simultaneous linkage analysis of chromosome 6 with Lp(a) levels in the three twin samples (▬▬).

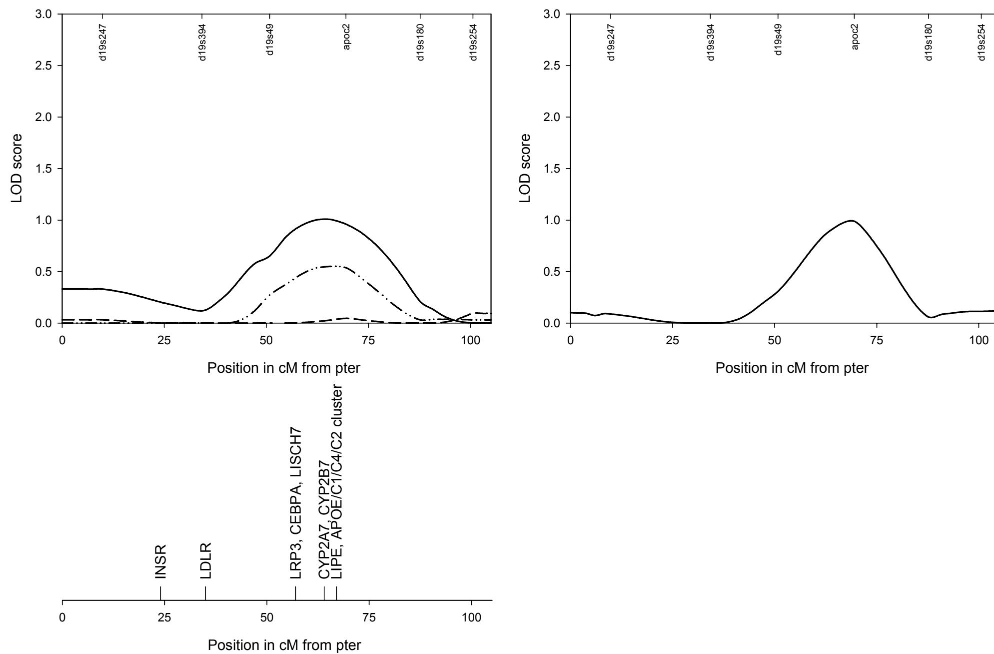


Table 5.
Significant linkage results (MLS ≥ 3.6).

Chromosome	Position in cM from pter	Phenotype	LOD score	Twin sample
2	38.3	Lp(a)	5.6	Australian
6	154.64	Lp(a)	4.4	Adult Dutch
6	181.66	Lp(a)	5.1	Australian
11	113.1	Lp(a)	4.1	Adult Dutch

Table 6.
Suggestive linkage results (2.2 ≤ MLS ≤ 3.6).

Chromosome	Position in cM from pter	Phenotype	LOD score	Twin sample
1	236.15	Lp(a)	2.6	Swedish
1	267.51	LDL cholesterol	2.6	Australian
2	205.6	Lp(a)	2.3	Australian
7	17.74	ApoA1	2.2	Swedish
7	90.52	ApoB	2.8	Australian
8	22.4	ApoA1	2.7	Adolescent Dutch
11	62.9	ApoB	2.6	Australian
15	63.3	ApoA1	2.3	Adolescent Dutch
19	54.5	LDL cholesterol	2.2	Adult Dutch

Table 7.
Replicated linkage results with MLS higher than 2.2 in one twin sample and MLS higher than 1.4 in a second twin sample within 20 cM.

Chromosome	Position in cM from pter	Phenotype	LOD score	Twin sample
1	236.15/251.29	Lp(a)	2.6/1.8	Swedish/Australian
2	17.9/38.3	Lp(a)	1.5/5.6	Swedish/Australian
19	54.5/34.3	LDL cholesterol	2.2/1.7	Adult Dutch/Australian

Dutch twins show positive linkage. On chromosome 2, the Australian sample show significant linkage (MLS=5.6) and the Swedish sample a MLS of 1.5 with Lp(a) levels. On chromosome 19, the adult Dutch and Australian twin samples show linkage with LDL cholesterol levels with MLS of 2.2 and 1.7, respectively. Furthermore, the Swedish twin sample also shows positive linkage (MLS 1.28) in the same chromosomal region with LDL cholesterol levels. For the Swedish twins sample, this MLS is very high, indicating that three twin samples show linkage of chromosome 19 with LDL cholesterol levels. Linkage graphs of the results shown in tables 5, 6 and 7 are presented in Appendix 7.

Discussion

We are conducting a 18.3 cM spaced genome scan with intermediate lipid phenotypes of cardiovascular disease in dizygotic twin pairs from the general population of The Netherlands, Sweden and Australia. In this paper, we reported significant linkage results (MLS higher than 3.6), suggestive linkage results (MLS between 2.2 and 3.6), and replicated linkage results. Since replication involves testing a prior hypothesis, the multiple testing problem associated with a genome scan does not apply. A point-wise p-value of 0.01 is needed for a 20 cM interval-wide significance level of 0.05[23], corresponding to a LOD score of 1.4. We, therefore, consider results with MLS higher than 2.2 in one twin sample and MLS higher than 1.4 in a second twin samples within 20 cM as replicated linkage.

From simulations resulted that 20 cM spacing would be sufficient to detect major gene loci in 500 sib pairs and to confirm this hypothesis we tested the *LPA* locus on chromosome 6q27 for

linkage with Lp(a) plasma levels. This locus is known to explain approximately 90% of the variance in Lp(a) levels[1]. We found significant linkage of chromosome 6q27 with Lp(a) levels in the adult Dutch and Australian twin samples. The simultaneous analysis of all samples increased the LOD score to 9.8 at 162 cM from pter, which is in close proximity of the *LPA* gene (160 cM from pter). We thus conclude that we are able to detect QTLs with major effects on a trait using our genome scan protocol.

In attempt to detect QTLs with moderate effects, the *APOE* locus on chromosome 19 was tested for linkage with plasma levels of apoE. This resulted in MLS of 1.0, 0.6 and 0.0 in the adolescent Dutch twins, adult Dutch twins and Australian, respectively. The simultaneous linkage analysis showed no evidence for heterogeneity and the MLS was 1.0. Since such MLS are expected to occur 20 times at random in a genome scan[23], we conclude that we lack power to detect QTLs with moderate effects using our genome scan protocol. However, the simulations for the inter-marker distance were carried out accounting for a QTL effect of 25% in 500 sib pairs, and therefore, it is reasonable that this QTL explaining approximately 15% of the variance in apoE levels cannot be detected at a significant level in three twin pair samples totalling 449 pairs.

Remarkably, the location of the MLS is in both analyses very close to the actual location of the presumably responsible gene.

From the four significant linkage results, two are actually linkages of chromosome 6 with Lp(a) levels, detecting the *LPA* locus. The other two linkage results involve also Lp(a) levels, but on chromosomes 2 and 11. Since MLS between 2.2 and 3.6 approximate the maximum estimated LOD scores in the separate sample, we report also the suggestive linkage results. Two loci show suggestive linkage with Lp(a) levels, 2 loci with LDL cholesterol levels, two loci with apoB levels and 3 loci with apoA1 levels. The locus on chromosome 8 possibly influencing apoA1 levels in adolescent Dutch twins overlaps with the location of the LPL gene. Variation in the LPL gene has been associated with HDL cholesterol levels[24-26]. Since apoA1 is a major constituent of HDL particles, these plasma levels are highly correlated traits. Variation in the LPL gene might thus underlie this apoA1 QTL on chromosome 8. One locus on chromosome 1 is suggestively linked to LDL cholesterol levels in the Australian twin sample, although the other samples do not show evidence for linkage. Additionally two loci, on chromosome 7 and 11, are suggestively linked to apoB levels in the Australian sample. On chromosome 7 also the adolescent and adult Dutch twins show positive linkage with apoB levels and on chromosome 11 adult Dutch and Swedish twins. It might be that we have not enough power to detect these LDL cholesterol and apoB loci in the other twin pair samples and to gain power in the linkage analyses, the information content of the linkage region should be increased by genotyping additional markers. This is also the case for the locus on chromosome 15, which is suggestively linked to apoA1 levels in adolescent Dutch twins and only weakly in the Swedish and Australian sample. No previous genome scan results are reported on chromosome 7 with plasma levels of apoB. However, on chromosome 11 the population based NHLBI Family Heart Study and a study in Dutch obese families reported linkage in the same area with LDL cholesterol levels[27,28]. Since LDL cholesterol and apoB levels are highly correlated phenotypes, this might be considered as an indication for replicated linkage. This locus on chromosome 11 has also been reported as a Dutch familial combined hypercholesterolaemia (FCH) locus[29]. Since in FCH hepatic apoB is overproduced mainly levels of VLDL and LDL cholesterol are elevated. This FCH locus could therefore also be involved in normal apoB levels. These replicated linkage results indicate that

this locus must be followed up by genotyping additional markers to gain power to detect this apoB QTL also in the other twin samples.

Replication of linkage results is necessary for their credibility and we have unique samples to test for replication. Three loci show our replicated linkage criterion. One locus on chromosome 1 was linked to Lp(a) levels in the Swedish and Australian sample. The adult Dutch sample showed little linkage in this region. Broeckel *et al*[30], reported also linkage in this region with Lp(a) levels and they reported evidence that this locus might be a second Lp(a) locus beside the *LPA* locus on chromosome 6. To investigate whether we can replicate this result, we will perform a two-locus linkage analysis in the follow up of this genome scan result.

Chromosome 2 also shows replicated linkage with Lp(a) levels in the Swedish and Australian samples. In this region on chromosome 2, the *APOB* gene is located, encoding apolipoprotein B, which is the major constituent of the Lp(a) particle. It, therefore, could be hypothesised that mutations in the APOB gene might affect Lp(a) levels. However, in the two Dutch samples we found no evidence for linkage of chromosome 2 with Lp(a) levels and to gain power in the linkage analysis in this region, additional markers need to be genotyped in all twin samples.

On chromosome 19, the adult Dutch and Australian twin samples show replicated linkage with LDL cholesterol levels within 20 cM. Also in the Swedish sample, we found positive linkage in the same chromosomal region with a MLS that can be considered as high for the Swedish sample, which has low power to detect linkage with LDL cholesterol levels. Since this makes this locus even more interesting, we will fine-map chromosome 19 by genotyping additional markers in all twin samples and heterogeneity will be tested in a simultaneous analysis. Our data suggest that on chromosome 19 may be a QTL with a major effect on LDL cholesterol levels. But also many genes with each a smaller effect on LDL cholesterol levels could underlie this linkage result[31,32].

Many candidate genes are located in this broad chromosomal region. The LDLR gene is located at 35 cM from pter (19p13), which is know for its mutations causing hypercholesterolaemia[33]. Furthermore, the *APOE* gene at 70 cM from pter (19q13) is known to influence the variation in LDL cholesterol levels[34]. Also in the linkage region the gene encoding the LDLR-related protein type 3 (LRP3) is located. LRP3 is a family member of the LDLR. It is also expressed in the liver and it binds apoE-containing particles. The function of the LRP3 remains unclear, but as a family member of LDLR, a role in lipid metabolism is assumed[35].

Previous genome scan studies in Pima Indians[36] and Hutterites[37] reported evidence for linkage of chromosome 19 with total cholesterol levels and LDL-C levels, respectively. Our study could extend these findings in genetic isolates to the general, Caucasian population.

The scanning of these 10 chromosomes reveal 11 putative QTLs, of which four would influence Lp(a) levels in addition to the *LPA* locus on chromosome 6. These loci should then all have a very small effect on Lp(a) levels, since the *LPA* locus almost explains all variation in Lp(a) levels. The linkage results for Lp(a) levels may hold some false positive result, possibly due to the non- normally distributed plasma levels for Lp(a), even after transformation.

In conclusion, we have to follow up 11 loci by genotyping additional markers to investigate whether the linkage remains and whether such linkage can be detected in all twin samples. Although, total cholesterol, HDL cholesterol, triglycerides and apoE levels are all highly heritable, no significant or suggestive linkage results have been found for these traits. This might indicate that the major gene loci are not present on these 10 chromosomes harbouring the majority of the candidate genes, or that the QTL effects on these traits are smaller that 25%.

When the whole genome scan is finished, we are able to estimate genome wide significance empirically[38], or evaluate the scan by locus counting[39].

Acknowledgements

This study was supported by grants to GP Vogler from the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976), to NG Martin and JB Whitfield from the National Heart Foundation of Australia, to NL Pedersen (AG04563 and AG10175) from NIA, to PE Slagboom (98.199) and DI Boomsma (86.083, 88.042 and 90.313) from the Netherlands Heart Foundation.

References

1. Boerwinkle E, Leffert CC, Lin J, Lackner C, Chiesa G, Hobbs HH. Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations. *J Clin Invest* 1992; 90: 52-60.
2. Neale MC, de Knijff P, Havekes LM, Boomsma DI. ApoE polymorphism accounts for only part of the genetic variation in quantitative ApoE levels. *Genet Epidemiol* 2000; 18: 331-340.
3. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet* 1988; 42: 104-112.
4. Vogler GP, McClearn GE, Snieder H, Boomsma DI, Palmer R *et al.* Genetics and behavioral medicine: risk factors for cardiovascular disease. *Behav Med* 1997; 22: 141-149.
5. Coon H, Leppert MF, Kronenberg F, Province MA, Myers RH *et al.* Evidence for a major gene accounting for mild elevation in LDL cholesterol: the NHLBI Family Heart Study. *Ann Hum Genet* 1999; 63 (Pt 5): 401-412.
6. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001; 69: 936-950.
7. Beekman M, Lakenberg N, Cherny SS, de Knijff P, Kluft CC *et al.* A powerful and rapid approach to human genome scanning using small quantities of genomic DNA. *Genet Res* 2001; 77: 129-134.
8. Sing CF, Davignon J. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation. *Am J Hum Genet* 1985; 37: 268-285.
9. Stengard JH, Clark AG, Weiss KM, Kardia S, Nickerson DA *et al.* Contributions of 18 additional DNA sequence variations in the gene encoding apolipoprotein e to explaining variation in quantitative measures of lipid metabolism. *Am J Hum Genet* 2002; 71: 501-517.
10. Risch N, Zhang H. Extreme discordant sib pairs for mapping quantitative trait loci in humans. *Science* 1995; 268: 1584-1589.
11. Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB *et al.* Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Res* 2002; 5: 87-97.
12. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; 17: 742-743.
13. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; 18: 499-502.
14. Rifai N, Warnick GR, McNamara JR, Belcher JD, Grinstead GF, Frantz ID. Measurement of low-density-lipoprotein cholesterol in serum: a status report. *Clin Chem* 1992; 38: 150-160.
15. Fulker DW, Cherny SS. An improved multipoint sib-pair analysis of quantitative traits. *Behav Genet* 1996; 26: 527-532.
16. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
17. Kruglyak L, Lander ES. A nonparametric approach for mapping quantitative trait loci. *Genetics* 1995; 139: 1421-1428.
18. Douglas JA, Boehnke M, Lange K. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 2000; 66: 1287-1297.
19. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; 58: 1347-1363.
20. Broman KW. Estimation of allele frequencies with data on sibships. *Genet Epidemiol* 2001; 20: 307-315.
21. Sham P. Statistics in human genetics. London: Arnold Publishers; 1998.
22. Purcell S, Cherny SS, Sham PC. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; 19: 149-150.
23. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11: 241-247.
24. Knoblauch H, Bauerfeind A, Krahenbuhl C, Daury A, Rohde K *et al.* Common haplotypes in five genes influence genetic variance of LDL and HDL cholesterol in the general population. *Hum Mol Genet* 2002; 11: 1477-1485.

25. Mattu RK, Needham EW, Morgan R, Rees A, Hackshaw AK *et al.* DNA variants at the LPL gene locus associate with angiographically defined severity of atherosclerosis and serum lipoprotein levels in a Welsh population. *Arterioscler Thromb* 1994; 14: 1090-1097.
26. Heizmann C, Kirchgessner T, Kwiterovich PO, Ladias JA, Derby C *et al.* DNA polymorphism haplotypes of the human lipoprotein lipase gene: possible association with high density lipoprotein levels. *Hum Genet* 1991; 86: 578-584.
27. Coon H, Eckfeldt JH, Leppert MF, Myers RH, Arnett DK *et al.* A genome-wide screen reveals evidence for a locus on chromosome 11 influencing variation in LDL cholesterol in the NHLBI Family Heart Study. *Hum Genet* 2002; 111: 263-269.
28. Reed DR, Nanthakumar E, North M, Bell C, Price RA. A genome-wide scan suggests a locus on chromosome 1q21-q23 contributes to normal variation in plasma cholesterol concentration. *J Mol Med* 2001; 79: 262-269.
29. Aouizerat BE, Allayee H, Cantor RM, Davis RC, Lanning CD *et al.* A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J Hum Genet* 1999; 65: 397-412.
30. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ *et al.* A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002; 30: 210-214.
31. Barton NH, Keightley PD. Understanding quantitative genetic variation. *Nat Rev Genet* 2002; 3: 11-21.
32. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 2002; 416: 326-330.
33. Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat* 1992; 1: 445-466.
34. Frikke-Schmidt R. Context-dependent and invariant associations between APOE genotype and levels of lipoproteins and risk of ischemic heart disease: a review. *Scand J Clin Lab Invest Suppl* 2000; 3-25.
35. Ishii H, Kim DH, Fujita T, Endo Y, Saeki S, Yamamoto TT. cDNA cloning of a new low-density lipoprotein receptor-related protein and mapping of its gene (LRP3) to chromosome bands 19q12-q13. 2. *Genomics* 1998; 51: 132-135.
36. Imperatore G, Knowler WC, Pettitt DJ, Kobes S, Fuller JH *et al.* A locus influencing total serum cholesterol on chromosome 19p: results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arterioscler Thromb Vasc Biol* 2000; 20: 2651-2656.
37. Ober C, Abney M, McPeek MS. The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 2001; 69: 1068-1079.
38. Hirschhorn JN, Lindgren CM, Daly MJ, Kirby A, Schaffner SF *et al.* Genomewide linkage analysis of stature in multiple populations reveals several regions with evidence of linkage to adult height. *Am J Hum Genet* 2001; 69: 106-116.
39. Wiltshire S, Cardon LR, McCarthy MI. Evaluating the results of genomewide linkage scans of complex traits by locus counting. *Am J Hum Genet* 2002; 71: 1175-1182.

5 | Two-locus linkage analysis applied to putative quantitative trait loci for Lipoprotein(a) levels

Marian Beekman, Bastiaan T. Heijmans, Nicholas G. Martin, John B. Whitfield, Nancy L. Pedersen, Ulf DeFaire, Harold Snieder, Nico Lakenberg, Peter de Knijff, Rune R. Frants, Gert Jan B. van Ommen, Cornelis Kluft, George P. Vogler, P. Eline Slagboom, Dorret I. Boomsma

Twin Research 2003; 6: 322-324

Abstract

Plasma levels of lipoprotein(a) (Lp(a)) are associated with cardiovascular risk[1] and were long believed to be influenced by the *LPA* locus on chromosome 6q27 only. However, a recent report of Broeckel *et al*[2] suggested the presence of a second quantitative trait locus on chromosome 1 influencing Lp(a) levels. Using a two-locus model, we found no evidence for an additional Lp(a) locus on chromosome 1 in a linkage study among 483 dizygotic twin pairs.

Introduction

Levels of Lp(a) are highly heritable. More than 90% of the variance is attributable to genetic factors[3,4]. Variation in the gene encoding apolipoprotein(a) (*LPA*) on chromosome 6q27 accounts for more than 85% of the variation in Lp(a) plasma levels[5-7]. Recently, Broeckel *et al*[2] reported a second chromosomal region linked with Lp(a) levels on chromosome 1 at 170cM from pter, which explained 16% of the variance in Lp(a), in a linkage study of 513 Western European families with 2 or more cases of premature myocardial infarction (on average 2.7 subjects per family). To test whether this putative quantitative trait locus (QTL) also affects Lp(a) levels in the general population, we performed a linkage study in four samples of dizygotic (DZ) and monozygotic (MZ) twin pairs recruited regardless of their health status.

Materials and Methods

Plasma levels of Lp(a) were obtained of an adolescent Dutch sample (90 DZ, 70 MZ, aged 13-22) and three adult samples from the Netherlands (117 DZ, 91 MZ, aged 34-61), Sweden (87 DZ, 69 MZ, aged 42-81) and Australia (651 DZ, 711 MZ, aged 31-80) recruited regardless of their health status[8]. In these samples, heritability estimates of Lp(a) levels ranged from 90 to 96%. Genotypes for the linkage study were obtained of DZ pairs of whom DNA was available, and not of MZ twins, since these are genetically identical, sharing two alleles identical-by-descent (IBD) at all loci. Markers from the Marshfield Screening Set 8 were genotyped on chromosome 1 and 6 with an average spacing of 18 cM[9]. To adjust for non-normality, Lp(a) levels were transformed by the natural logarithm. Variance components linkage analyses were carried out, including data from the 483 genotyped DZ twins as well as the phenotypical data from the 941 untyped MZ twin pairs. The inclusion of MZ twins provides a more accurate estimate of non-shared environmental effects (MZ correlations are <1), so that an upper limit for the estimation of genetic effects is obtained and overestimation is reduced. IBD status for the DZ pairs was estimated separately for each of the four samples with Genehunter 2.1[10] using population specific allele frequencies. All analyses were performed using Mx 1.52d[11] in a 4-sample simultaneous analysis. Mean Lp(a) levels, background genetic and non-shared environmental effects were estimated for each sample separately. The total variance was modelled as A+E+Q, where A represents additive genetic background factors, E non-shared environmental factors and Q the QTL effect. The covariance equals A+Q for MZ twins and $½A+\hat{\pi}Q$ for DZ twins, where $\hat{\pi}$ is estimated as ½P(IBD=1)+P(IBD=2). To test for linkage, the fit of the AE model was compared to the fit of the AEQ model. For the two-locus models, the equations for the total variance, covariance for MZ twins and covariance for DZ twins were $A+E+Q_1+Q_6$, $A+Q_1+Q_6$ and $½A+\hat{\pi}_1Q_1+\hat{\pi}_6Q_6$, respectively, where Q_1 and Q_6 refer to the putative QTLs on chromosomes 1 and 6, respectively. The proportion of the variance explained by the QTL was constrained to be equal over the samples. The power to replicate the putative Lp(a) QTL on chromosome 1 (QTL effect, 0.16; residual shared variance, 0.74) in 483 sib pairs at a significance level of 0.00074 (LOD-score of 2.2 constituting suggestive linkage according to Lander-Kruglyak critera) and assuming incomplete marker informativeness (theta=0.10) is 0.51, as calculated using the Genetic Power Calculator (http://statgen.iop.kcl.ac.uk/gpc/)[12].

Results

As expected, we obtained highly significant evidence for linkage at the *LPA* locus on chromosome 6 (maximum LOD-score (MLS) = 9.8, Figure 1; MLS of 1.7, 4.4, 1.8 and 5.1, in the adolescent Dutch and adult Dutch, Swedish and Australian samples, respectively). This QTL explained 82% of the total variance in Lp(a) levels (Table 1). Next, chromosome 1 was analysed and a MLS of 1.6 was obtained at 251 cM from pter (Figure 1; MLS of 0.0, 0.7, 2.2 and 1.8, in the adolescent Dutch and adult Dutch, Swedish and Australian samples, respectively).

This QTL explained 44% of the total variance in Lp(a) levels (Table 1). The two putative QTLs on chromosome 6 and 1 together explained 126% of the variation in Lp(a) levels, which likely results from overestimation of effect sizes, a common phenomenon in genome-wide linkage analyses[13]. We, therefore, analysed the effect of both QTLs simultaneously in a two-locus model (see figure 1 for details). In this two-locus model, evidence for linkage on chromosome 1 completely disappeared (LOD=0; Figure 1). Moreover, the chromosome 1 locus now explained none of the variation in Lp(a) plasma levels, while the estimates for the *LPA* locus remained unaffected (Table 1).

Figure 1.
Linkage analysis of chromosome 6 and 1 with plasma levels of Lp(a) in four twin samples simultaneously. ▬ *shows the result of the initial analysis of the chromosomes;* ····· *shows the result of the two-locus analysis.*

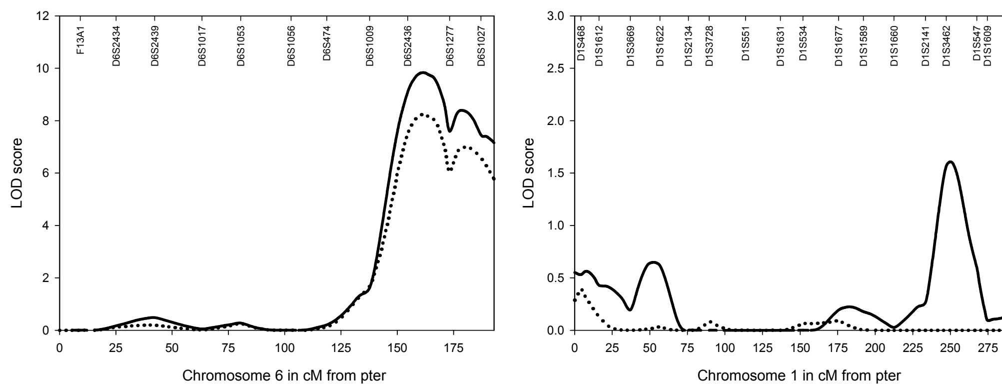


Table 1.
Effect size of putative Lp(a) QTLs at chromosome 6 and 1 at the position of the maximum LOD scores, in separate and simultaneous analyses.

	Maximum LOD score (position in cM from pter)		**Proportion of the variance attributable to the QTL at chromosome 6**	**Proportion of the variance attributable to the QTL at chromosome 1**
Chromosome analysed	Chr 6	Chr 1	(95% C.I.)	(95% C.I.)
6	9.8 (163)	-	0.82 (0.66-0.91)	-
1	-	1.6 (251)	-	0.44 (0.13-0.68)
6 and 1	8.2 (163)	0.0 (251)	0.82 (0.65-0.91)	0.00 (0.00-0.14)

Conclusion

In a recent study, Broeckel *et al*[2] reported a second QTL on chromosome 1 influencing Lp(a) levels in premature MI patients. Although the power to detect linkage at chromosome 1 in the four twin samples is low (0.51), did find suggestive linkage (MLS=2.2) in one of the samples and a MLS of 1.6 in the simultaneous analysis of chromosome 1 with Lp(a) levels. Nevertheless our simultaneous analysis of the two loci on chromosome 6 and 1, the linkage on chromosome 1 disappeared. Our analysis differs from that of Broeckel *et al* in the fact that the total genetic effect of the two loci is restricted to 100%. Our study in twin pair samples provides no support for a Lp(a) QTL on chromosome 1 acting in the general population.

Acknowledgements

This study was supported by grants to GP Vogler of the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976), to NG Martin and JB Whitfield of the National Heart Foundation of Australia, to NL Pedersen (AG04563) of NIA, to PE Slagboom (98.199) and DI Boomsma (86.083, 88.042 and 90.313) of the Netherlands Heart Foundation.

References

1. Danesh J, Collins R, Peto R. Lipoprotein(a) and coronary heart disease. Meta-analysis of prospective studies. *Circulation* 2000; 102: 1082-1085.
2. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ *et al.* A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002; 30: 210-214.
3. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.
4. Rainwater DL. Lp(a) concentrations are related to plasma lipid concentrations. *Atherosclerosis* 1996; 127: 13-18.
5. Boerwinkle E, Leffert CC, Lin J, Lackner C, Chiesa G, Hobbs HH. Apolipoprotein(a) gene accounts for greater than 90% of the variation in plasma lipoprotein(a) concentrations. *J Clin Invest* 1992; 90: 52-60.
6. Boomsma DI, Kaptein A, Kempen HJ, Gevers LJ, Princen HM. Lipoprotein(a): relation to other risk factors and genetic heritability. Results from a Dutch parent-twin study. *Atherosclerosis* 1993; 99: 23-33.
7. DeMeester CA, Bu X, Gray RJ, Lusis AJ, Rotter JI. Genetic variation in lipoprotein (a) levels in families enriched for coronary artery disease is determined almost entirely by the apolipoprotein (a) gene locus. *Am J Hum Genet* 1995; 56: 287-293.
8. Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB *et al.* Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Res* 2002; 5: 87-97.
9. Beekman M, Lakenberg N, Cherny SS, de Knijff P, Kluft CC *et al.* A powerful and rapid approach to human genome scanning using small quantities of genomic DNA. *Genet Res* 2001; 77: 129-134.
10. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; 58: 1347-1363.
11. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
12. Sham PC, Cherny SS, Purcell S, Hewitt JK. Power of linkage versus association analysis of quantitative traits, by use of variance-components models, for sibship data. *Am J Hum Genet* 2000; 66: 1616-1630.
13. Göring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 2001; 69: 1357-1369.

6 | Evidence for a QTL on chromosome 19 influencing LDL cholesterol levels in the general population

*Marian Beekman, *Bastiaan T. Heijmans, Nicholas G. Martin, John B. Whitfield, Nancy L. Pedersen, Ulf DeFaire, Harold Snieder, Nico Lakenberg, H. Eka D. Suchiman, Peter de Knijff, Rune R. Frants, Gert Jan B. van Ommen, Cornelis Kluft, George P. Vogler, Dorret I. Boomsma, P. Eline Slagboom.

*These authors contributed equally to this work.

European Journal of Human Genetics 2003; 11: 845-850

Abstract

Plasma levels of lipids and apolipoproteins are among the major quantitative risk factors for CVD and are well-established intermediate traits that may be more accessible to genetic dissection than clinical CVD endpoints. Chromosome 19 harbours multiple genes that have been suggested to play a role in lipid metabolism and previous studies indicated the presence of a quantitative trait locus for cholesterol levels in genetic isolates. We performed a linkage study in adolescent Dutch twins and adult Dutch, Swedish and Australian twins totalling 493 dizygotic twin pairs. In the three adult twin samples, we found consistent evidence for linkage of chromosome 19 with LDL cholesterol levels (maximum LOD scores of 4.5, 1.7 and 2.1 in the Dutch, Swedish and Australian sample, respectively). The simultaneous analysis of the samples increased the maximum LOD score to 5.7 at 60 cM pter. Bivariate analyses indicated that the putative LDL-C QTL also contributed to the variance in ApoB levels, consistent with the high genetic correlation between these phenotypes. Our study provides strong evidence for the presence of a QTL on chromosome 19 with a major effect on LDL-C plasma levels in out-bred Caucasian populations.

Introduction

Dissecting the genetic basis of cardiovascular disease (CVD) is complicated by the etiologic heterogeneity of seemingly undistinguishable clinical endpoints. Major quantitative risk factors for CVD such as cholesterol levels may be more accessible to genetic dissection[1]. Also, individuals do not need to be classified as affected or unaffected for these quantitative variables, which often leads to ambiguous or uncertain phenotype assignment. Genes described to date that play a role in lipid metabolism have mainly been identified in familial lipid disorders, many of which are monogenic diseases. For example, defects in the genes encoding the low-density-lipoprotein-receptor (*LDLR,* 19p13.2)[2] and apolipoprotein B (*APOB*, 2p24)[3] are found to be responsible for the most common forms of familial hypercholesterolemia, and mutations in the *ABC1* gene (9q31.1), encoding the cholesterol efflux regulatory protein, induce familial HDL cholesterol deficiency (Tangier disease)[4]. Although carriers of such mutations may develop severe disorders of lipid metabolism, these mutations explain only a minor proportion of the variation in plasma levels of lipids and apolipoproteins at the level of the general population[5]. The results of twin studies showing that 50-80% of the population variation in levels of lipids and apolipoproteins is attributable to genetic factors[6,7] thus imply that the majority of the genes determining these levels are still to be discovered.

Chromosome 19 harbours several genes that have been suggested to play a role in lipid metabolism including the *LDLR* (19p13.2)[2], the apolipoprotein E gene (*APOE*, 19q13.2)[8] and other genes from the *APOE/C1/C4/C2* cluster[9,10], the insulin-receptor gene (*INSR*, 19p13.3)[11], the hormone-sensitive-lipase gene (*LIPE*, 19q13.1)[12],and the LDLR-related-protein-type3 gene (*LRP3*, 19q12)[13]. Compatible with this wealth of candidate genes, evidence for linkage with plasma levels of total cholesterol was found in Pima Indians and with LDL cholesterol levels in Hutterites[14,15]. In contrast to the findings in these genetic isolates, however, no significant linkage with LDL-C levels on chromosome 19 was found in genome scans among out-bred Caucasians[16-20]. To gain insight in the effect of genetic variation on chromosome 19 on lipid metabolism in the general, Caucasian population, we performed a linkage study in four twin samples originating from The Netherlands, Sweden and Australia totalling 493 dizygotic twin pairs.

Subjects and methods

Subjects

We studied samples of adolescent Dutch twins and adult Dutch (sample on cardiovascular risk factors), Swedish (Swedish Adoption/Twin Study of Aging) and Australian twins (Semi-Structured Assessment for the Genetics of Alcoholism). The recruitment of the twins and the measurements of lipid and apolipoprotein levels in the different samples are described elsewhere[7]. In this study we used 83 dizygotic (DZ) young Dutch twin pairs (aged 13-22), 117 DZ adult Dutch twin pairs (aged 34-62), 44 DZ Swedish twin pairs (aged 42-81) and 249 DZ Australian twin pairs (aged 31-80). All relationships were confirmed with the Graphical Representation of Relationship software[21] using more than 100 short tandem repeats. Total cholesterol, HDL cholesterol, apolipoprotein B, apolipoprotein A1, triglycerides and apolipoprotein E levels were assessed in plasma. Concentrations of low density lipoprotein (LDL) cholesterol were calculated according to Friedewald[22]. LDL cholesterol level was considered missing, when the triglyceride concentration did exceed 4.52 mmol/L, which was the case for 1 individual of the middle-aged Dutch sample, for 82 individuals of the Australian sample and for 11 individuals of the Swedish sample[23].

Genotyping

In the four twin pair samples, 12 short tandem repeats with an average inter-marker distance of 8 centiMorgans (cM) were genotyped (D19S247, D19S1034, D19S394, D19S714, D19S49, D19S433, D19S47, APOC2, D19S246, D19S180, D19S210 and D19S254). In the two Dutch samples 4 additional markers were genotyped (D19S391, D19S865, D19S420, D19S178), resulting in an average spacing of 6 cM in the Dutch. The average heterozygosity for these markers was estimated at 0.78 and the Marshfield genetic map (http://research.marshfieldclinic.org/genetics/) was used.

The Cy5-labeled PCR products were electrophoretically separated on an automated-fluorescence DNA sequencer, ALF*express* (Amersham Pharmacia Biotech). Analysis and assignment of the marker alleles were performed with Fragment Analyser 1.02 (Amersham Pharmacia Biotech). To reduce genotyping errors, one known genotype was present on each gel, 5% of the genotypings were repeated and two independent individuals performed the allele calling. SIBMED[24] was used to identify unlikely double recombinants, the occurrence of which may be due to genotyping errors. After running SIBMED and checking the raw genotyping data, approximately 0.2% of the total genotypings appeared to be erroneous. Dependent on the error, these genotypes were changed in the right genotype or were set to missing.

Statistical analysis

Allele frequencies were estimated separately for the twin samples using marker data for all individuals[25]. Plasma levels of triglycerides and apoE showed a skewed distribution and these values were therefore transformed by natural logarithm prior to analysis. The full distribution of multipoint identity-by-descent (IBD) sharing probabilities was estimated every centiMorgan across chromosome 19 using Genehunter 2.1[26]. Linkage analysis of quantitative traits was performed with variance components analysis using structural equation modelling with maximum likelihood implemented in the software Mx 1.52d[27]. The weighted likelihood approach, which makes use of the full distribution of IBD-probabilities, with adjustments for age and sex was used[28]. In a 4-sample simultaneous analysis, IBD status for the DZ pairs was estimated separately for each of the four samples in Genehunter 2.1[26] using population specific allele frequencies. Mean LDL-C levels, background genetic and non-shared environmental effects were estimated for each sample separately. The absolute QTL effects were constrained to be equal over the different samples. Heterogeneity between the effects of the QTL in the different twin samples was tested by comparing the model estimating all parameters for the populations separately with the model in which the QTL effect was constrained to be equal. Bivariate analyses[29] were performed for correlated plasma levels of LDL-C and apoB.

Results

The characteristics of the adolescent Dutch and adult Dutch, Swedish and Australian dizygotic twins totalling 493 pairs are shown in table 1.

The four twin samples were analysed using variance components analyses, which revealed a consistent indication for linkage with LDL cholesterol (LDL-C) levels (Table 2). The maximum LOD scores (MLS) were 4.5, 1.7 and 2.1 in adult Dutch, Swedish and Australian twins, respectively; no linkage was observed in the adolescent Dutch twins. Lower levels of linkage were also observed for the LDL-C levels correlated phenotypes total cholesterol and apoB in the adult populations (Table 2).

As shown in figure 1, linkage in the adult twins samples was observed in the same chromosomal region suggesting that the same QTL may be involved. This was reinforced by the fact that the QTL effect was not significantly different in the three adult populations at the positions where the MLS were observed (p=0.50, 34 cM pter; p=0.13, 60 cM pter; p=0.06, 63 cM pter), whereas the QTL effect was significantly different in the adolescent Dutch population (e.g. p=0.0001, 60 cM pter). Therefore, the adult populations were combined in one linkage analysis including 410 dizygotic twin pairs (figure 1). This analysis increased the MLS for LDL-C levels to 5.7 at 60 cM pter. Of the twin pairs that individually contributed more than 0.20 to the MLS, 62% (18/29) were Dutch, 17% (5/29) were Swedish and 21% (6/29) were Australian.

Table 1.
Characteristics of adolescent Dutch and adult Dutch, Swedish and Australian dizygotic twin samples.

Phenotype	Adolescent Netherlands		Adult Netherlands		Adult Sweden		Adult Australia	
	(n=166)		(n=234)		(n=88)		(n=498)	
men, %	49.4		48.7		59.1		35.7	
Age, years – mean (range)	17	(13-22)	44	(34-59)	65	(42-81)	44	(31-80)
Body Mass Index, kg/m^2 – mean (SD)	20.28	(2.21)	24.64	(3.06)	25.21	(3.07)	25.44	(5.19)
LDL-C, mmol/L – mean (SD)	2.56	(0.65)	3.63	(0.99)	4.26	(1.07)	3.37	(1.07)
ApoB, g/L – mean (SD)	0.79	(0.17)	1.21	(0.35)	1.04	(0.21)	0.97	(0.28)
Total cholesterol, mmol/L – mean (SD)	4.15	(0.70)	5.39	(1.06)	6.50	(1.11)	5.64	(1.20)
HDL-C, mmol/L – mean (SD)	1.28	(0.26)	1.22	(0.38)	1.47	(0.34)	1.46	(0.41)
ApoA1, g/L – mean (SD)	1.38	(0.20)	1.67	(0.43)	1.39	(0.30)	1.44	(0.27)
ApoA2, g/L – mean (SD)	0.58	(0.08)	-	-	-	-	0.36	(0.06)
Triglycerides, mmol/L – mean (SD)	0.68	(0.30)	1.22	(0.71)	1.62	(0.87)	1.90	(1.69)
Lipoprotein (a), mg/L – mean (SD)	14.2	(21.2)	18.4	(26.6)	175.7	(158.1)	214.8	(267.3)
ApoE, mg/dL – mean (SD)	6.59	(2.39)	2.54	(1.02)	-	-	4.18	(1.88)

Figure 1.
Linkage of LDL cholesterol levels with chromosome 19 in adult Dutch (▬ · · ▬ · ·), Swedish (· · · ·) and Australian (▬ ▬ ▬) twins in separate analyses and a combined analysis (▬▬). The upper x-axe shows the location of candidate genes.

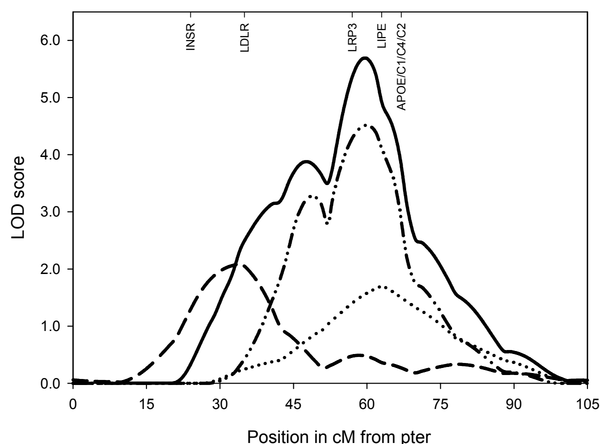


Table 2.
Maximum LOD scores observed at chromosome 19 for lipid and apolipoprotein levels in adolescent Dutch and adult Dutch, Swedish and Australian DZ twins.

Phenotype	Position from pter	Max LOD score
LDL cholesterol		
Adolescent Netherlands	98	0.5
Adult Netherlands	60	4.5
Adult Sweden	63	1.7
Adult Australia	34	2.1
Apolipoprotein B		
Adolescent Netherlands	98	0.3
Adult Netherlands	48	1.0
Adult Sweden	70	1.6
Adult Australia	80	0.5
Total cholesterol		
Adolescent Netherlands	53	0.2
Adult Netherlands	60	2.3
Adult Sweden	63	1.3
Adult Australia	30	1.7
HDL cholesterol		
Adolescent Netherlands	67	0.1
Adult Netherlands	32	0.5
Adult Sweden	34	0.2
Adult Australia	57	1.2
Apolipoprotein A1		
Adolescent Netherlands	0	0.1
Adult Netherlands	34	0.2
Adult Sweden	36	1.0
Adult Australia	51	1.5
Triglycerides		
Adolescent Netherlands	52	1.2
Adult Netherlands	45	0.6
Adult Sweden	63	0.1
Adult Australia	63	0.2
Apolipoprotein E		
Adolescent Netherlands	69	1.0
Adult Netherlands	63	2.8
Adult Sweden	-	-
Adult Australia	66	0.1

Table 3.
Proportion of the total variance of LDL-C and ApoB explained by the putative QTL on the basis of bivariate analyses at the position of the maximum LOD score for LDL-C in the adult twin samples.

Twin sample	Phenotypic Correlation LDL-C/ApoB	Genetic Correlation LDL-C/ApoB	Position (cM)		[a]V_{QTL}/[b]V_{total}	Proportion of variance attributable to QTL (95% CI)	
Adult Netherlands	0.89	0.92	60	LDL-C	0.80/0.88	0.92	(0.71-0.98)
				ApoB	0.07/0.11	0.63	(0.35-0.83)
Adult Sweden	0.79	0.79	63	LDL-C	0.94/1.15	0.84	(0.11-0.98)
				ApoB	0.02/0.04	0.68	(0.00-0.85)
Adult Australia	0.83	0.92	34	LDL-C	0.38/1.01	0.38	(0.05-0.67)
				ApoB	0.01/0.07	0.12	(0.00-0.38)

[a] Variance explained by the QTL
[b] Total variance

Table 4.
Overview of studies indicating linkage of plasma levels of LDL cholesterol and related phenotypes on chromosome 19.

Phenotype	Population	Location on chromosome 19 (cM)	Genome-wide p-value (Gwp) or maximum LOD score (MLS)	Reference
LDL cholesterol	Dutch	60	MLS = 4.5	This study
	Sweden	63	MLS = 1.7	This study
	Australian	34	MLS = 2.1	This study
	Above combined	60	MLS = 5.7	This study
LDL cholesterol	Hutterites	52	Gwp = 0.035	[15]
Total cholesterol	Pima Indians	21	MLS = 3.9	[14]
LDL1 cholesterol	Mexican Americans	38	MLS = 2.4	[33]
LDL2 cholesterol	Mexican Americans	62	MLS = 1.9	[33]
Total cholesterol	Non hispanic whites	27	MLS = 1.1	[34]

LDL-C and apoB levels are highly correlated phenotypes, both phenotypically (≥0.79) and genetically (≥0.79) (Table 3) indicating that genetic variation influencing LDL-C levels also influences apoB levels. Nevertheless, the MLSs were considerably higher for LDL-C levels than for apoB levels. To gain insight in the effect of the LDL-C QTL on apoB levels, we performed a bivariate linkage analysis of the two phenotypes at the position of MLS in the adult samples. Although the confidence intervals were wide and the estimates of the QTL effect not always significant (Table 3), this analysis showed that the putative QTL explained a considerable proportion of the variance in apoB levels in addition to that of LDL-C levels.

Discussion

We have studied chromosome 19 for linkage with intermediate lipid phenotypes of cardiovascular disease in dizygotic twin pairs from the general population of The Netherlands, Sweden and Australia. In all the adult twin samples, we found evidence for linkage of chromosome 19 with LDL-C levels with MLS ranging from 1.7 to 4.5. In spite of a large heritability of LDL cholesterol levels, which was previously estimated in the adolescent and adult twin samples between 0.60 and 0.85[7], no indication for linkage was observed in the adolescent twin sample. This may be due to partly different effects of genes on lipid levels at different ages as suggested by Snieder *et al.*[30] or to different gene-environment interaction on lipid levels at different ages as suggested by Zerba *et al.*[31]. The size of the QTL effect on LDL-C levels was not significantly different in the adult samples and a simultaneous analysis of these adult samples increased the LOD score to 5.7 at 60 cM from pter. This constitutes significant linkage according to the Lander-Kruglyak criteria[32]. A bivariate analysis indicates that the QTL also influences apoB levels as is consistent with the fact that the majority of apoB protein circulates as a constituent of LDL particles. Our linkage study thus emphasizes the relevance of genetic variation on chromosome 19 for cardiovascular risk in the general population.
Previous studies reported evidence for linkage of chromosome 19 with total cholesterol levels (MLS= 3.89) and LDL-C levels (Genome-wide p-value= 0.035) in Pima Indians[14] and Hutterites[15], respectively. Our study extends these findings in genetic isolates to the general, Caucasian population. A further inspection of other genome scans in Caucasians from the general population (Table 4) shows that, although the reported LOD scores for chromosome 19 are not significant, they are compatible with our findings. Linkage with LDL_1-C levels in the San Antonio Heart Study (LOD=2.26) and with LDL_2-C levels (LOD=1.86)[33], and linkage with total cholesterol levels in the Rochester Family Heart Study (LOD= 1.14)[34]. No suggestion for linkage

with total cholesterol or LDL-C levels, however, was found in selected samples of myocardial infarction patients, type 2 diabetes patients and patients with familial combined hyperlipidemia[16-20]. The LDL-C QTL on chromosome 19 thus constitutes one of the most replicated result from linkage studies, virtually ruling out the possibility that it is a false positive observation.

The region in which we and other groups (Figure 1, Table 4) found evidence for linkage is broad as is a general characteristic of results from twin and sib pair studies. QTL mapping in *Drosophila melanogaster* and *Saccharomyces cerevisiae* demonstrated that linkage results at a broad chromosomal region may be caused by several polymorphisms with a relatively small effect, which could be interpreted as one gene with a major effect[35,36]. Possibly, several loci on chromosome 19 are influencing LDL cholesterol levels and perhaps different loci play major roles in Australian and European LDL cholesterol levels, which might explain the different locations of the MLS. However, the overall effect size of these loci is the not significantly different in Dutch, Swedish and Australians. The main candidates for underlying the LDL-C QTL on chromosome 19 (Table 4) are the *ApoE/C1/C4/C2* gene cluster and the *LDLR* and *LRP3* gene loci. Apolipoprotein E is the major constituent of chylomicrons, VLDL and IDL particles and serves as ligand for the LDLR. When VLDL and IDL are not efficiently removed from the circulation by the LDLR, LDL particles will accumulate. The LDLR is also known to play an important role in the clearance of apoB-carrying-lipoproteins by the liver[2,37]. Apolipoprotein C1 is a constituent of VLDL and HDL particles. It inhibits the lipoprotein lipase (LPL)-mediated hydrolysis of the triglycerides from VLDL, which leads to lower levels of LDL-C[38]. ApoC2 has the opposite effect on LDL-C levels, since it is an activator of LPL[39]. Although the function of apoC4 is unknown, there are some indications that it might play a role in lipid metabolism[10,40]. The LRP3 is a family member of the LDLR. It is also expressed in the liver and it binds apoE-containing particles. The function of the LRP3 remains unclear, but as a family member of LDLR a role in lipid metabolism is assumed[13]. *INSR* and *LIPE* are candidates for influencing triglyceride levels, but not for LDL-C levels[11,12]. Near the LDLR locus, the ATHS locus has been mapped involved in atherogenic lipoprotein phenotype pattern B, which is characterized by small, dense LDL particles, increased levels of triglycerides and decreased levels of HDL-C[41]. Since we have no indication for linkage with levels of triglycerides and HDL-C, we assume that the ATHS locus is a different locus than our LDL-C QTL.

In conclusion, we report strong and significant evidence for the presence of genetic variation at chromosome 19 with a major effect on LDL-C plasma levels in out-bred Caucasian populations. Since high LDL-C level is a major risk factor for cardiovascular disease, this finding may significantly contribute to disentangling the complex architecture of cardiovascular disease. Candidate gene screening and the recently developed method for the combined analysis of linkage and association[42,43] may reveal the genetic variation underlying the QTL.

Acknowledgements

This study was supported by grants to GP Vogler from the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976), to NG Martin and JB Whitfield from the National Heart Foundation of Australia, to NL Pedersen (AG04563 and AG10175) from NIA, to PE Slagboom (98.199) and DI Boomsma (86.083, 88.042 and 90.313) from the Netherlands Heart Foundation.

References

1. Vogler GP, McClearn GE, Snieder H, Boomsma DI, Palmer R *et al.* Genetics and behavioral medicine: risk factors for cardiovascular disease. *Behav Med* 1997; 22: 141-149.
2. Hobbs HH, Brown MS, Goldstein JL. Molecular genetics of the LDL receptor gene in familial hypercholesterolemia. *Hum Mutat* 1992; 1: 445-466.
3. Hansen PS. Familial defective apolipoprotein B-100. *Dan Med Bull* 1998; 45: 370-382.
4. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K *et al.* Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 1999; 22: 336-345.
5. Williams RR, Hopkins PN, Hunt SC, Wu LL, Hasstedt SJ *et al.* Population-based frequency of dyslipidemia syndromes in coronary-prone families in Utah. *Arch Intern Med* 1990; 150: 582-588.
6. Hamsten A, Iselius L, Dahlen G, de Faire U. Genetic and cultural inheritance of serum lipids, low and high density lipoprotein cholesterol and serum apolipoproteins A-I, A-II and B. *Atherosclerosis* 1986; 60: 199-208.
7. Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB *et al.* Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Res* 2002; 5: 87-97.
8. Friedlander Y, Leitersdorf E. Influence of apolipoprotein E genotypes on plasma lipid and lipoprotein concentrations: results from a segregation analysis in pedigrees with molecularly defined familial hypercholesterolemia. *Genet Epidemiol* 1996; 13: 159-177.
9. Shachter NS. Apolipoproteins C-I and C-III as important modulators of lipoprotein metabolism. *Curr Opin Lipidol* 2001; 12: 297-304.
10. Kamboh MI, Aston CE, Hamman RF. DNA sequence variation in human apolipoprotein C4 gene and its effect on plasma lipid profile. *Atherosclerosis* 2000; 152: 193-201.
11. Walton C, Lees B, Crook D, Godsland IF, Stevenson JC. Relationships between insulin metabolism, serum lipid profile, body fat distribution and blood pressure in healthy men. *Atherosclerosis* 1995; 118: 35-43.
12. Raclot T, Holm C, Langin D. A role for hormone-sensitive lipase in the selective mobilization of adipose tissue fatty acids. *Biochim Biophys Acta* 2001; 1532: 88-96.
13. Ishii H, Kim DH, Fujita T, Endo Y, Saeki S, Yamamoto TT. cDNA cloning of a new low-density lipoprotein receptor-related protein and mapping of its gene (LRP3) to chromosome bands 19q12-q13. 2. *Genomics* 1998; 51: 132-135.
14. Imperatore G, Knowler WC, Pettitt DJ, Kobes S, Fuller JH *et al.* A locus influencing total serum cholesterol on chromosome 19p: results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arterioscler Thromb Vasc Biol* 2000; 20: 2651-2656.
15. Ober C, Abney M, McPeek MS. The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 2001; 69: 1068-1079.
16. Reed DR, Nanthakumar E, North M, Bell C, Price RA. A genome-wide scan suggests a locus on chromosome 1q21-q23 contributes to normal variation in plasma cholesterol concentration. *J Mol Med* 2001; 79: 262-269.
17. Aouizerat BE, Allayee H, Cantor RM, Davis RC, Lanning CD *et al.* A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J Hum Genet* 1999; 65: 397-412.
18. Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F *et al.* A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet* 2001; 10: 2751-2765.
19. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ *et al.* A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002; 30: 210-214.
20. Elbein SC, Hasstedt SJ. Quantitative trait linkage analysis of lipid-related traits in familial type 2 diabetes: evidence for linkage of triglyceride levels to chromosome 19q. *Diabetes* 2002; 51: 528-535.
21. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; 17: 742-743.
22. Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972; 18: 499-502.
23. Rifai N, Warnick GR, McNamara JR, Belcher JD, Grinstead GF, Frantz ID. Measurement of low-density-lipoprotein cholesterol in serum: a status report. *Clin Chem* 1992; 38: 150-160.
24. Douglas JA, Boehnke M, Lange K. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 2000; 66: 1287-1297.
25. Broman KW. Estimation of allele frequencies with data on sibships. *Genet Epidemiol* 2001; 20: 307-315.
26. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; 58: 1347-1363.
27. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
28. Neale MC. QTL Mapping with sib pairs: The flexibility of Mx, In: Spector TD, Snieder H, MacGregor AJ, editors. Advances in twin and sib pair analysis. London: Greenwich Medical Media; 2000.
29. Vogler GP, Tang W, Nelson TL, Hofer SM, Grant JD *et al.* A multivariate model for the analysis of sibship covariance structure using marker information and multiple quantitative traits. *Genet Epidemiol* 1997; 14: 921-926.

30. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.
31. Zerba KE, Ferrell RE, Sing CF. Complex adaptive systems and human health: the influence of common genotypes of the apolipoprotein E (ApoE) gene polymorphism and age on the relational order within a field of lipid metabolism traits. *Hum Genet* 2000; 107: 466-475.
32. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11: 241-247.
33. Rainwater DL, Almasy L, Blangero J, Cole SA, VandeBerg JL *et al.* A Genome Search Identifies Major Quantitative Trait Loci on Human Chromosomes 3 and 4 That Influence Cholesterol Concentrations in Small LDL Particles. *Arterioscler Thromb Vasc Biol* 1999; 19: 777-783.
34. Klos KL, Kardia SL, Ferrell RE, Turner ST, Boerwinkle E, Sing CF. Genome-wide linkage analysis reveals evidence of multiple regions that influence variation in plasma lipid and apolipoprotein levels associated with risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2001; 21: 971-978.
35. Barton NH, Keightley PD. Understanding quantitative genetic variation. *Nat Rev Genet* 2002; 3: 11-21.
36. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 2002; 416: 326-330.
37. Ludwig EH, Hopkins PN, Allen A, Wu LL, Williams RR *et al.* Association of genetic variations in apolipoprotein B with hypercholesterolemia, coronary artery disease, and receptor binding of low density lipoproteins. *J Lipid Res* 1997; 38: 1361-1373.
38. Pentikainen MO, Oksjoki R, Oorni K, Kovanen PT. Lipoprotein lipase in the arterial wall: linking LDL to the arterial extracellular matrix and much more. *Arterioscler Thromb Vasc Biol* 2002; 22: 211-217.
39. Shen Y, Lookene A, Nilsson S, Olivecrona G. Functional analyses of human apolipoprotein CII by site-directed mutagenesis: identification of residues important for activation of lipoprotein lipase. *J Biol Chem* 2002; 277: 4334-4342.
40. de Graaf J, Hoffer MJ, Stuyt PM, Frants RR, Stalenhoef AF. Familial chylomicronemia caused by a novel type of mutation in the APOE-CI-CIV-CII gene cluster encompassing both the APOCII gene and the first APOCIV gene mutation: APOCII-CIV(Nijmegen). *Biochem Biophys Res Commun* 2000; 273: 1084-1087.
41. Nishina PM, Johnson JP, Naggert JK, Krauss RM. Linkage of atherogenic lipoprotein phenotype to the low density lipoprotein receptor locus on the short arm of chromosome 19. *Proc Natl Acad Sci USA* 1992; 89: 708-712.
42. Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 1999; 64: 259-267.
43. Posthuma D., Geus E.J.C., Boomsma D.I., and Neal, M. C. Combined linkage and association in Mx. *Behav Genet* 2004; 34: 153-174.

7 | Combined association and linkage analysis applied to the *APOE* locus

Marian Beekman, Daniëlle Posthuma, Bastiaan T. Heijmans, Nico Lakenberg, H. Eka D. Suchiman, Harold Snieder, Peter de Knijff, Rune R. Frants, Gert Jan B. van Ommen, Cornelis Kluft, George P. Vogler, P. Eline Slagboom, Dorret I. Boomsma.

Genetic Epidemiology 2004; in press

Abstract

Combined association and linkage analysis is a powerful tool for pinpointing functional QTLs responsible for regions of significant linkage identified in genome wide scans. We applied this technique to apoE plasma levels and the *APOEε2/ε3/ε4* polymorphism in two Dutch twin cohorts of different age range. Across chromosome 19, short tandem repeats and the *APOEε2/ε3/ε4* polymorphism were genotyped in adolescent (aged 13-22 years) and adult (aged 34-62 years) Dutch twins. In both samples, evidence for indicative linkage with plasma apoE levels was found (MLS=0.8, MLS=2.5, respectively) at 19q13.32. These linkage regions included the *APOE* locus. As expected, the *APOEε2/ε3/ε4* polymorphism was strongly associated with apoE plasma levels in both samples. An extension of the between/within families association test developed by Fulker *et al.*[1] showed that these associations were not due to population stratification. The combined association and linkage analyses revealed that the association of the *APOEε2/ε3/ε4* polymorphism with apoE plasma levels completely explained the linkage in the adolescent twins and partly in the adult twins.

Introduction

Successful genome scans for quantitative traits in sibling pairs at best yield broad chromosomal regions, in which the loci influencing the quantitative trait (QTLs) need to be identified. These regions can hardly be narrowed by typing additional markers, since sib pairs regularly share large haplotypes of tens of centiMorgans[2]. Testing polymorphisms for association with the quantitative trait is therefore essential for gene identification in positive linkage regions. Combined association and linkage analysis can give insight in which variant in a set of polymorphisms, located in a linkage region, explains the linkage result by association with the trait. By comparing allelic effects both within and between sibling pairs, this method also provides a direct test for the presence of population stratification[1,3].

We applied the combined association and linkage approach to apoE plasma levels and the *APOEε2/ε3/ε4* polymorphism in adolescent (aged 13-22 years) and adult (aged 34-62 years) Dutch twin pairs. Investigation of cohorts of different age allows accounting for the finding that different genetic variation may influence lipid metabolism at different ages[4]. The *APOEε2/ε3/ε4* polymorphism was previously found to explain between 9 to 20% of the total variance in apoE levels in the population at large[5-8]. However, more than 80% of the total variance in apoE levels is determined by genetic factors[9,10], indicating that also other genetic variation must be involved. We investigated to what extent apoE plasma levels are influenced by the *APOEε2/ε3/ε4* polymorphism and whether other genetic variation influencing apoE levels is to be found at the *APOE* locus or elsewhere in the genome.

Subjects and methods

Subjects

Subjects were part of an adolescent and adult Dutch twin cohort of which the characteristics have been described in detail previously[10]. In this study, we used data from 90 dizygotic (DZ) and 70 monozygotic (MZ) adolescent Dutch twin pairs (aged 13-22 years), and 117 DZ and 96 MZ adult Dutch twin pairs (aged 34-62 years). The adolescent and adult twins were collected as separate samples; within the former twin sample, data were collected between 1988 and 1992, and within the latter twin sample between 1992 and 1996. Informed consent was obtained from all participants. Zygosity was confirmed with microsatellite data using the Graphical Representation of Relationship software[11]. Apolipoprotein E levels were assessed in plasma by enzyme-linked immunosorbent assays (ELISA)[12] in adolescent twins in 1994 and in the adult twins in 1998. The Netherlands Heart Foundation and the National Institutes of Health approved this study.

Genotyping

In the DZ twins, 16 STRs with an average inter-markers distance of 6.3 cM on chromosome 19 were genotyped (D19S247, D19S1034, D19S391, D19S865, D19S394, D19S714, D19S49, D19S433, D19S47, D19S420, D19S178, APOC2, D19S246, D19S180, D19S210 and D19S254). The average heterozygosity for these markers was estimated at 0.78. Information of the genetic map was obtained from the Marshfield linkage maps.

The Cy5-labeled PCR products were electrophoretically separated on an automated-fluorescence DNA sequencer, ALF*express* (Amersham Pharmacia Biotech). Analysis and assignment of the marker alleles were performed with Fragment Analyser 1.02 (Amersham Pharmacia Biotech). For monitoring possible genotyping errors, one known genotype was present on each gel, 5% of the genotypings were repeated and two independent observers performed the allele calling. SIBMED[13] was used to check for unlikely double recombinants indicative for genotyping errors.

After running SIBMED and checking the possible genotyping errors in the raw data, approximately 0.2% of the total genotypings appeared to be erroneous. These genotypes were corrected or set to missing values.

The *APOEε2/ε3/ε4* genotypes were determined on genomic DNA in both MZ and DZ twins and in parents of adolescent twins, as described previously[14]. Digestion products were separated on 5% agarose gels. For monitoring possible genotyping errors, two observers independently assessed the *APOEε2/ε3/ε4* genotypes. In addition, a randomly chosen 10% of the samples was reamplified and genotyped. In all cases the previous genotype was confirmed. Previously, in the adolescent twin sample including their parents *APOEε2/ε3/ε4* isoforms (i.e. *APOE* phenotypes) were determined by the use of iso-electric focussing instead of the DNA analysis[6]. In 4.7% of the cases, the original *APOE* phenotypes did not correspond to the genotype, which is in agreement with previous reports[15,16].

Statistical analysis

As the plasma levels of apoE had a skewed distribution, all values were transformed by natural logarithm prior to analysis. Since the data on these twin samples are separately collected during two different time spans and since in adolescence and adulthood lipid and apolipoprotein levels may be influenced by different genes[4], the two cohorts are analysed separately. Allele frequencies of the STRs and the *APOEε2/ε3/ε4* polymorphism were estimated per twin sample using marker data for DZ twins, ignoring their relationships[17]. The full distribution of multipoint identity-by-descent (IBD) sharing probabilities was estimated in DZ twins every cM using Genehunter 2.1[18]. All analyses were performed with the use of a variance components approach implemented in the software package Mx 1.50d[19].

Linkage analysis

The linkage model for the observed apoE levels is represented as:

$$y_{ij} = \mu + (\beta_1 \text{ x } age_{ij}) + (\beta_2 \text{ x } sex_{ij}) + e_{ij},$$

where y_{ij} is the observed apoE level for sib *j* in the *i*-th family, μ denotes the grand mean, β_1 denotes the regression coefficient for age, β_2 denotes the deviation of females, age_{ij} and sex_{ij} denote the age and sex (male = 0; female=1) respectively of sib *j* from the *i*-th family, and e_{ij} denotes a residual term that is not explained by the age and sex effects. The variance of e_{ij} is decomposed into additive genetic variance (A), non-shared environmental variance (E), and additive genetic variance due to a QTL in the vicinity of the marker (Q). No variation due to shared environmental influences was included, as we previously determined the absence of this influence on variation in apoE level[10]. A weighted likelihood approach, which uses the full distribution of IBD probabilities[20], was employed to estimate variation due to the QTL.

Association analysis

The association model for observed apoE levels as a function of the genotyped *APOEε2/ε3/ε4* polymorphism is represented as:

$$y_{ij} = \mu + (\beta_1 \text{ x } age_{ij}) + (\beta_2 \text{ x } sex_{ij}) + (a_b \text{ x } A_{bi}) + (d_b \text{ x } D_{bi}) + (a_w \text{ x } A_{wij}) + (d_w \text{ x } D_{wij}) + e_{ij},$$

where y_{ij} is the observed score for sib *j* in the *i*-th family, μ denotes the grand mean, β_1 denotes the regression coefficient for age, β_2 denotes the deviation of females, age_{ij} and sex_{ij} denote the observed age and sex respectively of sib *j* in the *i*-th family. A_{bi} is the derived coefficient for the additive genetic effect of the *APOEε2/ε3/ε4* polymorphism between families for the *i*-th family,

A_{wij} denotes the coefficient as derived for the additive genetic effects of the *APOEε2/ε3/ε4* polymorphism within families for sib *j* from the *i*-th family, D_{bi} is the coefficient for the dominant genetic effect of the *APOEε2/ε3/ε4* polymorphism between families for the *i*-th family, D_{wij} denotes the coefficient for the dominant genetic effects of the *APOEε2/ε3/ε4* polymorphism within families for sib *j* from the *i*-th family. a_b and a_w are the estimated additive effects between and within families, d_b and d_w are the estimated dominance effects between and within families, and e_{ij} denotes a residual that is not explained by the age, sex and the allelic effects of the *APOEε2/ε3/ε4* polymorphism. The variance of e_{ij} is decomposed into A and E components.

This association model for the observed apoE levels includes effects of *APOEε2/ε3/ε4* genotype on the observations. Allelic effects were calculated both between and within sib pairs, using respectively the genotypic mean of a sib pair and the deviation of a sib from the genotypic mean of a sib pair. The derivation of genotypic means of sib pairs and the differences of sibs from the genotypic means, for di-allelic loci, can be found in Fulker *et al.*[1]. The extensions to dominance effects, multi-allele loci and the use of parental genotypes if available, described by Posthuma[21]. For the adolescent twins, parental *APOEε2/ε3/ε4* genotypes were available and used to derive the between and within families coefficients A_{bi}, A_{wij}, D_{bi} and D_{wij}. For the adult twins, parental *APOEε2/ε3/ε4* genotypes were not available and the sibling genotypes were used to derive between and within families coefficients. Although MZ pairs are uninformative for linkage and uninformative for genotypic effects on observed scores within sib pairs, inclusion of MZ twins allows proper estimation of the additive genetic variance component and provides information for the genotypic effects on the observed scores between sib pairs.

Statistically testing the equivalence of the between and within families effects ($a_b = a_w$ and $d_b = d_w$) provides a test of the presence of population stratification, since in the absence of population stratification, genotypic effects operating within families are equal to the genotypic effects between families. The former represent the true genetic effects, whereas the latter contain both the true and the spurious genetic effects[1].

The presence of non-additive allelic effects of *APOEε2/ε3/ε4* genotype on apoE levels was evaluated by constraining the dominance coefficients to equal zero ($d_b = d_w = 0$). The evidence for an effect of the *APOEε2/ε3/ε4* polymorphism on apoE level is examined by constraining all association parameters to equal zero ($a_b = a_w = d_b = d_w = 0$).

Combined association and linkage analysis

The association model that provided the most parsimonious fit to the data as determined by likelihood ratio test, was taken as a starting point for the combined association and linkage analyses. In the combined model, the variance in apoE levels that was not accounted for by age, sex and allelic effects of the *APOEε2/ε3/ε4* polymorphism (e_{ij}) is decomposed into A, E and also Q components. It was thus tested for each position across chromosome 19 whether linkage with apoE levels is still present when modelled simultaneously with association of the *APOEε2/ε3/ε4* polymorphism.

Results

Table 1 gives the descriptive statistics in MZ and DZ twins, and their phenotypic resemblance (correlation). The high MZ correlations and the large differences between MZ and DZ correlations indicate substantial heritability for apoE levels. To find the most parsimonious

Table 1.
Descriptive statistics of the adolescent and adult Dutch monozygotic (MZ) and dizygotic (DZ) twin pair samples of which both members provide data on apoE level and APOE genotype.

	Adolescent twins		Adult twins	
	MZ	*DZ*	*MZ*	*DZ*
Number of pairs	65	83	88	114
Percentage males	46.2	49.4	47.7	48.2
Age, years – mean (range)	16 (13-22)	17 (13-22)	44 (34-62)	44 (34-59)
ApoE, mg/dL – mean (SD)	7.12 (2.42)	6.59 (2.39)	2.53 (1.03)	2.54 (1.02)
Ln(apoE) – mean (SD)	1.91 (0.34)	1.82 (0.36)	0.86 (0.39)	0.86 (0.37)
Correlation	0.88	0.37	0.87	0.41

model to describes the data on apoE plasma levels in adolescent and adult twins, a model in which the grand mean with regression deviations for age and sex is included and in which the residual variance is decomposed into additive genetic variance (A) and unique environmental variance (E), is fitted to the data. The results of the model fitting procedure are shown in table 2. Per analysis, the most extended model was simplified by leaving one factor from the model and when the simplification of the model has a significant effect ($p<0.05$), the factor has a significant effect on apoE levels and should not be excluded from the model. The models of the linkage and association analyses are compared to the model of the 'no linkage, no association' analysis to test, whether the QTL or the association had significant effect on the observed apoE levels. In the combined association and linkage analysis was tested whether the QTL had a significant effect on apoE levels in the presence of association.

In adolescent twins, sex (Table 2A, model 2) and age (Table 2A, model 3) both have a significant effect on apoE levels. ApoE levels significantly decrease with older age (β_1=-0.03) and are significantly higher in women (β_2=0.22). Model 1 thus describes the apoE levels observed in adolescent twins most parsimoniously and shows that 87% of the total variance in apoE levels is determined by additive genetic factors (Table 3A, model 1). In adult twins, sex has no significant effect on apoE levels (Table 2B, model 2) and age has a significant effect on apoE levels (Table 2B, model 3). Adult apoE levels increase significantly with older age (β_1=0.01). Model 2 describes the apoE levels observed in adult most parsimoniously and shows that 85% of the total variance in adult apoE levels is determined by additive genetic factors (Table 3B, model 2). These best fitting models, i.e. model 1 for the adolescents and model 2 for the adults, are extended for linkage, association and combined association and linkage analyses.

For linkage analysis, the variance of the residuals is decomposed into A, E and additionally, in additive genetic variance due to a QTL (Q). The adolescent twins provide an indication for linkage on chromosome 19q13.3, 70 cM from pter, with a maximum LOS score (MLS) of 0.8 (χ^2 of 3.49; p=0.06) (Table 2A, model 5) (Figure 1, 'Linkage only'). This putative QTL is estimated to explain 36% of the variance in adolescent apoE levels (Table 3A, model 5). The adult twins, provide evidence in favour of linkage at 63 cM from pter, with a MLS of 2.5 (χ^2 of 11.71; p< 0.000) (Table 2B, model 5) (Figure 2, 'Linkage only'). Both linkage peaks in adolescent and adult twins largely overlap the location of the *APOE* gene (19q13.32). This putative apoE QTL is estimated to explain 85% of the total variance in adult apoE levels (Table 3B, model 5).

For association analyses, the models 1 and 2 for the adolescent and adult twins, respectively, are

extended with additive and non-additive allelic effects of the *APOEε2/ε3/ε4* polymorphism between and within families. These extensions to model 1 in the adolescent twins, result in a significantly better fit to the observed apoE levels (Table 2A, model 6). To test whether this association is confounded by population stratification, the allelic effects of the *APOEε2/ε3/ε4*

Table 2.

Fit statistics of nested models in the adolescent and adult Dutch twins. The most parsimonious models are represented in bold. In the models, A represents additive genetic variance, E non-shared environmental variance, Q additive genetic variance due to a QTL in the vicinity of a marker, μ = grand mean; β_1 regression weight of age; β_2= female deviation. Column headings represent: df= degrees of freedom; -2ll= -2 times the log likelihood fit statistic; vs.= versus, which represents the model to which the comparison has been made; Δdf= difference in degrees of freedom between the two tested models; χ^2= difference in -2ll between the two tested models, which follows a χ^2 distribution; p= p-*value.*

2A	Model	df	-2ll	vs.	Δdf	χ^2	p
ADOLESCENT TWINS	**No Linkage, no association**						
	1: AE + $\mu + \beta_1 + \beta_2$	**289**	**78.14**	**-**	**-**		
	2: AE + $\mu + \beta_1$	290	103.28	2 vs 1	1	25.14	0.00
	3: AE + $\mu + \beta_2$	290	84.84	3 vs 1	1	6.71	0.01
	4: AE + μ	291	109.21	4 vs 2	1	24.37	0.00
	Linkage at maximum χ^2 difference						
	5: AEQ + $\mu + \beta_1 + \beta_2$	**288**	**74.65**	**1 vs 5**	**1**	**3.49**	**0.06**
	Association						
	6: AE + $\mu + \beta_1 + \beta_2 + a_b + a_w + d_b + d_w$	279	-44.58	1 vs 6	10	122.72	0.00
	7: AE + $\mu + \beta_1 + \beta_2 + (a_b = a_w) + (d_b = d_w)$	**284**	**-35.46**	**7 vs 6**	**5**	**9.12**	**0.10**
				1 vs 7	**5**	**113.60**	**0.00**
	8: AE + $\mu + \beta_1 + \beta_2 + (a_b = a_w) + (d_b = d_w = 0)$	287	-21.42	8 vs 7	3	14.05	0.00
	Linkage at maximum χ^2 + Association						
	9: AEQ + $\mu + \beta_1 + \beta_2 + (a_b = a_w) + (d_b = d_w)$	**283**	**-36.75**	**7 vs 9**	**1**	**1.29**	**0.26**
2B							
ADULT TWINS	**No Linkage, no association**						
	1: AE + $\mu + \beta_1 + \beta_2$	399	191.76	-	-		
	2: AE + μ + β1	**400**	**191.97**	**2 vs 1**	**1**	**0.21**	**0.65**
	3: AE + $\mu + \beta_2$	400	203.31	3 vs 1	1	11.55	0.00
	4: AE + μ	401	203.39	4 vs 2	1	11.42	0.00
	Linkage at maximum χ^2 difference						
	5: AEQ + $\mu + \beta_1$	**399**	**180.26**	**2 vs 5**	**1**	**11.71**	**0.00**
	Association						
	6: AE + $\mu + \beta_1 + a_b + a_w + d_b + d_w$	390	71.13	2 vs 6	10	120.84	0.00
	7: AE + $\mu + \beta_1 + (a_b = a_w) + (d_b = d_w)$	395	73.76	7 vs 6	5	2.63	0.76
	8: AE + $\mu + \beta_1 + (a_b = a_w) + (d_b = d_w = 0)$	**398**	**78.46**	**8 vs 7**	**3**	**4.70**	**0.20**
				2 vs 8	**2**	**113.51**	**0.00**
	Linkage at maximum χ^2 + Association						
	9: AEQ + $\mu + \beta_1 + (a_b = a_w) + (d_b = d_w = 0)$	**397**	**70.70**	**8 vs 9**	**1**	**7.76**	**0.01**

polymorphism between families is equated to the effects within families. This simplified association model describes the observed apoE levels in adolescent not significantly (Table 2A, model 7), indicating that the association of the *APOEε2/ε3/ε4* polymorphism with adolescent apoE levels is not confounded by population stratification.

Furthermore, dominant effects of the *APOEε2/ε3/ε4* polymorphism on adolescent apoE levels could not be equated to zero (Table 2A, model 8), indicating that dominance plays a role in adolescent apoE levels. Carriers of the *ε3/ε4* and *ε4/ε4* genotype had lower levels of apoE as compared to the common *ε3/ε3* genotype (Table 4a). Carriers of the *ε2* allele have higher levels as compared to the common *ε3/ε3* genotype, but the effect size of the *APOEε2* allele appears to depend on the accompanying allele. When an *ε2* allele accompanies the *ε2* allele, apoE levels are intermediate high as compared with an accompanying *ε3* or *ε4* allele. In adolescent twins, the *APOEε2/ε3/ε4* polymorphism is strongly associated with apoE levels (Table 2A, model 1 vs model 7), which explains 39% of the total apoE variance (Table 3A).

The extensions to model 2 in the adult twins, also resulted in a significantly better fit to the observed apoE data (Table 2B, model 6). This association of the *APOEε2/ε3/ε4* polymorphism with apoE levels in adults is not confounded with population stratification (Table 2A, model 7). Furthermore, there is no evidence that dominant effects of the *APOEε2/ε3/ε4* polymorphism play a significant role in adult apoE levels (Table 2B, model 8). The alleles thus showed an additive

Table 3.
Variance components estimates in the most parsimonious models for the situations of no linkage and no association, linkage at 70 cM from pter, association with APOEε2/ε3/ε4 polymorphism, and simultaneous linkage at 70 cM from pter and association in the adolescent and adult Dutch twins.

3A		Model	Variance (corrected for age and/or sex)	% of variance due to association	% of variance due to A	% of variance due to E	% of variance due to Q
ADOLESCENT TWINS	No Linkage, no association	1	0.1106	-	87	13	-
	Linkage at maximum LOD score	5	0.1106	-	51	13	36
	Association	7	0.1106	39	49	13	-
	Association + Linkage at peak Linkage only (70 cM from pter)	9	0.1106	39	49	13	0
	Association + Linkage at maximum LOD score (43 cM from pter)	9	0.1106	4	64	13	19
3B							
ADULTS TWINS	No Linkage, no association	2	0.1319	-	85	15	-
	Linkage at maximum LOD score	5	0.1319	-	0	15	85
	Association	8	0.1319	30	56	15	-
	Association + Linkage at peak Linkage only (63 cM from pter)	9	0.1319	32	0	15	53
	Association + Linkage at maximum LOD score (60 cM from pter)	9	0.1319	30	0	15	55

Table 4A.
Observed and expected means of apoE levels according to APOEε2/ε3/ε4 genotype in adolescent Dutch twins.

APOEε2/ε3/ε4 genotype	Observed Mean (mg/dL)	Expected Mean (mg/dL)
ε2ε2 (n=1)	7.20	7.17
ε2ε3 (n=58)	9.16	9.87
ε2ε4 (n=4)	9.24	10.59
ε3ε3 (n=152)	6.27	6.62
ε3ε4 (n=75)	5.15	5.47
ε4ε4 (n=6)	4.50	4.90

Table 4B.
Observed and expected mean of apoE levels according to APOEε2/ε3/ε4 genotype in adult Dutch twins.

APOEε2/ε3/ε4 genotype	Observed Mean (mg/dL)	Expected Mean (mg/dL)
ε2ε2 (n=3)	5.40	5.37
ε2ε3 (n=55)	3.48	3.49
ε2ε4 (n=12)	3.43	3.06
ε3ε3 (n=218)	2.30	2.27
ε3ε4 (n=109)	1.99	1.99
ε4ε4 (n=7)	1.51	1.75

effect on apoE levels with a decreasing pattern from *ε2* to *ε3* to *ε4*. Carriers of the *ε2* allele had higher apoE plasma levels and carriers of the *ε3/ε4* and *ε4/ε4* genotype had lower levels of apoE as compared to the common *ε3/ε3* genotype (Table 4b). The *APOEε2/ε3/ε4* polymorphism is also strongly associated with apoE levels in adult twins (Table 2B, model 2 vs 8), which explains 30% of the total apoE variance (Table 3B).
When the allelic association effects are simultaneously modelled with linkage, model 7 for adolescents and 8 for adults are extended with the decomposition of the residual variance into A, E and additionally in Q. In this combined association and linkage analysis, all linkage at the peak location of the 'linkage only' analysis in the adolescent twins (70 cM from pter), is explained by the association of the *APOEε2/ε3/ε4* polymorphism with apoE levels, since evidence for linkage completely disappears (Table 2A, model 9) (Figure 1, 'combined association and linkage'). Table 3A also shows that at 70 cM from pter no variance is left due to a QTL and that 39% of the apoE variance can be explained by its association with the *APOEε2/ε3/ε4* polymorphism in adolescent twins.
In adult twins, this combined association and linkage analysis shows that part of the apoE variance linked to 19q13.3 can be explained by its association with the *APOEε2/ε3/ε4* polymorphism (Table 2B, model 9) (Figure 2, 'combined association and linkage'). Table 3B shows that at 63 cM from pter, 53% of the variance is still linked to 19q13.3 and that 32% of the apoE variance can be explained by its association with the *APOEε2/ε3/ε4* polymorphism in adult twins. All residual genetic variance in adult apoE levels is still linked to chromosome 19q13.3 (Table 3), indicating that this chromosomal region harbours additional genetic variation influencing apoE plasma levels in adults.

Discussion

A crucial step following a positive linkage result from a genome scan for quantitative traits, is to identify which gene variants in the chromosomal region explain the linkage. Fulker *et al.*

Figure 1.
'Linkage only' analysis (solid lines: ▬▬ and ——) and 'combined association and linkage' analysis (dashed lines: - - - - and · · · ·) of chromosome 19 with apoE levels in the adolescent Dutch twins. Thick lines correspond with the left axe, indicating the LOD score. Thin lines correspond with the right axe, indicating the percentage of the variance (corrected for age and sex) that is accounted for by the QTL.

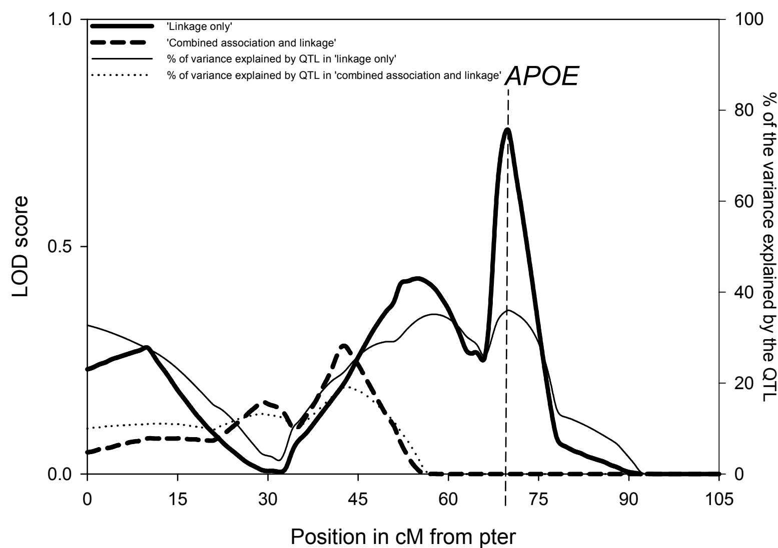
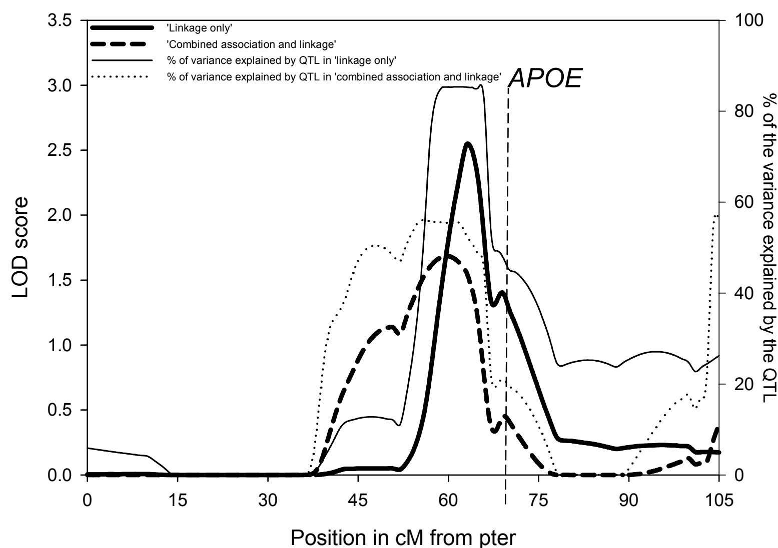


Figure 2.
'Linkage only' analysis (solid lines: ▬▬ and ——) and 'combined association and linkage' analysis (dashed lines: - - - - and · · · ·) of chromosome 19 with apoE levels in the adult Dutch twins. Thick lines correspond with the left axe, indicating the LOD score. Thin lines correspond with the right axe, indicating the percentage of the variance (corrected for age) that is accounted for by the QTL.



developed a method for simultaneous modelling of association and linkage for quantitative traits using sib pair data that controls for population stratification[1]. Until now, most studies performed linkage analyses to find QTLs, but some studies also carried out combined association and linkage analyses[22,23]. However, these studies did not control for population stratification.
Recently, the combined association and linkage analysis that includes a test for population stratification[1], was extended to incorporate effects of multi allelic loci, parental genotypes and genetic dominance by Posthuma *et al.*[21,24]. We applied the extended approach for quantitative traits in sib pairs to the *APOE* locus and its effect on apoE plasma levels.
We found evidence for linkage of chromosome 19q13.3 with apoE levels in both adolescent and adult Dutch twins, although it did not reach genome-wide significance according to Lander and Kruglyak criteria[25]. The peak LOD score in adolescents was exactly on the location where the *APOE* gene is mapped. The peak LOD score in adults is located 7 cM pter from the *APOE* gene, but the linkage region included the *APOE* gene. Both adolescent and adult apoE levels were strongly associated with the *APOEε2/ε3/ε4* polymorphism and these association were not confounded by population stratification. Using the combined association and linkage approach, we showed that linkage of 19q13.32 with apoE plasma levels is completely explained by the *APOEε2/ε3/ε4* polymorphism in adolescents, but only partly in adults. This finding in young age indicates that the *APOEε2/ε3/ε4* polymorphism is the only relevant genetic variant in *APOE*, not taking into account variants in complete LD. Later in life, however, other genetic variation may become relevant, such as other polymorphisms in the *APOE* gene, which have recently been shown to influence the variance in apoE levels in addition to *APOEε2/ε3/ε4* polymorphism in adults[8]. Of special interest are *APOE* promoter polymorphisms, –491A/T and –219G/T, that may influence transcription of the *APOE* gene[26]. Previously, the –219G/T polymorphism has been associated with apoE plasma levels[8,27]. This SNP is in partial LD with the *APOEε2/ε3/ε4* polymorphism[28,29] and thus may have contributed to the association we observed.
In adolescent twins, not all variance linked to chromosome 19 can be explained by the *APOEε2/ε3/ε4* polymorphism, since at 43 cM from pter a QTL explaining 19% of the total variance in apoE levels remains, although this is not a significant QTL. Possibly, on chromosome 19p13, additional genetic variants are influencing adolescent apoE levels, but for a complete explanation of the genetic variance in adolescent apoE levels, additional genetic variation should be found on other chromosomes. Since in adult twins, part of the variance linked to chromosome 19 can be explained by the *APOEε2/ε3/ε4* polymorphism and the remaining putative apoE QTL at 60 cM from pter explains 55% of the total variance, no residual heritability remains to be explained outside chromosome 19. The genetic architecture of adolescent apoE levels thus seems to differ from that of adult levels. This is consistent with previous studies, describing that lipid and apolipoprotein levels may be influenced by different genes in adolescence and adulthood[30] and that a quantitative trait locus on chromosome 19 seem to play only a role in adult and not in adolescent LDL cholesterol levels[33]. Such differential influence of genes at different age might also be the case for apoE levels.
We found that age has a decreasing effect on the plasma levels of apoE in the adolescent twins while it has an increasing effect in adult twin cohort. This differential effect of age may reflect the different interactions of the genes influencing apoE levels with the environment at different ages (gene-environment interaction) as assumed by Zerba *et al.*[31,32]. This finding may also refer to the influence of different genes, including *APOE*, at different age on apoE levels and their

interactions (*APOE* gene-gene interaction), reinforcing our observation that genes have differential influence on measures of lipid metabolism at different ages.
Much is known about the association of the *APOEε2/ε3/ε4* polymorphism in adult apoE levels. Adolescent apoE levels, however, are rarely investigated. Neale *et al.* also studied this adolescent Dutch twin cohort[6] and found that the *APOEε2/ε3/ε4* polymorphism explained 16% of the total variance in apoE levels, without evidence for non-additive allelic effects. In their analyses, Neale *et al.* used *APOEε2/ε3/ε4* phenotypes, which lead to genotype discrepancy, which may have influenced the proportion of the variance due to the *APOEε2/ε3/ε4* polymorphism. Furthermore, they included parental apoE levels and equated the association effects in the parental and offspring generation[6], which diluted the non-additive allelic effect in the adolescent twins to non significance. The absence of such effects in the parental generation, converges with our results from the adult Dutch twin sample.
To our knowledge, this is the first time that it in adolescents, non-additive allelic effect of the *APOEε2/ε3/ε4* polymorphism on apoE levels are found. Additionally, the finding that the *APOEε2/ε3/ε4* polymorphism is the only variant in the *APOE* gene influencing apoE levels in adolescents, not taking into account variants in complete LD, is different from the findings in adults. This differential architecture across cohorts as well as differential effects of age across cohorts reveal the dynamic nature of genetic and environmental effects.
We showed that using a combined association and linkage analysis, linkage results could be explained by genetic variation in positional candidate genes. It is possible to find out whether the variant tested is the only genetic variant influencing the quantitative trait, or that additional variants have to be found for a complete explanation of the linkage result.

Acknowledgements

We would like to thank R. Buitenhek and J.J. van der Boom for the apoE plasma level measurements. This study was supported by grants to GP Vogler from the National Institutes of Health, National Heart, Lung, and Blood Institute (HL55976), to PE Slagboom (98.199) and DI Boomsma (86.083, 88.042 and 90.313) from the Netherlands Heart Foundation.

References

1. Fulker DW, Cherny SS, Sham PC, Hewitt JK. Combined linkage and association sib-pair analysis for quantitative traits. *Am J Hum Genet* 1999; 64: 259-267.
2. Atwood LD, Heard-Costa NL. Limits of fine-mapping a quantitative trait. *Genet Epidemiol* 2003; 24: 99-106.
3. Abecasis GR, Cardon LR, Cookson WO. A general test of association for quantitative traits in nuclear families. *Am J Hum Genet* 2000; 66: 279-292.
4. Mensenkamp AR, Jong MC, van Goor H, van Luyn MJ, Bloks V *et al.* Apolipoprotein E participates in the regulation of very low density lipoprotein-triglyceride secretion by the liver. *J Biol Chem* 1999; 274: 35711-35718.
5. Sing CF, Davignon J. Role of the apolipoprotein E polymorphism in determining normal plasma lipid and lipoprotein variation. *Am J Hum Genet* 1985; 37: 268-285.
6. Neale MC, de Knijff P, Havekes LM, Boomsma DI. ApoE polymorphism accounts for only part of the genetic variation in quantitative ApoE levels. *Genet Epidemiol* 2000; 18: 331-340.
7. Boerwinkle E, Utermann G. Simultaneous effects of the apolipoprotein E polymorphism on apolipoprotein E, apolipoprotein B, and cholesterol metabolism. *Am J Hum Genet* 1988; 42: 104-112.
8. Stengard JH, Clark AG, Weiss KM, Kardia S, Nickerson DA *et al.* Contributions of 18 additional DNA sequence variations in the gene encoding apolipoprotein e to explaining variation in quantitative measures of lipid metabolism. *Am J Hum Genet* 2002; 71: 501-517.
9. Duggirala R, Blangero J, Almasy L, Dyer TD, Williams KL *et al.* A major susceptibility locus influencing plasma triglyceride concentrations is located on chromosome 15q in Mexican Americans. *Am J Hum Genet* 2000; 66: 1237-1245.
10. Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB *et al.* Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Res* 2002; 5: 87-97.

11. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; 17: 742-743.
12. Bury J, Vercaemst R, Rosseneu M, Belpaire F. Apolipoprotein E quantified by enzyme-linked immunosorbent assay. *Clin Chem* 1986; 32: 265-270.
13. Douglas JA, Boehnke M, Lange K. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 2000; 66: 1287-1297.
14. O'Dell SD, Humphries SE, Day IN. Rapid methods for population-scale analysis for gene polymorphisms: the ACE gene as an example. *Br Heart J* 1995; 73: 368-371.
15. Kim JH, Lee EJ, Kwon OH. Apolipoprotein E genotyping and phenotyping in type II diabetes mellitus patients with hypertriglyceridemia. *Clin Biochem* 1997; 30: 47-52.
16. Kardaun JW, White L, Resnick HE, Petrovitch H, Marcovina SM *et al.* Genotypes and phenotypes for apolipoprotein E and Alzheimer disease in the Honolulu-Asia aging study. *Clin Chem* 2000; 46: 1548-1554.
17. Broman KW. Estimation of allele frequencies with data on sibships. *Genet Epidemiol* 2001; 20: 307-315.
18. Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; 58: 1347-1363.
19. Neale MC, Boker SM, Xie G, Maes HH. Mx: Statistical modeling. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University; 1999.
20. Neale MC. QTL Mapping with sib pairs: The flexibility of Mx, In: Spector TD, Snieder H, MacGregor AJ, editors. Advances in twin and sib pair analysis. London: Greenwich Medical Media; 2000.
21. Posthuma D., Geus E.J.C., Boomsma D.I., and Neal, M. C. Combined linkage and association in Mx. *Behav Genet* 2004; 34: 153-174. In Press
22. Klos KL, Kardia SL, Ferrell RE, Turner ST, Boerwinkle E, Sing CF. Genome-wide linkage analysis reveals evidence of multiple regions that influence variation in plasma lipid and apolipoprotein levels associated with risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2001; 21: 971-978.
23. Soria JM, Almasy L, Souto JC, Bacq D, Buil A *et al.* A quantitative-trait locus in the human factor XII gene influences both plasma factor XII levels and susceptibility to thrombotic disease. *Am J Hum Genet* 2002; 70: 567-574.
24. Abecasis GR, Cookson WO, Cardon LR. Pedigree tests of transmission disequilibrium. *Eur J Hum Genet* 2000; 8: 545-551.
25. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11: 241-247.
26. Artiga MJ, Bullido MJ, Sastre I, Recuero M, Garcia MA *et al.* Allelic polymorphisms in the transcriptional regulatory region of apolipoprotein E gene. *FEBS Lett* 1998; 421: 105-108.
27. Lambert JC, Brousseau T, Defosse V, Evans A, Arveiler D *et al.* Independent association of an APOE gene promoter polymorphism with increased risk of myocardial infarction and decreased APOE plasma concentrations-the ECTIM study. *Hum Mol Genet* 2000; 9: 57-61.
28. Fullerton SM, Clark AG, Weiss KM, Nickerson DA, Taylor SL *et al.* Apolipoprotein E variation at the sequence haplotype level: implications for the origin and maintenance of a major human polymorphism. *Am J Hum Genet* 2000; 67: 881-900.
29. Heijmans BT, Slagboom PE, Gussekloo J, Droog S, Lagaay AM *et al.* Association of APOE epsilon2/epsilon3/epsilon4 and promoter gene variants with dementia but not cardiovascular mortality in old age. *Am J Med Genet* 2002; 107: 201-208.
30. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.
31. Zerba KE, Ferrell RE, Sing CF. Genotype-environment interaction: apolipoprotein E (ApoE) gene effects and age as an index of time and spatial context in the human. *Genetics* 1996; 143: 463-478.
32. Zerba KE, Ferrell RE, Sing CF. Complex adaptive systems and human health: the influence of common genotypes of the apolipoprotein E (ApoE) gene polymorphism and age on the relational order within a field of lipid metabolism traits. *Hum Genet* 2000; 107: 466-475.
33. Beekman M, Heijmans BT, Martin NG, Whitfield JB, Pedersen NL *et al*. Evidence for a QTL on chromosome 19 influencing LDL cholesterol levels in the general population. *Eur J Hum Genet* 2003; 11: 845-850.

8 | Summary and General Discussion

Summary and discussion of results

Increased plasma levels of LDL cholesterol, total cholesterol, TG, apoB and apoE, as well as decreased plasma levels of HDL cholesterol, apoA1 and apoA2 are intermediate phenotypes of CVD. These quantitative traits in the lipid metabolism pathway are considered to be among the main risk factor for development of atherosclerosis[1,2], the primary cause of CVD. Over the last years, many candidate gene studies and genome-wide scans have been performed to determine the contribution of genetic loci to the clinical endpoints or intermediate phenotypes of CVD. In general, these studies differ in samples size, sample ethnicity, statistical methods and the way the study subjects and families in those studies were selected for various clinical endpoints of CVD. These differences in design, samples, and approaches often make both association studies and whole genome scan results difficult to replicate[3,4].

The project reported in this thesis is aimed at the mapping of major genes involved in the regulation of the quantitative traits of lipid metabolism in the general population. Phenotypes and marker data were collected in four population-based samples; an adolescent and adult twin pair sample from the Netherlands, one adult twin sample from Sweden and an adult twin sample from Australia. These four samples can be used to investigate replication, as well as to test QTL effects simultaneously.

Heritability of lipid and apolipoprotein levels

To investigate the influence of genes and environment on the variation of apolipoprotein and lipid levels, the heritability of these traits were estimated in the four twin pair samples (Chapter 2). Between 48-87 % of the variance in lipid and apolipoprotein levels are attributable to genetic factors in the four twin pair samples. Also the heritabilities of apoE and apoA2 levels, which have rarely been reported before, were found to be very high, comparable to other lipid and apolipoprotein levels. Common environment does not influence the variances and there are no sex differences in the heritabilities of these intermediate phenotypes. Overall, the heritability estimates were the highest in the Dutch samples and the lowest in the Australian sample. These differences in heritability were significant, indicating a sample specific effect on heritability of lipid and apolipoprotein levels or on the influence of unique environmental factors, which was overall the largest in the Australian sample. Higher unique environmental effects lead concomitantly to lower estimates of genetic factors. This larger influence of unique environmental factors could be explained by larger environmental effects, interactions with genetic factors or measurement errors of the trait. Overall, these twin cohorts from three different countries thus provide a unique sample for QTL mapping of the genes involved.

Genome scanning protocol

Since genetic variation seems to play an important role in lipid and apolipoprotein levels, a genome-wide scan is a logical subsequent analysis to localise the chromosomal regions harbouring putative quantitative trait loci (QTL). In our study, a total number of 1,000 DZ subjects need to be genotyped. Regularly, 10 cM spaced marker sets, including about 400 STRs, are used for genome scanning, which would result in roughly 400,000 genotypings in our genome scan. In Chapter 3, power simulations are carried out to estimate the optimal inter marker distance, resulting in an optimal balance between power to detect major genes and genotyping effort. From these simulations resulted that using an average spacing of 20 cM, there is 80% power to detect a QTL effect of 25% at the significance level of 5%. Because of the large samples size of the study, a specific and practically feasible protocol was developed at the start of the genome scan based on the simulations. Therefore, a screening set of 229 STRs with an average inter marker distance of 18.3 cM was chosen. A complete genome scan using this protocol requires 80 multiplex PCR reactions, which can all be carried out using one set of conditions and would consume 2.5 µg of genomic DNA, which is an exceptionally small amount for a genome wide scan. Current automated methodology allows for much higher throughput. In spite of the simulation study we would now prefer a much denser marker spacing, for example 5 or 10 cM inter-marker distance, which will not only gain the power in linkage analyses, to obtain genome wide significance, but also enables a more accurate search for possible genotyping errors and taking into account that not all markers will be completely informative.

QTL mapping

Linkage analyses were carried out on the data obtained from chromosomes 1, 2, 6, 7, 8, 11, 15, 16, 17 and 19. Before the linkage analyses, all family relationships were checked and an estimate was obtained of the genotyping error rate. For the current study, this is especially important since only few parental data were available, so Mendelian inheritance of marker alleles cannot be checked. All genotypings were used as input for the GRR software[5], which calculates the

proportion of shared alleles for each twin pair across all markers. The linkage analyses were finally performed in 83 DZ pairs for the adolescent Dutch, 117 for the adult Dutch, 44 for the Swedish and 249 for the Australian twins. Furthermore, marker data were checked for unusual double recombinants using SIBMED[6], the occurrence of which may be due to genotyping errors. After running SIBMED and checking the possible genotyping errors with the raw data, approximately 0.01% of the total genotypings appeared to be erroneous. In this way, only a small proportion of the potential genotyping errors can be found. The large inter-marker distance in our study increases the probability of a double recombination. The erroneous genotypes were corrected when possible or otherwise set to missing. Linkage analyses were performed with these cleaned data sets and reported in Chapter 4.

To investigate whether the genome scan with 18.3 cM spacing had indeed sufficient power to detect major loci, the *LPA* locus on chromosome 6q27, which is known to explain approximately 90% of the variance in Lp(a) levels, was tested for linkage. In all four samples linkage was found at chromosome 6q; in the adolescent Dutch and Swedish samples there was suggestive linkage[3] (MLS between 2.2 and 3.6) with MLS of 2.75 and 2.28 respectively. Additionally, in the adult Dutch twins as well as in the Australian twins significant linkage was found with MLSs of 4.40 and 5.13 respectively. Consequently, we have confirmed linkage according to the criteria of Lander and Kruglyak[3] in independent samples, which gives strong evidence that we located a QTL influencing Lp(a) levels. Even more, in a simultaneous variance components linkage analysis, in which the proportion of the QTL effect was equated over the four samples, the MLS increased to 9.8.

To investigate whether we could detect QTLs with a moderate effect using our genome scan protocol, the *APOE* locus was tested for linkage. The *APOE* gene is the structural coding gene for apolipoprotein E and is located on chromosome 19q13.2. Between 9 to 20% of the total variation in apoE levels has been ascribed to this candidate gene[7-9], a QTL effect much smaller then assumed in the simulation analyses (QTL effect of 25%) that formed the basis for the choice of the marker spacing. At this locus we found indications for linkage in the adolescent Dutch twins (MLS = 1.0) and in the adult Dutch twins (MLS = 0.6). No evidence for linkage was found in the Australian twins. The heritability in this population, however, was also reasonably low, raising doubts as to the quality of the apoE phenotypings performed in this twin sample. Since apoE levels were not measured in the Swedish twins, no linkage analysis with apoE levels could be performed in this sample. In the simultaneous analysis of the three twin samples, no evidence for heterogeneity was present and the MLS was 1.0. Since a MLS of 1.0 is expected to occur approximately 8 times at random in a genome scan[10], we conclude that we are unable to detect this QTL with a moderate effect on apoE plasma levels. With the 18 cM spacing in our studies, we cannot, as expected, detect a QTL explaining approximately 15% of the variance in apoE levels at a significant level in three twin pair samples totalling 449 pairs. The *APOE* example indeed illustrates that in our genome scanning approach, only QTLs with a major effect (>25%) on the trait can be detected. Consequently, we will miss as false negatives QTLs with smaller effects. Reconsidering, we are convinced that denser marker spacing would have improved the study since effects of 25% for a single QTL may be somewhat unrealistically high.

Analysing the 10 finished chromosomes for suggestive or significant linkage, we detected 11 putative QTLs; 4 loci influencing Lp(a) levels, 3 loci influencing apoA1 levels, 2 loci influencing apoB levels and 2 loci influencing LDL cholesterol levels. Two loci, in addition to the *LPA* locus, showed significant linkage (MLS>3.6) with Lp(a) levels. The significance criterion may

be too stringent for detecting QTLs in the separate twin samples, while only part of the analyses has the power to actually reach MLS higher than 3.6. This is why we consider that suggestive linkage results (MLS between 2.2 and 3.6) may also be valuable. Of the eleven loci showing significant or suggestive linkage, three loci could also be replicated in an other twin sample with MLS higher than 1.4. According to Lander & Kruglyak, a point-wise p-value of 0.01, corresponding with a LOD score of 1.4, is needed for a 20 cM interval-wide significance level of 0.05[3]. Two loci, on chromosome 1 and 2, show replicated linkage with Lp(a) levels, and chromosome 19 shows replicated linkage with LDL cholesterol levels. Although, total cholesterol, HDL cholesterol, triglycerides and apoE levels are all highly heritable, no significant or suggestive linkage results have been found. This might indicate that the major gene loci are not present on these 10 chromosomes, or that the QTL effects on these traits are smaller that 25%.

Replicated linkage with Lp(a) levels

Among the replicated loci are two loci influencing Lp(a) levels, on chromosome 1 and 2. In the region on chromosome 2, the *APOB* gene is located, encoding apolipoprotein B, which is the major constituent of the Lp(a) particle. It, therefore, could be hypothesised that mutations in the APOB gene might affect Lp(a) levels. However, in the two Dutch samples we found no evidence yet for linkage of chromosome 2 with Lp(a) levels.
Our genome scan results showed also a putative QTL influencing Lp(a) levels on chromosome 1. The Swedish and the Australian sample show MLS of 2.6 and 1.8 for Lp(a) levels, respectively. Broeckel *et al*[11] performed a genome scan for myocardial infarction and its quantitative risk factors. From this genome scan was concluded that on chromosome 1 a second QTL was located influencing Lp(a) levels. In attempt to confirm the result from Broeckel, we carried out advanced linkage analyses (Chapter 5). Simultaneous analyses of the four twin samples resulted in a MLS of 1.6 for Lp(a) levels on chromosome 1. This QTL would explain 44% of the total variance in Lp(a) levels, although the LPA gene already explained 82% of the variance. It is a common phenomenon that the QTL effect sizes are overestimated[12], resulting in more than 100% of the total trait variance explained by several QTLs found in one genome scan. Therefore, a two-locus analysis was conducted to analyse the chromosome 1 QTL together with the *LPA* locus in one model. It could be concluded that 82% of the variance in Lp(a) levels can be explained by the *LPA* locus and that the putative QTL on chromosome 1 was a false positive result. This may point at low power. Alternatively, it may indicate that two-locus analyses constitute a valuable tool to detect false positive linkage results.

Fine mapping of chromosome 19

On chromosome 19, the adult Dutch and Australian twin samples show linkage with LDL cholesterol levels with MLS of 2.2 at 54.5 cM from pter and 1.7 at 34.3 cM from pter, respectively. Also in the Swedish sample, we found positive linkage in the same chromosomal region, which makes this locus even more interesting. Previous studies reported evidence for linkage of chromosome 19 with total cholesterol levels and LDL-C levels in Pima Indians[13] and Hutterites[14], respectively. Our study could extend these findings in genetic isolates to the general, Caucasian population. However, the chromosomal region of linkage we found is very broad. To narrow down this region and also to gain power in the linkage analysis in this region, additional markers were genotyped in all twin samples, resulting in an inter-marker spacing of 6 cM in the

Dutch twin samples and 8 cM in the Swedish and Australian twin samples. MLS increased to 4.5, 1.7 and 2.1 in the adult Dutch, Swedish and Australian samples, respectively (Chapter 6). No evidence for linkage was found in the adolescent Dutch twins. After fine mapping, the linkage region is still very broad and was not narrowed down by the genotyping of additional STRs of additional family members (data not shown). From QTL mapping in *Drosophila melanogaster* and *Saccharomyces cerevisiae*, it appeared that linkage results are often caused by several polymorphisms with a small effect, which could be interpreted as one gene with a major effect[15,16]. This might explain that the position of the MLS in the Australian sample is different from that detected in the Dutch and Swedish samples. Possibly, two loci at 35 and 55 cM play a role in LDL cholesterol levels, and the former plays a major role in Australians and the latter in the European samples.

A simultaneous analysis of the three adult samples showing linkage on chromosome 19 with LDL cholesterol levels increased the MLS to 5.7 at 60 cM. In the Dutch sample, this QTL explained approximately 90% of the total variance of LDL cholesterol levels, in the Swedish 80% and in the Australian sample 40%. Hence, we located a major QTL influencing LDL cholesterol levels on chromosome 19. Since apoB is a constituent of LDL particles, apoB levels are highly correlated to LDL cholesterol levels. We performed bivariate linkage analysis of chromosome 19 with both parameters. The putative LDL cholesterol QTL on chromosome 19, indeed contributes also to the variance in apoB levels. Considering the peak-width even after fine mapping and bivariate analysis, various candidate loci may exert combined and pleiotropic effects on LDL cholesterol and apoB levels.

Previously, evidence for linkage of chromosome 19 with total cholesterol and LDL cholesterol levels were reported in Pima Indians[13] and Hutterites[14], respectively. Our study extends these findings in genetic isolates to the general, Caucasian population. Genome scans in the San Antonio Heart Study[17] and the Rochester Family Heart Study[18] were also performed in Caucasians from the general population. In Mexican Americans, linkage with LDL1 and LDL 2 cholesterol levels were found with MLS of 2.4 and 1.9[17], respectively, and in non-Hispanic whites, linkage with total cholesterol levels was found with MLS of 1.1[18]. Although the reported LOD scores for chromosome 19 were not significant in these studies, the findings are compatible with our findings. No suggestion for linkage with total cholesterol or LDL-C levels, however, was found in selected samples of myocardial infarction patients, type 2 diabetes patients and patients with familial combined hyperlipidemia[11,19-22]. The LDL-C QTL on chromosome 19 constitutes one of the most replicated result from linkage studies, virtually ruling out the possibility that it is a false positive observation. Our study in three twin samples strongly suggest that a locus influencing LDL cholesterol levels in both isolated and out bred populations is to be found.

Most obvious candidate genes that might explain the linkage on chromosome 19 are the *LDLR* gene, the *APOE/C1/C4/C2* gene cluster and *LRP3* gene. The gene product of the *LDLR* gene and the *APOE/C1/C4/C2* gene cluster are known to play a role in lipid metabolism and the *LRP3* gene is a member of the *LDLR* gene family, which leads to the speculation of a role in lipid metabolism[23]. The contribution of these loci to the positive linkage obtained could be investigated by combined association and linkage approach as described in Chapter 7. Since mutations in the LDLR gene are known to cause hypercholesterolaemia, the complete coding sequence of this gene was searched for mutations in the three sib pairs contributing most to the linkage of chromosome 19 with LDL cholesterol levels. These pairs were extreme discordant for

LDL cholesterol levels. In one sib pair a rare mutation was found, which was previously identified in an English patient with mild symptoms of FH[24]. However, the mutation was found in both members of the discordant sib pair, indicating that it is not the causative mutation. In the other two pairs, no *LDLR* mutations were found.
The contribution of the APOEε2/ε3/ε4 polymorphism to LDL cholesterol levels for example was investigated in a combined association and linkage approach as described in Chapter 7. The APOEε2/ε3/ε4 polymorphism explained part of the total variance in LDL cholesterol levels, decreasing the MLS from 5.7 to 3.5 at 60 cM from pter. So, LDL cholesterol levels are influenced by the APOEε2/ε3/ε4 polymorphism, but still other genetic variants remain to be found on chromosome 19 affecting LDL cholesterol levels (data not shown).

Combined association and linkage analysis

After the fine mapping on chromosome 19, we reanalysed linkage for apoE levels at this chromosome. The MLS in adolescent Dutch twins was still 1.0 and 0.1 in the Australian twins, but in the adult Dutch it increased to 2.6. The three twin samples could not be analysed simultaneously, since the analysis showed evidence for heterogeneity (data not shown). Therefore, only the adolescent and adult Dutch samples where analysed simultaneously, showing no evidence for heterogeneity. The MLS in the Dutch simultaneous analysis increased to 3.1 at the *APOE* locus, indicating suggestive linkage with apoE levels.
We then investigated what proportion of the linkage result of the apoE levels could be explained by the *APOE*ε2/ε3/ε4 polymorphism. A combined linkage-association approach was used. The combined association and linkage study in the two Dutch samples, showed that the linkage in the adolescent Dutch twins could entirely be explained by the APOEε2/ε3/ε4 polymorphism, indicating that probably no other polymorphisms on chromosome 19 influences apoE levels in adolescent Dutch twins. Only polymorphisms in complete linkage disequilibrium (LD) with this APOEε2/ε3/ε4 polymorphism can not be excluded to play a role in plasma levels of apoE. In the adult Dutch twins, however, only part of linkage could be explained by the APOEε2/ε3/ε4 polymorphism, indicating that other genetic variation in the QTL region influences apoE levels. Other polymorphisms that might play a role in determining apoE levels are, for example, the functional *APOE* promoter variants -291G/T and -491A/T[25].
Although the Australian twins did not show any linkage, association of the APOEε2/ε3/ε4 polymorphism with apoE plasma levels could be detected. This illustrates well that association studies have more power than linkage studies. However, this polymorphism explained merely 5% of the variance in Australian apoE levels, which is much less than known from literature. These results in the Australian sample, together with the much lower heritability of apoE levels (57% vs 86-87% Dutch samples), might suggest that the apoE measurements may not be accurate.
Additionally, the finding that the *APOE*ε2/ε3/ε4 polymorphism is the only variant in the APOE gene influencing apoE levels in adolescents, not taking into account variants in complete LD, is different from the findings in adults. Thus, the genetic architecture of adolescent apoE levels seems to differ from that of adult levels. This is consistent with a previous quantitative genetic studies, describing that some lipid and apolipoprotein levels are influenced by partly different genes in adolescence and adulthood[26], which might also be the case for apoE levels.

Conclusion and Future perspectives

The studies in this thesis indicate that lipid and apolipoprotein levels are highly heritable in Caucasian samples and that thus the genes involved in these traits have to be identified. We were able to localise several putative QTLs for lipid and apolipoprotein levels and we identified one major LDL cholesterol QTL on chromosome 19, probably also influencing apoB levels. It is expected that in our approach, QTLs with smaller effect have been missed. Since high throughput genotypings are now feasible, the remaining chromosomes will be scanned with a 10 cM inter-marker distance. In the fine mapping of positive linkage regions, we have applied several statistic tools of which the combined association and linkage approach seems to be the most valuable tool. Using this approach, genes might be identified to have their influence on quantitative parameters of the human lipid metabolism.

References

1. Khot UN, Khot MB, Bajzer CT, Sapp SK, Ohman EM *et al.* Prevalence of conventional risk factors in patients with coronary heart disease. *JAMA* 2003; 290: 898-904.
2. Greenland P, Knoll MD, Stamler J, Neaton JD, Dyer AR *et al.* Major risk factors as antecedents of fatal and nonfatal coronary heart disease events. *JAMA* 2003; 290: 891-897.
3. Lander E, Kruglyak L. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat Genet* 1995; 11: 241-247.
4. Altmuller J, Palmer LJ, Fischer G, Scherb H, Wjst M. Genomewide scans of complex human diseases: true linkage is hard to find. *Am J Hum Genet* 2001; 69: 936-950.
5. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. GRR: graphical representation of relationship errors. *Bioinformatics* 2001; 17: 742-743.
6. Douglas JA, Boehnke M, Lange K. A multipoint method for detecting genotyping errors and mutations in sibling-pair linkage data. *Am J Hum Genet* 2000; 66: 1287-1297.
7. Larson IA, Ordovas JM, DeLuca C, Barnard JR, Feussner G, Schaefer EJ. Association of apolipoprotein (Apo)E genotype with plasma apo E levels. *Atherosclerosis* 2000; 148: 327-335.
8. Lambert JC, Brousseau T, Defosse V, Evans A, Arveiler D *et al.* Independent association of an APOE gene promoter polymorphism with increased risk of myocardial infarction and decreased APOE plasma concentrations-the ECTIM study. *Hum Mol Genet* 2000; 9: 57-61.
9. Kaprio J, Ferrell RE, Kottke BA, Kamboh MI, Sing CF. Effects of polymorphisms in apolipoproteins E, A-IV, and H on quantitative traits related to risk for cardiovascular disease. *Arterioscler Thromb* 1991; 11: 1330-1348.
10. Wiltshire S, Cardon LR, McCarthy MI. Evaluating the results of genomewide linkage scans of complex traits by locus counting. *Am J Hum Genet* 2002; 71: 1175-1182.
11. Broeckel U, Hengstenberg C, Mayer B, Holmer S, Martin LJ *et al.* A comprehensive linkage analysis for myocardial infarction and its related risk factors. *Nat Genet* 2002; 30: 210-214.
12. Göring HH, Terwilliger JD, Blangero J. Large upward bias in estimation of locus-specific effects from genomewide scans. *Am J Hum Genet* 2001; 69: 1357-1369.
13. Imperatore G, Knowler WC, Pettitt DJ, Kobes S, Fuller JH *et al.* A locus influencing total serum cholesterol on chromosome 19p: results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arterioscler Thromb Vasc Biol* 2000; 20: 2651-2656.
14. Ober C, Abney M, McPeek MS. The genetic dissection of complex traits in a founder population. *Am J Hum Genet* 2001; 69: 1068-1079.
15. Barton NH, Keightley PD. Understanding quantitative genetic variation. *Nat Rev Genet* 2002; 3: 11-21.
16. Steinmetz LM, Sinha H, Richards DR, Spiegelman JI, Oefner PJ *et al.* Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 2002; 416: 326-330.
17. Rainwater DL, Almasy L, Blangero J, Cole SA, VandeBerg JL *et al.* A Genome Search Identifies Major Quantitative Trait Loci on Human Chromosomes 3 and 4 That Influence Cholesterol Concentrations in Small LDL Particles. *Arterioscler Thromb Vasc Biol* 1999; 19: 777-783.
18. Klos KL, Kardia SL, Ferrell RE, Turner ST, Boerwinkle E, Sing CF. Genome-wide linkage analysis reveals evidence of multiple regions that influence variation in plasma lipid and apolipoprotein levels associated with risk of coronary heart disease. *Arterioscler Thromb Vasc Biol* 2001; 21: 971-978.
19. Reed DR, Nanthakumar E, North M, Bell C, Price RA. A genome-wide scan suggests a locus on chromosome 1q21-q23 contributes to normal variation in plasma cholesterol concentration. *J Mol Med* 2001; 79: 262-269.
20. Aouizerat BE, Allayee H, Cantor RM, Davis RC, Lanning CD *et al.* A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am J Hum Genet* 1999; 65: 397-412.

21. Francke S, Manraj M, Lacquemant C, Lecoeur C, Lepretre F *et al.* A genome-wide scan for coronary heart disease suggests in Indo-Mauritians a susceptibility locus on chromosome 16p13 and replicates linkage with the metabolic syndrome on 3q27. *Hum Mol Genet* 2001; 10: 2751-2765.
22. Elbein SC, Hasstedt SJ. Quantitative trait linkage analysis of lipid-related traits in familial type 2 diabetes: evidence for linkage of triglyceride levels to chromosome 19q. *Diabetes* 2002; 51: 528-535.
23. Ishii H, Kim DH, Fujita T, Endo Y, Saeki S, Yamamoto TT. cDNA cloning of a new low-density lipoprotein receptor-related protein and mapping of its gene (LRP3) to chromosome bands 19q12-q13. 2. *Genomics* 1998; 51: 132-135.
24. Sun XM, Patel DD, Knight BL, Soutar AK. Comparison of the genetic defect with LDL-receptor activity in cultured cells from patients with a clinical diagnosis of heterozygous familial hypercholesterolemia. The Familial Hypercholesterolaemia Regression Study Group. *Arterioscler Thromb Vasc Biol* 1997; 17: 3092-3101.
25. Stengard JH, Clark AG, Weiss KM, Kardia S, Nickerson DA *et al.* Contributions of 18 additional DNA sequence variations in the gene encoding apolipoprotein e to explaining variation in quantitative measures of lipid metabolism. *Am J Hum Genet* 2002; 71: 501-517.
26. Snieder H, van Doornen LJ, Boomsma DI. The age dependency of gene expression for plasma lipids, lipoproteins, and apolipoproteins. *Am J Hum Genet* 1997; 60: 638-650.

9 | Samenvatting

Inleiding

Hart- en vaatziekten zijn de meest voorkomende doodsoorzaak in Westerse landen. De hoofdoorzaak van hart- en vaatziekten is het ophopen van vet in de wand van slagaders in combinatie met een ontstekingsreactie. Dit proces wordt atherosclerose genoemd, ook wel aderverkalking. Omdat de vetten cholesterol en triglyceriden, ook lipiden genoemd, niet oplosbaar zijn in water, worden ze door het bloed getransporteerd in bolletjes die lipoproteïnen worden genoemd. Er zijn lipoproteïnen die lipiden van de lever naar de spieren en organen brengen (*Very Low Density Lipoproteins* en *Low Density Lipoproteins*) en lipoproteïnen die overtollige lipiden weer terug brengen naar de lever (*High Density Lipoproteins*). Op de oppervlakte van de lipoproteïnen zijn eiwitten aanwezig die zorgen dat de verschillende soorten lipoproteïnen op de juiste bestemming aankomen. Deze eiwitten noemen we apolipoproteïnen. Op de VLDL- en LDL-deeltjes zijn als belangrijkste wegwijzers de apolipoproteïnen B en E aanwezig en op de HDL-deeltjes de apolipoproteïnen A1 en A2. De voornaamste risicofactoren voor de ontwikkeling van aderverkalking zijn verhoogde bloedplasmaspiegels van de lipiden LDL-cholesterol, totaal-cholesterol, triglyceriden, en de apolipoproteïnen B en E, en verlaagde bloedplasmaspiegels van het lipide HDL-cholesterol en de apolipoproteïnen A1 en A2. Deze risicofactoren zijn een voorbode van hart- en vaatziekten en representeren het vetmetabolisme als één van de betrokken processen die leiden tot hart- en vaatziekten. Om meer inzicht te krijgen in welke rol genen spelen bij het ontstaan van hart- en vaatziekten, zijn lipiden- en apolipoproteïnen-spiegels onderzocht als risicofactoren voor hart- en vaatziekten. De studies beschreven in dit proefschrift zijn gericht op het vinden van genen die een belangrijke rol spelen in het bepalen van de lipiden- en apolipoproteïnen-bloedplasmaspiegels in de algemene bevolking.

In deze studies zijn vier tweelingpopulaties onderzocht; één populatie van adolescente Nederlandse tweelingen, één van volwassen Nederlandse tweelingen, één van volwassen Zweedse tweelingen en één van volwassen Australische tweelingen. Deze vier populaties kunnen worden gebuikt om te onderzoeken of een resultaat gerepliceerd kan worden en of in elk van de tweelingpopulaties dezelfde genen een rol spelen bij de variatie in lipiden- en apolipoproteïnen-spiegels.

Erfelijkheid van plasma spiegels van cholesterol en apolipoproteïnen

Om de grootte van de invloed van genen en omgeving op de variatie in bloedplasmaspiegels van lipiden en apolipoproteïnen te onderzoeken, is de erfelijkheid van deze eigenschappen geschat in de vier tweelingpopulaties (Hoofdstuk 2). Dit kan door de gelijkheid van de spiegels van eeneiige en twee-eiige tweelingen met elkaar te vergelijken. Eeneiige tweelingen hebben identiek DNA materiaal en twee-eiige tweelingen hebben gemiddeld de helft van hun DNA materiaal gemeenschappelijk. Wanneer de bloedplasmaspiegels van eeneiige tweelingen meer gelijk zijn dan de bloedplasmaspiegels van twee-eiige tweelingen, dan mag geconcludeerd worden dat bij twee-eiige tweelingen meer verschil in de hoogte van de spiegels is, omdat er meer verschil in het DNA materiaal is. Dit houdt in dat variatie in genen bijdraagt aan de hoogte van de lipiden- en apolipoproteïnen-spiegels, dus dat die erfelijk is. Wij vonden dat tussen de 48 en 87% van de variatie in bloedplasmaspiegels van lipiden en apolipoproteïnen wordt veroorzaakt door genetische factoren. Ook de erfelijkheid van apolipoproteïne E en apolipoproteïne A2 spiegels, die tot nu toe slechts zelden zijn beschreven, is vrij hoog, net zoals dat van de andere cholesterol en apolipoproteïnen spiegels. Familieomgeving, zoals bijvoorbeeld het voedingspatroon of sociaal economische klasse, is niet van invloed op deze risicofactoren voor hart- en vaatziekten en de erfelijkheid van de spiegels is niet verschillend voor mannen en vrouwen, maar is wel verschillend in de vier landen. In de Nederlandse tweelingen is de erfelijkheid het hoogst en in de Australische tweelingen het laagst. Het verschil in erfelijkheid van de cholesterol en apolipoproteïnen spiegels zou kunnen komen doordat de invloed van de omgeving, zoals bijvoorbeeld eetpatronen, groter is in Australië of doordat verschillende factoren uit de omgeving een interactie aangaan met de genen.

In het algemeen blijkt echter dat de variatie in de lipiden- en apolipoproteïnen-spiegels tussen mensen voor een heel groot deel bepaald wordt door de variatie in genen. De tweelingpopulaties uit drie verschillende landen vormen uniek materiaal om te onderzoeken welke genen invloed hebben op de hoogte van de lipiden- en apolipoproteïnen-spiegels.

Protocol voor het scannen van het totale genoom

Erfelijkheid speelt dus een belangrijke rol in het bepalen van de plasmaspiegels van cholesterol en apolipoproteïnen, maar welke genen deze rol precies vervullen is nog onduidelijk. Er zijn ongeveer 30.000 genen aanwezig in het menselijk genoom, dat is het complete menselijke DNA. Het zou te veel moeite kosten om al die genen één voor één te onderzoeken en daarom zijn we begonnen met een systematische scan van het hele genoom. In een totale genoomscan kunnen chromosomale gebieden gevonden worden waar genen liggen die bepalend zijn voor lipiden- en apolipoproteïnen-spiegels.

Voor een genoomscan is het nodig dat wordt onderzocht welke chromosomale gebieden twee-eiige tweelingen gemeenschappelijk hebben geërfd van hun ouders. Twee-eiige tweelingen hebben net als een gewoon broer-broer, broer-zus en zus-zus paar gemiddeld de helft van hun genetisch materiaal gemeenschappelijk. Maar op specifieke delen van een chromosoom kunnen ze genetisch helemaal gelijk, ongelijk of voor de helft gelijk zijn. Om de mate van genetische gelijkheid op specifieke chromosoom gebieden te onderzoeken, worden op regelmatige afstand van elkaar genetische markers gemeten. Genetische markers zijn stukjes DNA die van lengte kunnen variëren zonder dat het een schadelijk uitwerking heeft. Wanneer een tweeling een genetische marker van dezelfde lengte heeft, dan is het met grote zekerheid te zeggen dat ze dat

stukje DNA van dezelfde ouder geërfd hebben, dus gemeenschappelijk hebben. Meer over het principe van de genetische analyse volgt later in deze samenvatting.
Normaal gesproken worden in een totale genoomscan 400 van dit soort DNA stukjes bekeken, waarvan de onderlinge afstand ongeveer 10 centiMorgan is. De centiMorgan is een maat voor genetische afstand en 1 centiMorgan komt overeen met ongeveer 1 miljoen basenparen, de bouwstenen van het DNA. Aangezien binnen onze studie 1000 individuen moeten worden gescand, zouden 400.000 bepalingen nodig zijn. Daarom is onderzocht wat een optimale onderlinge afstand van de markers is, om te komen tot een optimale balans tussen de hoeveelheid metingen en de kans op het lokaliseren van genen met een belangrijk effect op de plasmaspiegels van lipiden en apolipoproteïnen (Hoofdstuk 3).
Uit berekeningen bleek dat als genetische markers een onderlinge afstand hebben van ongeveer 20 centiMorgan, de helft van het aantal bepalingen nodig is, terwijl de kans op het vinden van een interessant chromosomaal gebied toch nog voldoende hoog is. Vandaar dat we een aangepast protocol hebben ontwikkeld om sneller het genoom af te zoeken naar genen. We hebben 229 genetische markers geselecteerd met een gemiddelde onderlinge afstand van 18.3 centiMorgan. Het protocol is zo aangepast dat ongeveer 3 genetische markers in dezelfde reactie kunnen worden gemeten, zodat slecht 2.5 µg DNA nodig is om het hele genoom te onderzoeken, dat uitzonderlijk weinig is voor een totale genoom scan.
Door nieuwe geautomatiseerde technieken is de snelheid van de metingen nu al veel hoger dan ten tijde van deze studie. Met dezelfde inspanning zouden we tegenwoordig een genoomscan kunnen uitvoeren met genetische markers met een onderlinge afstand van ongeveer 5 tot 10 cM.

Koppelingsanalyse met lipiden- en apolipoproteïnen-spiegels

In de vier tweelingpopulaties zijn DNA markers gemeten op de chromosomen 1, 2, 6, 7, 8, 11, 15, 16, 17 en 19 met een onderlinge afstand van gemiddeld 18.3 centiMorgan, zoals beschreven in hoofdstuk 3. Genetische analyses zijn gedaan op 83 adolescente Nederlandse twee-eiige tweelingen, 117 volwassen Nederlandse twee-eiige tweelingen, 44 volwassen Zweedse twee-eiige tweelingen en 249 volwassen Australische twee-eiige tweelingen. Met gebruik van deze DNA markerdata en de lipiden- en apolipoproteïnen-spiegels in de tweelingpopulaties kan onderzocht worden of de lokatie bepaald kan worden van genen die en effect hebben op één van de lipiden- of apolipoproteïnen-spiegels.
Het principe van koppelingsanalyse is als volgt; wanneer variatie in een bepaald gen zorgt voor variatie in bijvoorbeeld LDL-cholesterolspiegels, dan zullen twee-eiige tweelingen die dezelfde variant van dat gen geërfd hebben van hun ouders, een meer gelijke LDL-cholesterolspiegel hebben dan wanneer ze beiden een andere genvariant geërfd hebben. Door te bepalen welke stukjes DNA de tweelingen gemeenschappelijk hebben geërfd, hopen we een deel van het menselijk genoom te kunnen identificeren waar twee-eiige tweelingen met een vergelijkbaar LDL-cholesterolspiegel hetzelfde genetische materiaal dragen. Of omgekeerd: een deel van het menselijk genoom waar twee-eiige tweelingen met verschillen in de LDL-cholesterolspiegel van elkaar verschillen in genetische materiaal. Van elk chromosomaal gebied kan worden berekend hoe groot de waarschijnlijkheid is dat het een gen bevat dat één van de bloedplasmaspiegels beïnvloedt. Deze waarschijnlijkheid wordt uitgedrukt in een LOD score. Een LOD score van 1 geeft weer dat er een aanwijzing is dat in dat gebied een gen ligt dat een rol speelt in de lipiden- en apolipoproteïnen-spiegels. Een LOD score van 2.2 geeft weer dat er een suggestie voor bewijs is en een LOD van 3.6 wordt gezien als bewijs dat in dat chromosomale gebied een gen ligt dat

bepalend is voor de bloedplasmaspiegels. De koppelingsanalyses zijn gedaan voor de chromosomen 1, 2, 6, 7, 8, 11, 15, 16, 17 en 19 met bloedplasmaspiegels van lipiden- en apolipoproteïnen-spiegels in alle vier de tweelingpopulaties (Hoofdstuk 4).

Om te bevestigen dat de genoomscan met markers met een gemiddelde onderlinge afstand van 18.3 centiMorgan tot betrouwbare resultaten kan leiden, is de lange arm van chromosoom 6, waar het *LPA* gen ligt, onderzocht. Van dit gen is bekend dat het ongeveer 90% van de variatie in de spiegels van lipoproteïne(a) bepaalt. Lipoproteïne(a) is een LDL-deeltje met een extra apolipoproteïne op de oppervlakte, namelijk apolipoproteïne(a). In alle vier de populaties is een koppeling gevonden tussen chromosoom 6 en lipoproteïne(a) spiegels. In de adolescente Nederlandse en de Zweedse tweelingen vonden we een aanwijzing voor koppeling en in zowel de volwassen Nederlandse als de Australische tweelingen vonden we zelfs sterk bewijs voor koppeling tussen chromosoom 6 en lipoproteïne(a) spiegels. Het feit dat hetzelfde resultaat gevonden wordt in onafhankelijke populaties, is een zeer sterk bewijs dat op deze lokatie een gen moet liggen dat lipoproteïne(a) spiegels beïnvloedt. Wanneer alle vier de populaties tezamen worden geanalyseerd, dan vinden we een LOD score van 9.8, wat aangeeft dat we heel sterk bewijs hebben dat hier een gen ligt dat lipoproteïne(a) spiegels beïnvloedt, dan dat het gen lipoproteïne(a) spiegels niet beïnvloedt.

Om te onderzoeken of we met het door ons ontwikkelde genoomscan protocol, een chromosomaal gebied kunnen vinden waar een gen zou moeten liggen met een niet al te groot effect, testten we het gebied waar het *APOE* gen ligt. Het *APOE* gen is het gen dat codeert voor apolipoproteïne E en is gelokaliseerd op de lange arm van chromosoom 19. Tussen de 9 en 20% van de variatie in apolipoproteïne E spiegels is toe te schrijven aan genetische variatie in het *APOE* gen. In dit gebied vonden we een aanwijzing voor koppeling in de adolescente en de volwassen Nederlandse tweelingen, hoewel minder sterk dan bij het *LPA* gen. We vonden geen bewijs voor koppeling in de Australische tweelingen. Koppelingsanalyse kon niet worden uitgevoerd in de Zweedse tweelingpopulatie, omdat daarin geen apolipoproteïne E spiegels zijn gemeten. Wanneer de drie tweelingpopulaties tezamen werden geanalyseerd, was de LOD score 1.0, een score is die per toeval ongeveer 8 keer zou kunnen voorkomen in een totale genoom scan. Hieruit concluderen we dat we niet in staat zijn om het *APOE* gen te detecteren in de drie tweelingpopulaties als een gen dat een matig effect heeft op apolipoproteïne E spiegels. Dus gebruik makend van ons protocol voor een genoomscan met markers met een onderlinge afstand van 18.3 centiMorgan, zijn we alleen in staat om genen met een groot effect op de bloedspiegels te detecteren en zullen we genen met een klein effect mogelijk niet kunnen herkennen.

Na de koppelingsanalyse van de eerste 10 chromosomen vonden we 11 gebieden waar mogelijk een gen zou kunnen liggen dat betrokken is bij één van de cholesterol of apolipoproteïne spiegels; 4 chromosomale gebieden lijken betrokken bij lipoproteïne(a) spiegels, 3 gebieden bij apolipoproteïne A1 spiegels, 2 gebieden bij apolipoproteïne B spiegels en 2 gebieden bij LDL-cholesterolspiegels. Van deze 11 gebieden, zijn 3 gebieden in meer dan 1 tweelingpopulaties gevonden. Voor 2 gebieden op chromosoom 1 en 2 was een suggestie voor bewijs voor koppeling met lipoproteïne(a) spiegels in zowel de Zweedse als de Australische tweelingpopulatie en voor 1 gebied op chromosoom 19 vonden we een suggestie bewijs voor koppeling met LDL-cholesterolspiegels in de volwassen Nederlandse en een aanwijzing voor koppeling in de Australische tweelingpopulatie. Hoewel de erfelijkheid van totaal cholesterol, HDL-cholesterol, triglyceriden en apolipoproteïne E spiegels heel erg hoog is, vonden we geen bewijs voor koppeling in deze 10 chromosomen. Dit zou kunnen betekenen dat de genen die een

groot effect op deze bloedplasmaspiegels hebben op andere chromosomen liggen, of dat veel genen met kleine effecten invloed hebben op deze bloedplasmaspiegels, die we door onze onderzoeksbenadering mogelijk niet hebben kunnen detecteren.

Gerepliceerde koppeling met lipoproteïne(a) spiegels

Voor het gebied op chromosoom 2 vonden wij bewijs voor koppeling met lipoproteïne(a) spiegels in de Zweedse en Australische tweelingpopulaties. In dit gekoppelde chromosomale gebied is het *APOB* gen gelokaliseerd, dat codeert voor apolipoproteïne B, een eiwit dat een belangrijk onderdeel vormt van het lipoproteïne(a) deeltje. Het zou zo kunnen zijn dat afwijkend apolipoproteïne B een effect heeft op lipoproteïne(a) spiegels. Het feit dat in de twee Nederlandse tweelingpopulaties totaal geen bewijs voor koppeling van chromosoom 2 met lipoproteïne(a) spiegels is gevonden, maakt het *APOB* gen echter een stuk minder interessant als kandidaat-gen voor het bepalen van de variatie in lipoproteïne(a) spiegels.

Het gebied op chromosoom 1 dat gekoppeld lijkt te zijn met lipoproteïne(a) spiegels vinden we ook in slechts twee van onze tweelingpopulaties, maar Broeckel *et al.* vond in zijn populatie ook bewijs voor de koppeling van dit gebied met lipoproteïne(a) spiegels. Broeckel concludeerde dat op chromosoom 1 een gen moest liggen dat een effect heeft op lipoproteïne(a) spiegels onafhankelijk van het grote effect van het bekende *LPA* gen op de lange arm van chromosoom 6. Om te onderzoeken of inderdaad het gen op chromosoom 1 en het *LPA* gen beiden effect hebben op lipoproteïne(a) spiegels, zijn de effecten tegelijkertijd in één analyse bestudeerd. Zo'n analyse is nog nauwelijks toegepast in de wetenschappelijke literatuur. De analyse van chromosoom 1 voor de vier tweelingpopulaties tezamen resulteerde in een LOD score van 1.6. Het gen dat mogelijk betrokken is bij lipoproteïne(a) spiegels zou 44% van de variatie in lipoproteïne(a) spiegels kunnen verklaren, hoewel het *LPA* gen op chromosoom 6 zelf al 82% van de variatie in lipoproteïne(a) spiegels verklaart. Het is een bekend verschijnsel dat als chromosomale gebieden apart worden geanalyseerd, het cumulatieve effect van de gebieden meer dan 100% kan zijn. Daarom is het nodig om in één analyse de twee gebieden op chromosoom 1 en 6 tegelijkertijd te analyseren, zodat de gebieden tezamen maximaal 100% van de variatie kunnen verklaren (Hoofdstuk 5). Het resultaat van deze zogenaamde twee-locus analyse was dat 82% van de variatie in lipoproteïne(a) spiegels te verklaren is door de genetische variatie in het *LPA* gen op chromosoom 6 en dat de koppeling op chromosoom 1 een vals positief resultaat was in onze tweelingpopulaties. Los van de uitkomst geeft dit resultaat een indicatie dat met behulp van een twee-locus analyse vals positieve resultaten gedetecteerd zouden kunnen worden.

Fijn kartering van chromosoom 19

Op chromosoom 19 was een suggestie voor bewijs voor koppeling gevonden met LDL-cholesterolspiegels in de volwassen Nederlandse en een aanwijzing voor koppeling in de Australische en Zweedse tweelingen. Deze bevindingen in drie onafhankelijke populaties maken het gebied op chromosoom 19 aantrekkelijk om nader te onderzoeken.

Van de 23 chromosomen zijn nu 10 chromosomen bestudeerd en het blijkt dat een breed gebied op chromosoom 19 een belangrijke rol speelt in het bepalen van LDL-cholesterolspiegels. In een poging om het gebied te verkleinen en om de betrouwbaarheid van de resultaten te toetsen, zijn extra DNA markers getypeerd in alle vier de tweelingpopulaties. Uiteindelijk is de gemiddelde onderlinge afstand tussen de markers 6 centiMorgan in de twee Nederlandse populaties geworden en 8 centiMorgan in de Zweedse en Australische populaties. Het bewijs voor koppeling werd

door deze extra DNA-markerinformatie aanzienlijk groter in de volwassen Nederlandse, Zweedse en Australische tweelingpopulaties. In de adolescente Nederlandse tweelingen vonden we echter geen bewijs voor koppeling tussen chromosoom 19 en LDL cholesterolspiegels (Hoofdstuk 6). Hoewel het bewijs voor koppeling versterkt is, is het koppelingsgebied niet smaller geworden door het typeren van extra DNA markers. Dit kan verklaard worden door het feit dat wanneer twee-eiige tweelingen eenmaal een stuk genetische materiaal gemeenschappelijk hebben geërfd, het dan direct een groot stuk DNA is. Wanneer meerdere tweelingparen in een koppelingsanalyse worden betrokken, dan is het koppelingsgebied het overlappende gebied van gemeenschappelijke gebieden in alle paren. Hoe meer tweelingparen, des te kleiner is het overlappende gebied van alle gemeenschappelijke chromosomale gebieden. Een koppelingsgebied als resultaat van een tweelingstudie, kan dus niet smaller worden gemaakt door het typeren van extra genetische markers, wat ten tijde van deze studie nog niet duidelijk was. Daarbij zou een breed koppelingsgebied verklaard kunnen worden doordat koppelingsresultaten mogelijk het opgetelde resultaat zijn van de effecten van meerdere genen in het koppelingsgebied. Het resultaat wordt dan niet veroorzaakt wordt door het grote effect van één gen, zoals een koppelingsresultaat gewoonlijk geïnterpreteerd wordt.
De analyse van de drie volwassen tweelingpopulaties tezamen resulteert in een LOD score van 5.7 op 60 centiMorgan vanaf de top van chromosoom 19. Dit is zeer sterk bewijs dat in dit gebied een gen ligt dat invloed heeft op LDL-cholesterolspiegels dan dat het een vals positief resultaat is. In de Nederlandse, Zweedse en Australische tweelingen zou de genetische variatie van dit gen meer dan de helft van de variatie in LDL-cholesterolspiegels verklaren. Op chromosoom 19 ligt dus een gen met een grote effect op LDL-cholesterolspiegels. Omdat apolipoproteïne B een onderdeel vormt van het LDL-deeltje, zijn apolipoproteïne B spiegels sterk gecorreleerd met LDL cholesterolspiegels. Om te onderzoeken of dit gen ook invloed heeft op apolipoproteïne B spiegels, hebben we een analyse gedaan waarbij zowel het effect op LDL cholesterol spiegels als op apolipoproteïne B spiegels wordt bekeken. De analyse bevestigde dat dit gen ook een effect heeft op apolipoproteïne B spiegels.
Eerder is door anderen ook bewijs voor koppeling tussen chromosoom 19 en totaal cholesterol en LDL-cholesterolspiegels gevonden in respectievelijk Pima Indianen en Hutterieten. Dit zijn twee genetisch geïsoleerde bevolkingsgroepen uit Noord Amerika. Pima Indianen leiden een boeren bestaan en Hutterieten leven in streng gelovige gemeenschappen, zonder radio of televisie. Door hun specifieke leefstijlen mengen deze bevolkingsgroepen niet snel met andere bevolkingsgroepen. De resultaten uit onze studie vormen een belangrijke aanwijzing dat het chromosomale gebied op chromosoom 19 niet alleen in zulke beperkte bevolkingsgroepen een rol speelt, maar ook bij het grote aantal mensen van Europese afkomst. Andere studies die ook de niet-geïsoleerde bevolking onderzochten vonden ook aanwijzingen voor koppeling van chromosoom 19 met totaal cholesterol of LDL-cholesterolspiegels, maar vaak waren dit niet de belangrijkste bevindingen in hun totale genoomscans.Aangezien wij en anderen steeds uitkomen op chromosoom 19, is het zeer waarschijnlijk dat daar een gen ligt dat bepalend is voor LDL-cholesterolspiegels. In het gebied liggen een aantal genen waarvan al bekend is dat ze een rol spelen bij het lipidenmetabolisme. De volgende stap is na te gaan of genetische variatie in deze genen inderdaad heeft bijgedragen aan het koppelingsresultaat. Dit kan op de manier zoals is beschreven in hoofdstuk 7 door middel van gecombineerde associatie- en koppelingsanalyse.
Op chromosoom 19 liggen een aantal genen die een kandidaat zijn voor het beïnvloeden van LDL-cholesterolspiegels. De meest opvallende kandidaat-genen zijn de genen die coderen voor

de LDL receptor en het LDL receptor gerelateerd eiwit en ook het genencluster dat codeert voor de apolipoproteïnen E, C1, C4 en C2. Van deze genproducten, behalve van het LDL receptor gerelateerd eiwit, is het bekend dat ze een rol spelen in het vetmetabolisme. Omdat echter het LDL receptor gerelateerd eiwit uit dezelfde genfamilie komt als de LDL receptor, wordt gespeculeerd dat het ook een rol speelt binnen het vetmetabolisme.
De bijdrage van deze genen aan het koppelingsresultaat zou kunnen worden onderzocht met behulp van een gecombineerde associatie en koppelingsanalyse zoals die beschreven is in Hoofdstuk 7. Omdat het bekend is dat zeldzame mutaties in het LDL receptor gen hypercholesterolemie kunnen veroorzaken, hebben we de drie tweeling paren die het meest bijdroegen aan het koppelingsresultaat onderzocht op mutaties in dit gen. In deze paren had één helft van de tweeling een heel hoge LDL cholesterolspiegel en de andere helft een heel lage LDL cholesterolspiegel. In één tweeling met een hoge LDL-cholesterolspiegel vonden we inderdaad een mutatie, die één keer eerder is gevonden in een Engelse patiënt met een matig verhoogde LDL-cholesterolspiegel. Het bleek echter dat zijn tweelingbroer met een lage LDL-cholesterolspiegel ook drager was van deze mutatie. Deze mutatie kon dus niet de oorzaak zijn van het verschil in LDL-cholesterolspiegel binnen deze tweeling. Zeldzame mutaties in het LDL receptor gen spelen dus geen rol in de variatie in LDL-cholesterolspiegels.
De bijdrage van de *APOEε2/ε3/ε4* genvarianten aan het koppelingsresultaat met LDL cholesterolspiegels is inmiddels onderzocht in een gecombineerde associatie- en koppelinganalyse, zoals beschreven in hoofdstuk 7. Deze genvarianten verklaarden inderdaad een deel van de koppeling tussen chromosoom 19 en LDL-cholesterolspiegels. LDL cholesterolspiegels worden dus voor een deel beïnvloed door de *APOEε2/ε3/ε4* genvarianten, maar op chromosoom 19 liggen nog onbekende genen die het grootste deel van het effect op LDL-cholesterolspiegels verklaren.

Gecombineerde associatie- en koppelingsanalyse

Na de fijn-kartering van chromosoom 19 is opnieuw de koppelingsanalyse gedaan van chromosoom 19 en apolipoproteïne E spiegels. In de adolescente Nederlandse tweelingen en in de Australische tweelingen vonden we hetzelfde resultaat als voor de fijn-kartering, maar het bewijs voor koppeling van het gebied waar het *APOE* gen is gelokaliseerd was in de volwassen Nederlandse tweelingen flink sterker geworden. Het is bekend dat variatie in het *APOE* gen de variatie in apolipoproteïne E spiegels beïnvloedt. Een bekende variatie in het *APOE* gen wordt ook wel de *APOEε2/ε3/ε4* variatie genoemd, dat 3 versies van het apolipoproteïne E onderscheidt. Om te onderzoeken in welke mate deze variatie in het *APOE* gen bijdraagt aan het koppelingsresultaat, kan een gecombineerde associatie- en koppelingsanalyse gedaan worden. We hebben deze methode toegepast op de Nederlandse tweelingpopulaties, waar we koppeling vonden tussen chromosoom 19 en apolipoproteïne E spiegels (Hoofdstuk 7).
Eerst wordt met behulp van associatieanalyse onderzocht of één van de *APOEε2/ε3/ε4* genvarianten van invloed is op de hoogte van apolipoproteïne E spiegels. Het blijkt dat de dragers van de *APOEε2* vorm van apolipoproteïne E hogere apolipoproteïne E spiegels hebben dan dragers van de *APOEε3* vorm. De dragers van de *APOEε4* vorm hebben juist lagere apolipoproteïne E spiegels dan de dragers van de *APOEε3* vorm. De *APOEε2/ε3/ε4* genvariatie is dus geassocieerd met apolipoproteïne E spiegels. Vervolgens is in een gecombineerd associatie- en koppelingsanalyse onderzocht in welke mate dit *APOEε2/ε3/ε4* effect bijdraagt aan de koppeling tussen chromosoom 19 en apolipoproteïne E spiegels.

Wanneer in de adolescente tweelingen het effect van de associatie verdisconteerd wordt in de koppelingsanalyse dan verdwijnt de koppeling in zijn geheel. Dit houdt in dat geen andere genetische variatie op chromosoom 19 invloed heeft op apolipoproteïne E spiegels in adolescente Nederlanders. In de volwassen tweelingen ligt de situatie iets anders. Wanneer het effect van de associatie verdisconteerd werd in de koppelingsanalyse, bleef een groot deel van de koppeling tussen chromosoom 19 en apolipoproteïne E spiegels over. Met andere woorden, andere genetische variatie, in of buiten het *APOE* gen, draagt ook bij aan de variatie in apolipoproteïne E spiegels in de volwassen Nederlandse tweelingpopulatie.
Het is dus zo dat de variatie in apolipoproteïne E spiegels op adolescente leeftijd door alleen de *APOEε2/ε3/ε4* genvarianten wordt beïnvloed, terwijl op latere leeftijd ook andere genvarianten een rol gaan spelen. Dit komt overeen met eerder onderzoek waarin beschreven wordt dat later in het leven andere genen gaan bijdragen aan verschillen in lipiden- en apolipoproteïnen-spiegels.

Conclusie

De studies beschreven in dit proefschrift laten zien dat cholesterol en apolipoproteïne spiegels in grote mate erfelijk zijn. Wij zijn in staat geweest om een aantal chromosomale gebieden aan te wijzen waar mogelijk één of meerdere genen liggen die een effect hebben op de spiegels van één van de bloedparameters uit het vetmetabolisme. Onze belangrijkste bevinding is dat op chromosoom 19 één of meerder genen liggen die bepalend zijn voor LDL-cholesterolspiegels. Identificatie van de onderliggende genen dankzij het toepassen van de gecombineerde associatie- en koppelingsanalyse kan uiteindelijk een belangrijke bijdrage leveren aan preventie van hart- en vaatziekten.

Abbreviations

ABC1	ATP-binding cassette transporter 1 gene
ALF*express*	Automated laser fluorescent DNA analyser
Apo	Apolipoprotein
APOA1	Apolipoprotein A1
APOB	Apolipoprotein B
APOE	Apolipoprotein E
ARH	Autosomal recessive hypercholesterolaemia
BSCL	Berardinelli-Seip syndrome
CETP	Cholesterol ester transfer protein
CI	Confidence interval
cM	CentiMorgan
CVD	Cardiovascular disease
DNA	Deoxyribonucleic acid
DOS	Dizygotic (twin pair) of opposite sex
DZ	Dizygotic (twin pair)
FCH	Familial Combined Hypercholesterolemia
FD	Familial Dysbetahyperlipidemia
FDS	Family Diabetes Study
FFA	Free fatty acids
FH	Familial Hypercholesterolemia
FHS	Family Heart Study
GRR	Graphical representation of relationship software
Gwp-value	Genome wide p-value
HDL	High-density lipoprotein
HDL-UC	HDL unesterified cholesterol
HE	Haseman-Elston (sib pair analysis)
HL	Hepatic lipase
HSPG	Heparan sulphate proteoglycan
IBD	Identical by descent
IDL	Intermediate-density lipoprotein
LCAT	Lecithin:cholesterol acyltransferase
LD	Linkage disequilibrium
LDL	Low-density lipoprotein

LDLR	Low-density lipoprotein receptor
LOD	^{10}Logarithm of the odds ratio between likelihood of linkage and the likelihood of no linkage
Lp(a)	Lipoprotein (a)
LPA	Apolipoprotein(a) gene
LPL	Lipoprotein lipase
LRP	LDLR-related protein
MLS	Maximum LOD score
mRNA	Messenger ribonucleic acid
MZ	Monozygotic (twin pair)
PCR	Polymerase chain reaction
QTL	Quantitative trait locus
r_{DZ}	The correlation of a trait in dizygotic twin pairs
r_{MZ}	The correlation of a trait in monozygotic twin pairs
SNP	Single nucleotide polymorphism
STR	Short tandem repeat
TG	Triglycerides
VLDL	Very-Low-density lipoprotein

N.B.: In this thesis, gene names are written in italic capitals, proteins are written in regular fonts.

List of Publications

Van Schothorst EM, **Beekman M**, Torremans P, Kuipers-Dijkshoorn NJ, Wessels HW, Bardoel AF, Van der Mey AG, Van der Vijver MJ, Van Ommen GJ, Devilee P, Cornelisse CJ. Paragangliomas of the head and neck region show complete loss of heterozygosity at 11q22-q23 in chief cells and the flow-sorted DNA aneuploid fraction. *Human Pathology* 1998; 29: 1045-1049.

De Maat MP, Jukema JW, Ye S, Zwinderman AH, Moghaddam PH, **Beekman M**, Kastelein JJ, Van Boven AJ, Bruschke AV, Humphries SE, Kluft C, Henney AM. Effect of the stromelysin-1 promoter on efficacy of pravastatin in coronary atherosclerosis and restenosis. *American Journal of Cardiology* 1999; 83: 852-856.

Van Schothorst EM, Prins DE, Baysal BE, **Beekman M**, Licht JD, Waxman S, Zelent A, Cornelisse CJ, Van Ommen GJ, Richard CW 3rd, Devilee P. Genomic structure of the human PLZF gene. *Gene* 1999; 236: 21-24.

Slagboom PE, Heijmans BT, **Beekman M**, Westendorp RG, Meulenbelt I. Genetics of human aging. The search for genes contributing to human longevity and diseases of the old. *Annals of the New York Academy of Science* 2000; 908: 50-63.

Beekman M, Heijmans BT, Martin NG, Pedersen NL, Whitfield JB, DeFaire U, Van Baal GCM, Snieder H, Vogler GP, Slagboom PE, Boomsma DI. Heritabilities of apolipoprotein and lipid levels in three countries. *Twin Research* 2002; 5: 87-97.

Beekman M, Lakenberg N, Cherny SS, De Knijff P, Kluft CC, Van Ommen GJB, Vogler GP, Frants RR, Boomsma DI, Slagboom PE. A powerful and rapid approach to human genome scanning in small quantities of genomic DNA. *Genetical Research* 2001; 77:129-134.

Beekman M, Heijmans BT, Martin NG, Whitfield JB, Pedersen NL, DeFaire U, Snieder H, Lakenberg N, De Knijff P, Frants RR, Van Ommen GJB, Kluft CC, Vogler GP, Slagboom PE, Boomsma DI. Two-locus linkage analysis applied to putative quantitative trait locus for lipoproten(a) levels. *Twin Research* 2003; 6: 322-324.

Beekman M, Heijmans BT, Martin NG, Whitfield JB, Pedersen NL, DeFaire U, Snieder H, Lakenberg N, Suchiman HED, De Knijff P, Frants RR, Van Ommen GJB, Kluft CC, Vogler GP, Boomsma DI, Slagboom PE. Evidence for a QTL on chromosome 19 influencing LDL cholesterol levels in the general population. *European Journal of Human Genetics* 2003; 11: 845-850.

Beekman M, Posthuma D, Heijmans BT, Lakenberg N, Suchiman HED, Snieder H, De Knijff P, Frants RR, Van Ommen GJB, Kluft CC, Vogler GP, Slagboom PE, Boomsma DI. Combined association and linkage analysis applied to the *APOE* locus. *Genetic Epidemiology* 2004: in press.

Mooijaart SP, Van Heemst D, Schreuder J, Van Gerwen S, **Beekman M**, Brandt BW, Slagboom PE, Westendorp RGJ. on behalf of the "Long Life" study group.Variation in the SHC1 gene and longevity in human. *Experimental Gerontology* 2004; in press.

Van Heemst D, Mooijaart SP, **Beekman M**, Van Gerwen S, Schreuder J, Brandt BW, Slagboom PE, Westendorp RGJ. Variation in the human TP53 gene affects old age survival and cancer mortality. *Submitted for publication.*

Suchiman HED, Heijmans BT, **Beekman M**, Meulenbelt I, Feskens EJM, Slagboom PE. Whole genome amplification of genomic DNA from buccal cells and serum – reliability of genotypes. *Submitted for publication.*

Basrak B, Klaassen C, **Beekman M**, Martin NG, Boomsma DI. Copulas in QTL mapping. *Submitted for publication*.

Curriculum Vitae

Marian Beekman werd geboren op 19 november 1972 te Maarn. In 1991 behaalde zij het gymnasium β diploma aan het Revius Lyceum te Doorn en begon zij met de studie Biomedische Wetenschappen aan de Universiteit Leiden. In het derde en vierde jaar werd drie maal 3 maanden onderzoek verricht bij de afdeling Humane Genetica onder begeleiding van dr. P. Devilee, bij de afdeling Infectieziekten onder begeleiding van dr. P. Nibbering en bij TNO Preventie en Gezondheid, afdeling Vaat- en Bindweefsel Onderzoek onder begeleiding van dr. M. de Maat. Voor haar afstudeerstage van 9 maanden keerde Marian vervolgens terug bij dr. Devilee, waar zij werkte aan het lokaliseren van het genetisch defect dat glomus tumoren veroorzaakt. Na haar afstuderen in 1996, is zij als bursaal en vervolgens vanaf mei 1998 aangesteld als assistent in opleiding bij de afdeling Humane en Klinische Genetica van de Universiteit Leiden. Het project waar zij aan kwam te werken betrof een groot samenwerkingsverband tussen het *Center for Developmental and Health Genetics*, Pennsylvania (Vereningde Staten van America), *Karolinska Institutet*, Stockholm (Zweden), *Queensland Institute for Medical Research*, Brisbane (Australië), TNO Preventie en Gezondheid, Leiden (Nederland), Leids Universitair Medisch Centrum, Leiden (Nederland) en de Vrije Universiteit van Amsterdam (Nederland) gefinancierd door *National Institutes of Health* (USA) onder leiding van prof. dr. D.I. Boomsma (Biologische Psychology, VU, Amsterdam) en prof. dr. P.E. Slagboom (Moleculaire Epidemiology, LUMC, Leiden). De resultaten van dit promotieonderzoek staan beschreven in dit proefschrift. Sinds 1 oktober 2002 is zij als postdoc werkzaam bij de afdeling Moleculaire Epidemiologie van het LUMC aan het LangLeven project, gefinancierd door Senter/IOP. Binnen dit project wordt onderzocht welke biologische mechanismen een belangrijke rol spelen bij het behalen van een hoge leeftijd.

Appendix 1

The exogenous pathway

In the exogenous pathway, dietary cholesterol and TG are absorbed by the intestine and packed into large lipoproteins, called chylomicrons. These chylomicrons are very rich in TG and have as their major protein constituent apoA1, apoA4, apoB48 when secreted into the lymph. When chylomicrons enter subsequently the circulation, these particles acquire apoE, apoC1, apoC2 and apoC3. The chylomicron-TG are hydrolysed by lipoprotein lipase (LPL), which is anchored to heparan sulphate proteoglycans (HSPG) at the endothelial cells of the vascular wall, and its cofactor apoC2. The resulting FFA are used as an energy source in muscle and other peripheral tissue, or can either be stored as TG in adipose tissue. As the TG in the chylomicrons are hydrolysed, the chylomicrons becomes smaller and are called chylomicron remnants. These particles have transferred their excess of surface components, such as phospholipids, apoA1 and apoA4, to HDL, and are enriched in cholesterol and apoE, through which the particles can be rapidly taken up by the LDL receptor (LDLR) and the LDL receptor-related protein (LRP) on the liver. The liver utilises the dietary cholesterol and TG for the synthesis of bile acids and VLDL.

The endogenous pathway

In the endogenous pathway, the liver secretes cholesterol and TG packed in VLDL, that has apoB100, apoC1, apoC2, apoC3 and apoE as its major protein constituents. Similar to chylomicrons, VLDL is processed by LPL resulting in IDL particles, which can be cleared from the circulation by using apoE as ligand for either the LDLR or LRP on the liver, or can be further hydrolysed by LPL or hepatic lipase (HL) to form LDL, which mainly contains cholesterol. In the formation of LDL, the apoC1, apoC2, apoC3 and apoE disappear from the particle, while only apoB remains to serve as the ligand for LDLR so that LDL can be cleared from the circulation. In addition, lipoprotein(a) (Lp(a)) is a subclass of LDL, which has apolipoprotein(a) coupled to the apoB protein on a LDL particle.

The reverse cholesterol pathway

In the reverse cholesterol pathway, cholesterol from the peripheral tissues is transported back to the liver packed in HDL. Nascent HDL particles, produced by the liver and by lipolysis of chylomicrons and VLDL, consists primarily of phospholipids, apoA1 and apoA2. These nascent, disc-like HDL particles take up free cholesterol, which is released in the interstitial fluid by extra hepatic tissues via active cellular cholesterol efflux, in which the ATP-binding cassette transporter 1 (ABC1) plays an essential role[1,2]. Subsequently, this free cholesterol is esterified by the enzyme lecithin:cholesterol acyltransferase (LCAT) with the use of apoA1 as a cofactor, and these cholesterol esters enter the core of the HDL particle, changing it into a spherical small

HDL_3 particle. By further uptake of cholesterol, the HDL particles become larger (HDL_2). Cholesterol esters from HDL_2 can be transported to apoB containing particles (VLDL, IDL, LDL) by the enzyme cholesterol ester transfer protein (CETP), which will convert the HDL_2 back into HDL_3. HDL_2 can, otherwise, be taken up by the liver via the scavenger receptor class B type I (SR-B1)[3], which uses apoA1 as ligand. But by any further uptake of cholesterol, HDL_2 will finally convert into apoE rich HDL_1 particles, which can be cleared from the circulation by the liver by the LDLR or LRP.

References

1. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K *et al.* Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 1999; 22: 336-345.
2. Rust S, Rosier M, Funke H, Real J, Amoura Z *et al.* Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. *Nat Genet* 1999; 22: 352-355.
3. Trigatti B, Rigotti A, Krieger M. The role of the high-density lipoprotein receptor SR-BI in cholesterol metabolism. *Curr Opin Lipidol* 2000; 11: 123-131.

Appendix 2

Lipoprotein disorders

A number of genes encoding proteins involved in lipid metabolism, was identified in familial lipoprotein disorders. Mutations in the LDL receptor gene[1,2] are known to result in a) absence of mRNA or protein, b) defective transport of the protein to the cell membrane, c) defective binding of LDL particles, d) defective receptor-ligand complex internalisation, or e) defective recycling of the receptor to the cell surface, which consequently result in elevated LDL and total cholesterol levels (familial hypercholesterolaemia). One mutation is known in the *APOB* gene to affect the LDLR-binding-site[3] also resulting in elevated LDL and total cholesterol levels (familial defective apoB100). Many other mutations in the *APOB* gene, are leading to decreased synthesis or increased clearance rate of apoB-containing particles, resulting in decreased levels of LDL cholesterol and apoB (familial hypobetalipoproteinaemia)[4]. Mutations in the *APOE* gene may cause familial dysbetalipoproteinaemia, like the *APOEε2* allele of the *APOEε2/ε3/ε4* polymorphism, which has a poor binding capacity to the LDLR. Therefore, VLDL and IDL can not efficiently be cleared from the circulation, leading to elevated levels of β-VLDL, IDL, total cholesterol and TG[5,6]. Known mutations in *LPL* and *APOC2* result in inactivity or deficiency of LPL and its co-factor apoC2, leading to lower levels of LDL. Furthermore, through LPL's function in also the exogenous pathway, chylomicron-TG will then not be hydrolysed and accumulate in the circulation, which leads to an increase of TG and total cholesterol levels (familial chylomicronaemia)[7].

Defects and deficiency of ABC1 cause accumulation of cholesterol in macrophages, since cholesterol can not be actively transported through the membrane. No cholesterol can be taken up by the nascent HDL particles leading to very low plasma levels of HDL (Tangier disease)[8-10]. Also, large deletions and rearrangements in the *APOA1* gene leading to apoA1 deficiency are known to cause low plasma levels of HDL cholesterol and apoA1 as well as *LCAT* mutations causing loss of LCAT activity (Fish eye disease)[6,11]. CETP and HL deficiency are two known genetic causes of increased levels of HDL cholesterol[11-13].

Families with one of these genetic lipoprotein disorders form approximately 10% of the general population[14]. Thus, these mutations causing major changes in lipoprotein levels are relatively rare and play consequently a limited role in causing cardiovascular diseases in the population at large. Population studies should reveal whether these genes or (combinations of) other genes are major regulators of lipid and apolipoprotein levels in the population at large

References

1. Soutar AK. Familial hypercholesterolaemia and LDL receptor mutations. *J Intern Med* 1992; 231: 633-641.
2. Villeger L, Abifadel M, Allard D, Rabes JP, Thiart R *et al.* The UMD-LDLR database: additions to the software and 490 new entries to the database. *Hum Mutat* 2002; 20: 81-87.

3. Innerarity TL, Bostrom K. Mutations and variants of apolipoprotein B that affect plasma cholesterol levels. *Adv Exp Med Biol* 1991; 285: 25-31.
4. Farese RV, Jr., Linton MF, Young SG. Apolipoprotein B gene mutations affecting cholesterol levels. *J Intern Med* 1992; 231: 643-652.
5. Rall SC, Jr., Mahley RW. The role of apolipoprotein E genetic variants in lipoprotein disorders. *J Intern Med* 1992; 231: 653-659.
6. Breslow JL. Genetics of lipoprotein abnormalities associated with coronary artery disease susceptibility. *Annu Rev Genet* 2000; 34: 233-254.
7. Fojo SS, Brewer HB. Hypertriglyceridaemia due to genetic defects in lipoprotein lipase and apolipoprotein C-II. *J Intern Med* 1992; 231: 669-677.
8. Bodzioch M, Orso E, Klucken J, Langmann T, Bottcher A *et al.* The gene encoding ATP-binding cassette transporter 1 is mutated in Tangier disease. *Nat Genet* 1999; 22: 347-351.
9. Brooks-Wilson A, Marcil M, Clee SM, Zhang LH, Roomp K *et al.* Mutations in ABC1 in Tangier disease and familial high-density lipoprotein deficiency. *Nat Genet* 1999; 22: 336-345.
10. Rust S, Rosier M, Funke H, Real J, Amoura Z *et al.* Tangier disease is caused by mutations in the gene encoding ATP-binding cassette transporter 1. *Nat Genet* 1999; 22: 352-355.
11. Tall AR. Metabolic and genetic control of HDL cholesterol levels. *J Intern Med* 1992; 231: 661-668.
12. Brown ML, Inazu A, Hesler CB, Agellon LB, Mann C *et al.* Molecular basis of lipid transfer protein deficiency in a family with increased high-density lipoproteins. *Nature* 1989; 342: 448-451.
13. Saha N, Tay JS, Humphries SE. Apolipoprotein B-gene DNA polymorphisms (XbaI and EcoRI), serum lipids, and apolipoproteins in healthy Chinese. *Genet Epidemiol* 1992; 9: 1-10.
14. Williams RR, Hopkins PN, Hunt SC, Wu LL, Hasstedt SJ *et al.* Population-based frequency of dyslipidemia syndromes in coronary-prone families in Utah. *Arch Intern Med* 1990; 150: 582-588.

Appendix 3

List of loci in multiplex reactions.

Reaction number	Chromosome	Marker name	Locus name
1	8	GAAT1A4	
	11	GATA28D01	D11S2000
	1	GATA72H07	D1S2134
2	8	GATA26E03	D8S1132
	11	GATA90D07	D11S2371
	1	GATA87F04	D1S2141
3	8	143XD8	D8S264
	19	GATA29B01	D19S589
	11	GATA6B09	D11S1392
	1	GATA4A09	D1S547
4	1	ATA29D04	D1S1631
	8	GATA14E09	D8S2324
5	1	GGAA3A07	D1S1612
	8	UT721	D8S373
	11	GATA26H10	D2S2739
6	8	COS140D4	D8S136
	11	GATA23F06	D19S1999
	18	ACT1A01	D18S843
	6	ATA50C05	D6S2434
7	8	GATA25C10	D8S1130
	1	GATA50F11	D1S1609
	17	AFM044XG3	D17S784
8	8	GGAA20C10	D8S1477
	11	ATA27C09	D11S2359
	1	ATA29C07	D1S3462
	2	GATA71D01	D2S1776
9	6	GATA31	D6S474
	1	GGAA22G10	D1S1677
	19	GATA66B04	D19S714
10	17	GATA185H04	D17S2196
	7	GATA189C06	D7S3070
	1	GATA48B01	D1S1660
11	10	ATA31G11	D10S1412
	1	GATA12A07	D1S534
	4	GATA28F03	D4S3248
	3	GGAT2G03	D3S2406
12	13	GATA51B02	D13S796
	1	ATA4E02	D1S1589
	5	AFM164X68	D5S408

Reaction number	Chromosome	Marker name	Locus name
13	2	GGAA20G04	D2S1399
	8	AFM073Y67	D8S256
	1	ATA20F08	D1S1622
	6	GATA64D02	D6S1053
14	6	GGAT3H10	D6S1017
	7	MFD265	D7S559
	1	GATA165C03	D1S3728
15	X	GATA31F01	DXS6789
	1	GATA29A05	D1S3669
	4	ATA26B08	D4S2394
16	1	AFM280WE5	D1S468
	5	GATA145D09	D5S2848
17	19	GATA44F10	D19S591
	1	GATA6A05	D1S551
	X	GATA69C12	DXS6810
18	2	GATA65C03	D2S1391
	16	AFM031XA5	D16S402
	8	GGAA8G07	D8S1113
19	8	GATA7G07	D8S1179
	9	AFM308VB1	D9S282
20	11	AFM157XH6	D11S912
	16	GATA22F09	D16S3253
	6	GATA68H04	D6S1056
	7	GATA118G10	D7S3046
21	11	GATA23E06	D11S1998
	X	AFM150F10	DXS1047
22	11	GGAA17G05	D11S1984
	19	GATA21G05	D19S1034
23	11	GATA48E02	D11S1981
24	16	AFM350VD1	D16S516
	7	GATA3F01	D7S820
	6	GATA81B01	D6S1277
25	6	ATA22G07	D6S1027
	7	GATA24F03	D7S3056
	16	GATA71F09	D16S2621
26	6	GATA165G02	D6S2436
	7	GATA41G07	D7S1802
	16	GATA71H05	D16S769
27	16	ATA41E04	D16S2616
	7	GATA31A10	D7S2846
28	16	GATA42E11	D16S764
	7	GATA32C12	D7S1824
	6	GATA32B03	D6S1009
29	3	GATA22G12	D3S2387
	7	GATA5D08	D7S821
30	7	AFM217YC5	D7S513
	6	GATA163B10	D6S2439
31	7	GGAA6D03	D7S3061
	6	F13A1	SE30
	16	GATA67G11	D16S2620
32	15	ACTC	ACTC
	17	ATA43A10	D17S2193
	2	GATA116B01	D2S2952
33	2	AFM112YD4	D2S125
	17	AFM217YD10	D17S928
	15	GATA27A03	D15S642

Reaction number	Chromosome	Marker name	Locus name
34	2	GATA30E06	D2S2944
	15	GATA63A03	D15S659
	17	GATA64B04	D17S1303
35	17	ATC6A06	D17S2180
	2	GATA86E02	D2S1788
	15	GATA88H02	D15S822
36	2	GATA11H10	D2S1360
	15	AFM323YD9	D15S211
	17	GTAT1A05	D17S1308
37	17	GATA28D11	D17S1301
	2	GATA165C07	D2S2976
	15	ATA24A08	D15S652
38	17	GATA49C09	D17S1290
	2	GATA4D07	D2S1334
	15	GATA22F01	D15S657
39	2	GATA4E11	D2S410
	15	GATA151F03	D15S1507
	17	GAAT2C03	D17S1298
40	10	GGAA5D10	D10S1213
	2	GATA69E12	D2S1394
	3	GATA22F11	D3S2427
	9	GATA62F03	D9S2169
41	2	GATA178G09	D2S2968
	5	ATA20G07	D5S2488
42	19	Mfd238	D19S254
	2	GATA23D03	D2S1363
43	19	Mfd232	D19S246
	2	GATA88G05	D2S1790
44	5	GATA21D04	D5S1457
	9	GATA27A11	D9S925
	14	ATA19H08	D14S592
45	5	GATA52A12	D5S1501
	14	GATA136B01	D14S1426
	9	GATA87E02	D9S1121
46	9	ATA18A07	D9S910
	14	GATA4B04	D14S306
47	5	GATA67D03	D5S2500
	9	AFM73YB11	D9S158
	14	GATA74E02	D14S742
48	14	MFD190	D14S53
	5	MFD154	D5S211
	9	GATA21F05	D9S922
49	14	GGAA21G11	D14S617
	9	GATA7D12	D9S301
	5	GATA3E10	D5S817
50	14	GATA43H01	D14S608
	9	ATA59H06	D9S2157
	5	GATA68A03	D5S2501
51	3	GATA27C08	D3S2432
	5	GATA6E05	D5S820
	9	GATA48D07	D9S930
52	20	GATA81E09	D20S604
	5	GATA3A04	D5S807
	12	PAH	PAH

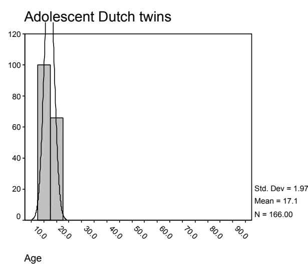
Reaction number	Chromosome	Marker name	Locus name
53	12	GATA26D02	D12S1052
	5	GATA89G08	D5S1725
	13	GATA43H03	D13S793
54	20	GATA51D03	D20S482
	5	GATA11A11	D5S1456
	18	GATA6D09	D18S851
55	3	AFM254VE1	D3S1311
	4	GATA107	D4S1625
	10	GATA115E01	D10S2470
56	4	GATA22G05	D4S2366
	3	GATA68D03	D3S2459
	10	GGAA23C05	D10S1248
57	4	GATA72G09	D4S2632
	3	ATA34G06	D3S4523
	10	GATA48G07	D10S1237
58	4	GATA90B10	D4S2639
	10	GATA70E11	D10S1423
	3	GATA3H01	D3S1763
59	4	GATA10G07	D4S3243
	3	GATA6F06	D3S1766
	10	GATA87G01	D10S1435
60	4	GATA2F11	D4S1647
	10	ATA24F10	D10S1225
	3	GATA4A10	D3S1764
61	10	ATA5A04	D10S1208
	4	AFM165XC11	D4S408
	3	GATA6G12	D3S2398
62	4	GATA5B02	D4S1652
	X	GATA175D03	DXS9902
	12	GATA11H08	D12S391
63	4	GATA8A05	D4S1629
	10	AFM189ZB4	D10S212
	3	AFM234TF4	D3S1304
64	X	GATA124E07	DXS9896
	4	GGAA19H07	D4S2431
65	12	ATA25F09	D12S2070
	18	ATA7D07	D18S1357
	13	ATA26D07	D13S779
66	12	ATA27A06	D12S1042
	18	GATA88A12	D18S976
	13	GGAA29H03	D13S1493
67	12	GGAT2G06	D12S398
	19	Mfd235	D19S245
	13	GATA29A09	D13S788
68	18	ATA1H06	D18S844
69	12	GATA63D12	D12S1064
70	18	GATA64H04	D18S877
	12	GATA4H03	D12S372
	13	GATA23C03	D13S787
71	13	AFM309VA9	D13S285
	12	GATA13D05	D12S392
	18	ATA82B02	
72	X	GATA165B12	
	13	GATA7G10	D13S317
	12	GATA32F05	D12S2078
	18	GATA178F11	

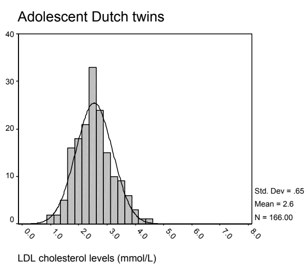
Reaction number	Chromosome	Marker name	Locus name
73	20	AFM046XF6	D20S171
	22	GCT10C10	D22S1685
	Y	GATA30F10	DYS389
	X	GATA72E05	DXS7132
74	22	GGAT3C10	D22S445
	X	GGAT3F08	DXS9900
	21	GATA70B08	D21S1446
75	21	GGAA3C07	D21S1437
	22	GATA6F05	D22S685
	X	GATA182E04	DXS9908
76	21	GATA129D11	D21S2052
	18	GATA11A06	D18S542
	20	GATA42A03	D20S478
77	21	GATA188F04	D21S2055
	XY	SDF1	DXYS154
	20	GATA45B10	D20S480
78	22	AFM217XF4	D22S420
	3	AFM036YB8	D3S1259
	5	GATA2H09	D5S816
79	20	AFM077XD3	D20S103
	21	GATA11C12	D21S1432
	Y	GGAAT1B07	
80	X	GATA172D05	
	10	GATA87G01	D10S1432

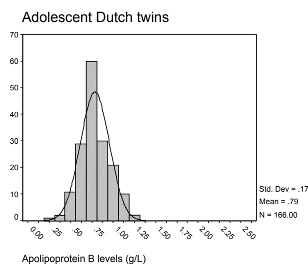
Appendix 4

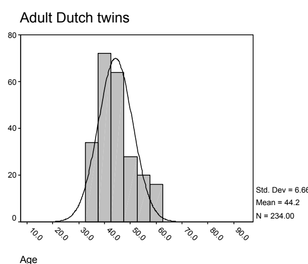
Histograms of the age, LDL cholesterol and apolipoprotein B levels in the adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively.

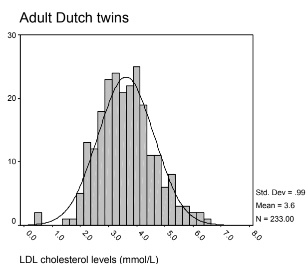
Age: **LDL cholesterol:** **Apolipoprotein B:**

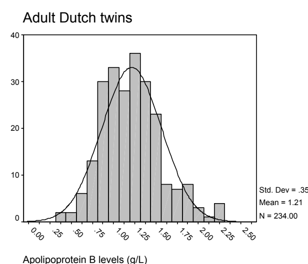


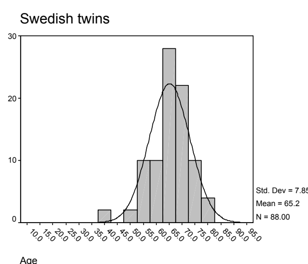


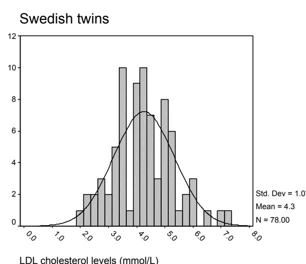


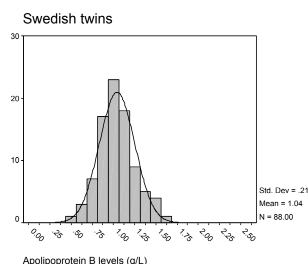


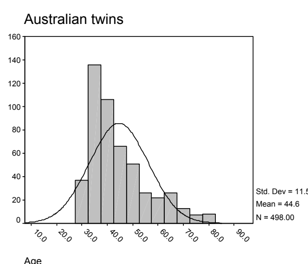


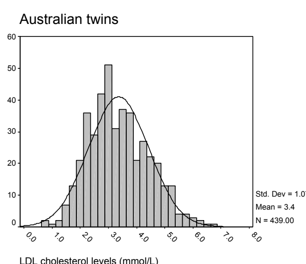


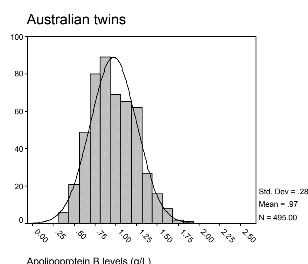






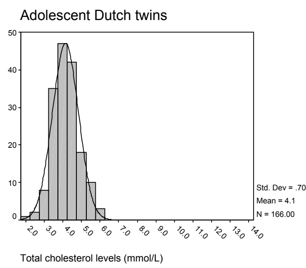


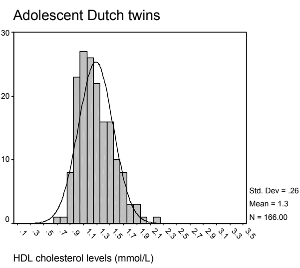


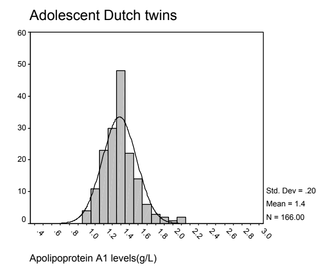


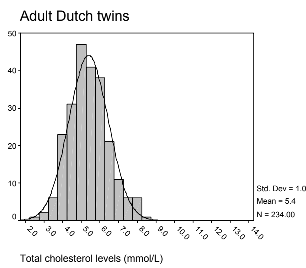
Histograms of the total cholesterol, HDL cholesterol and apolipoprotein A1 levels in the adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively.

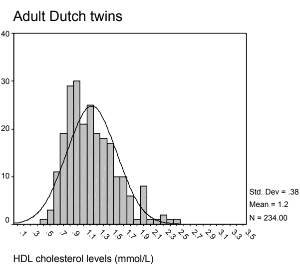
Total cholesterol: **HDL cholesterol:** **Apolipoprotein A1:**

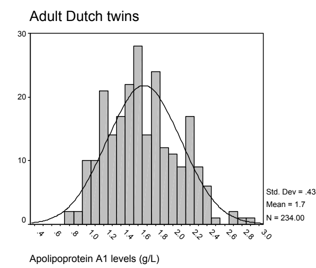


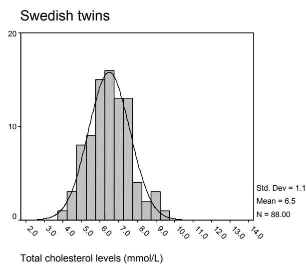


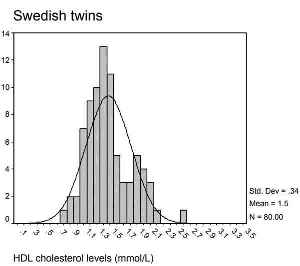


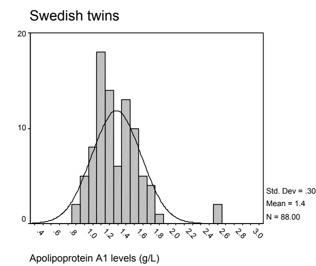


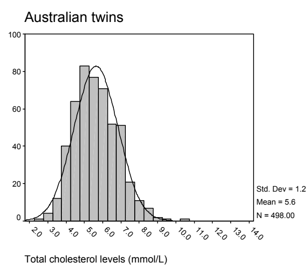


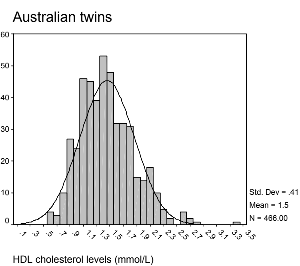


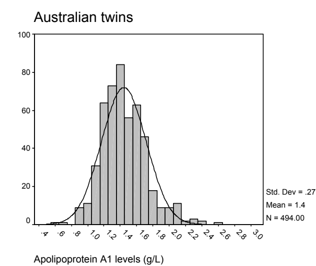






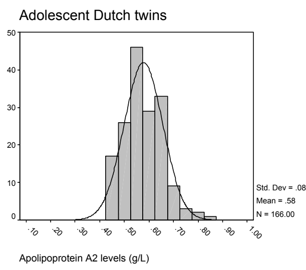


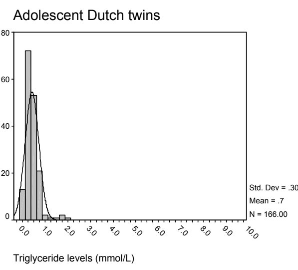


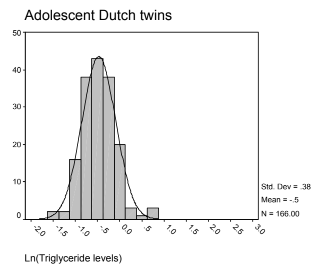


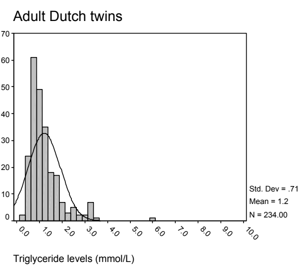
Histograms of the apolipoprotein A2, triglyceride and Ln(triglyceride) levels in the adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively.

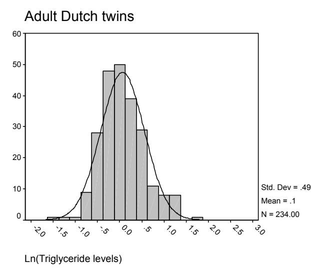
Apolipoprotein A2: | **Triglyceride:** | **Ln(triglyceride):**

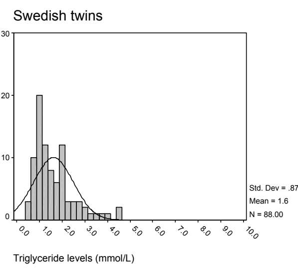


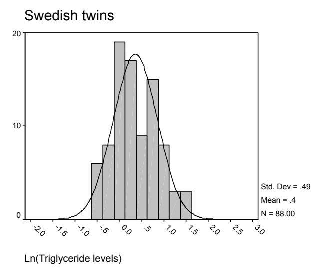


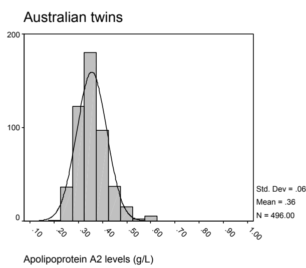


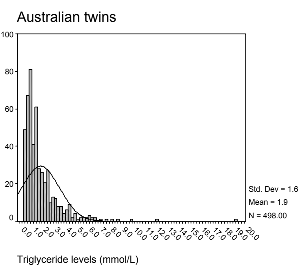


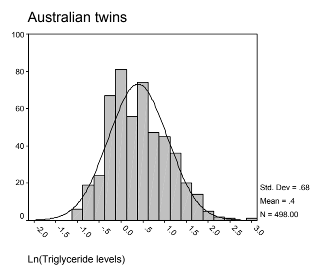






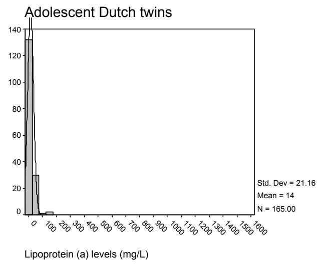


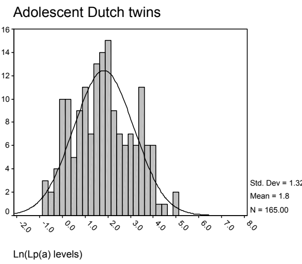


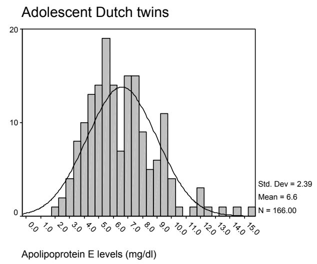


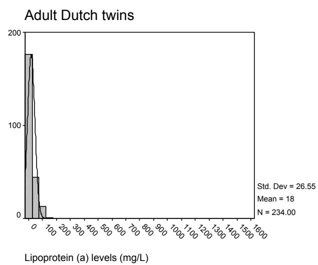
Histograms of the lipoprotein (a), Ln(lipoprotein(a)) and apolipoprotein E levels in the adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively.

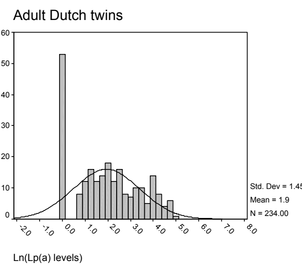
Lipoprotein (a): **Ln(lipoprotein(a)):** **Apolipoprotein E:**

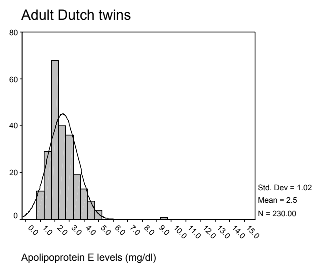


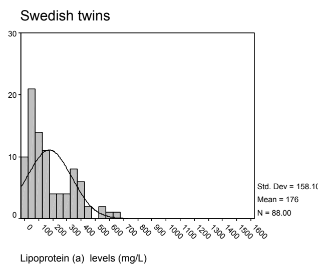


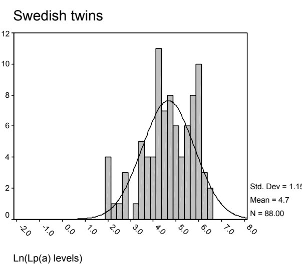


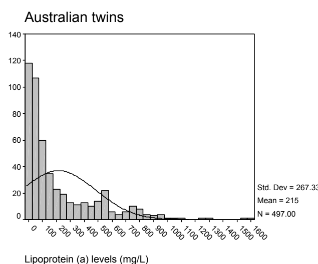


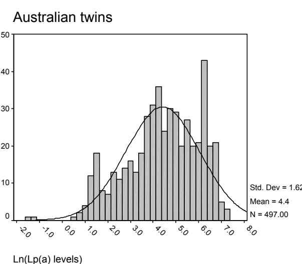


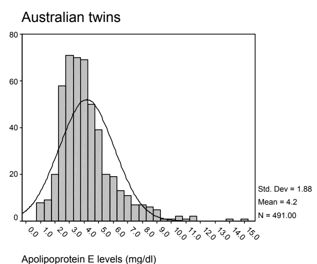






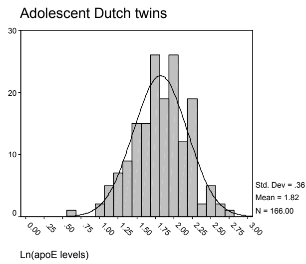


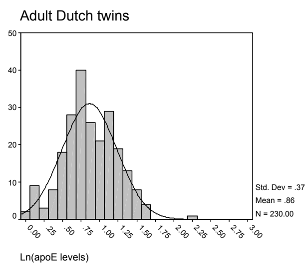


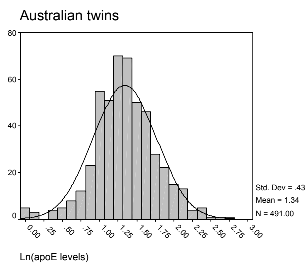


Histograms of the lipoprotein (a), Ln(lipoprotein(a)) and apolipoprotein E levels in the adolescent Dutch, adult Dutch, Swedish and Australian twins, respectively.

Ln(Apolipoprotein E):







Appendix 5

Maximum LOD scores per chromosome per twin sample per phenotype.

Gray background indicates the confirmed influence of the *LPA* gene. **Bold font** indicates MLS higher that 3.6, ***bold as well as italic font*** indicates MLS between 2.2 and 3.6, underlined font indicates linkage results with MLS higher than 2.2 in one twin sample and higher than 1.4 in an other twin samples within 20 cM.

Chromosome 1	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.49	136.88	0.99	238.92	0.35	16.22	1.63	89.49
LDL cholesterol	0.28	0.00	0.62	236.15	0.49	113.69	***2.58***	***267.51***
ApoB	0.83	75.66	0.76	37.05	0.41	9.02	0.73	99.17
HDL cholesterol	1.36	37.05	0.41	267.51	0.66	233.38	1.40	52.80
ApoA1	1.02	200.21	0.62	94.33	0.39	233.38	0.25	40.99
ApoA2	0.88	132.24	-	-	-	-	0.75	236.15
Ln(TG)	1.54	220.82	1.33	113.69	0.17	113.69	0.23	192.05
Ln(Lp(a)	0.92	4.22	0.71	247.23	2.62	236.15	1.83	251.29
Ln(apoE)	0.28	233.38	1.64	259.398	-	-	0.64	99.17

Chromosome 2	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	1.23	99.4	0.58	168.8	0.69	55.5	0.28	55.5
LDL cholesterol	0.76	99.4	0.60	173.0	0.92	38.3	1.63	173.0
ApoB	0.62	99.4	1.08	186.2	0.51	55.5	1.27	55.5
HDL cholesterol	0.45	186.2	1.02	67.4	0.39	6.6	0.35	213.7
ApoA1	1.53	131.1	0.88	3.8	0.49	70.3	0.76	41.8
ApoA2	0.87	17.9	-	-	-	-	1.00	9.4
Ln(TG)	1.29	70.3	0.95	145.1	0.98	210.4	0.26	251.9
Ln(Lp(a)	0.36	17.9	0.00	210.4	1.46	17.9	5.57	38.3
Ln(apoE)	0.25	126.7	0.69	210.4	-	-	0.17	17.9

Gray background indicates the confirmed influence of the *LPA* gene. **Bold font** indicates MLS higher that 3.6, ***bold as well as italic font*** indicates MLS between 2.2 and 3.6, underlined font indicates linkage results with MLS higher than 2.2 in one twin sample and higher than 1.4 in an other twin samples within 20 cM.

Chromosome 6	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	1.89	173.31	0.47	151.26	0.11	122.46	0.99	63.28
LDL cholesterol	1.13	173.31	0.36	147.88	0.18	9.18	0.67	50.67
ApoB	1.08	42.27	0.07	102.81	0.29	126.28	1.25	46.47
HDL cholesterol	0.09	173.31	0.21	9.18	0.78	154.64	0.87	165.84
ApoA1	0.41	18.72	0.24	9.18	0.54	63.28	0.73	9.18
ApoA2	0.76	21.9	-	-	-	-	0.65	118.40
Ln(TG)	0.82	42.27	1.36	154.64	0.26	25.08	0.54	181.66
Ln(Lp(a)	2.75	164.31	4.40	154.64	2.28	176.09	5.13	181.66
Ln(apoE)	0.10	25.08	0.21	137.74	-	-	0.95	173.31

Chromosome 7	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	1.34	97.41	0.56	149.90	0.37	124.55	0.99	38.03
LDL cholesterol	0.42	109.12	0.25	141.30	0.11	116.84	0.85	30.02
ApoB	0.28	181.97	0.34	128.41	0.24	17.74	***2.81***	***90.52***
HDL cholesterol	0	-	0	-	0.61	7.44	0.77	181.97
ApoA1	0.00	57.79	0.14	181.97	***2.24***	***17.74***	1.63	181.97
ApoA2	0.57	42.97	-	-	-	-	0.06	149.90
Ln(TG)	0.08	112.98	0.80	90.52	0.25	33.09	0.29	47.91
Ln(Lp(a)	0.32	149.90	1.27	98.44	0.32	98.44	0.78	17.74
Ln(apoE)	0.63	181.97	0.27	181.97	-	-	1.04	128.41

Chromosome 8	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.40	148.1	0.16	164.5	0.44	110.2	1.98	142.9
LDL cholesterol	0.46	148.1	0.29	164.5	0.55	119.2	0.86	67.4
ApoB	0.34	154.7	0.29	164.5	0.63	119.2	0.70	26.7
HDL cholesterol	1.03	26.7	0.45	164.5	0.87	148.1	0.11	164.5
ApoA1	***2.68***	***22.4***	1.32	157.9	0.93	148.1	0.17	60.3
ApoA2	0.15	164.5	-	-	-	-	1.43	164.5
Ln(TG)	1.05	5.1	0.45	119.2	0.97	135.1	0.15	164.5
Ln(Lp(a)	0.50	77.9	0.92	47.2	2.12	60.3	0.31	161.2
Ln(apoE)	0.16	0.7	0.00	167.9	-	-	0	-

Gray background indicates the confirmed influence of the *LPA* gene. **Bold font** indicates MLS higher that 3.6, ***bold as well as italic font*** indicates MLS between 2.2 and 3.6, underlined font indicates linkage results with MLS higher than 2.2 in one twin sample and higher than 1.4 in an other twin samples within 20 cM.

Chromosome 11	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.18	131.3	0.82	17.2	1.03	100.6	0.38	56.3
LDL cholesterol	0.73	131.3	0.86	17.2	1.34	108.1	0.22	76.1
ApoB	0.20	131.3	0.54	81.0	0.76	134.6	***2.59***	***62.9***
HDL cholesterol	0	-	0.21	21.5	0	-	1.26	144.5
ApoA1	0.07	17.2	0.51	113.1	0	-	0.00	17.2
ApoA2	0.59	69.5	-	-	-	-	0.78	131.3
Ln(TG)	0.33	14.2	0.32	56.3	0.23	21.5	0.62	69.5
Ln(Lp(a)	0.03	147.8	**4.13**	**113.1**	0.74	147.8	0.68	147.8
Ln(apoE)	0.28	8.1	0.27	18.9	-	-	0.57	0.0

Chromosome 15	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.41	72.7	0	-	0.11	16.1	0.33	31.5
LDL cholesterol	0.56	69.6	0.03	31.5	0.38	12.3	0	-
ApoB	0.16	90.0	0.02	31.5	0.55	27.6	0.01	121.1
HDL cholesterol	0.01	60.2	0	-	0.65	122.1	0.68	84.4
ApoA1	***2.33***	***63.3***	0.27	122.1	1.86	122.1	0.38	90.0
ApoA2	2.10	43.5	-	-	-	-	1.31	75.9
Ln(TG)	1.35	27.6	0	-	0.32	122.1	0	-
Ln(Lp(a)	0.45	122.1	0.21	12.3	1.90	63.3	0.83	60.2
Ln(apoE)	0.58	56.8	0.00	12.3	-	-	0.03	104.9

Chromosome 16	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.86	92.7	0.40	130.4	0.34	113.5	1.57	130.4
LDL cholesterol	0.50	88.8	0.35	130.4	0.93	113.5	0.63	123.7
ApoB	0.00	71.8	0.02	130.4	0.15	130.4	0.84	0.0
HDL cholesterol	0.17	100.4	0.30	130.4	0.02	50.6	1.27	73.6
ApoA1	0.47	130.4	0.48	67.5	0.48	71.8	0.08	116.9
ApoA2	0.19	113.5	-	-	-	-	0	-
Ln(TG)	0.38	71.8	0.67	73.6	0.65	113.5	0.38	11.5
Ln(Lp(a)	0	-	0.75	130.4	0	-	0.68	0.0
Ln(apoE)	0.00	113.5	0.15	130.4	-	-	1.93	9.2

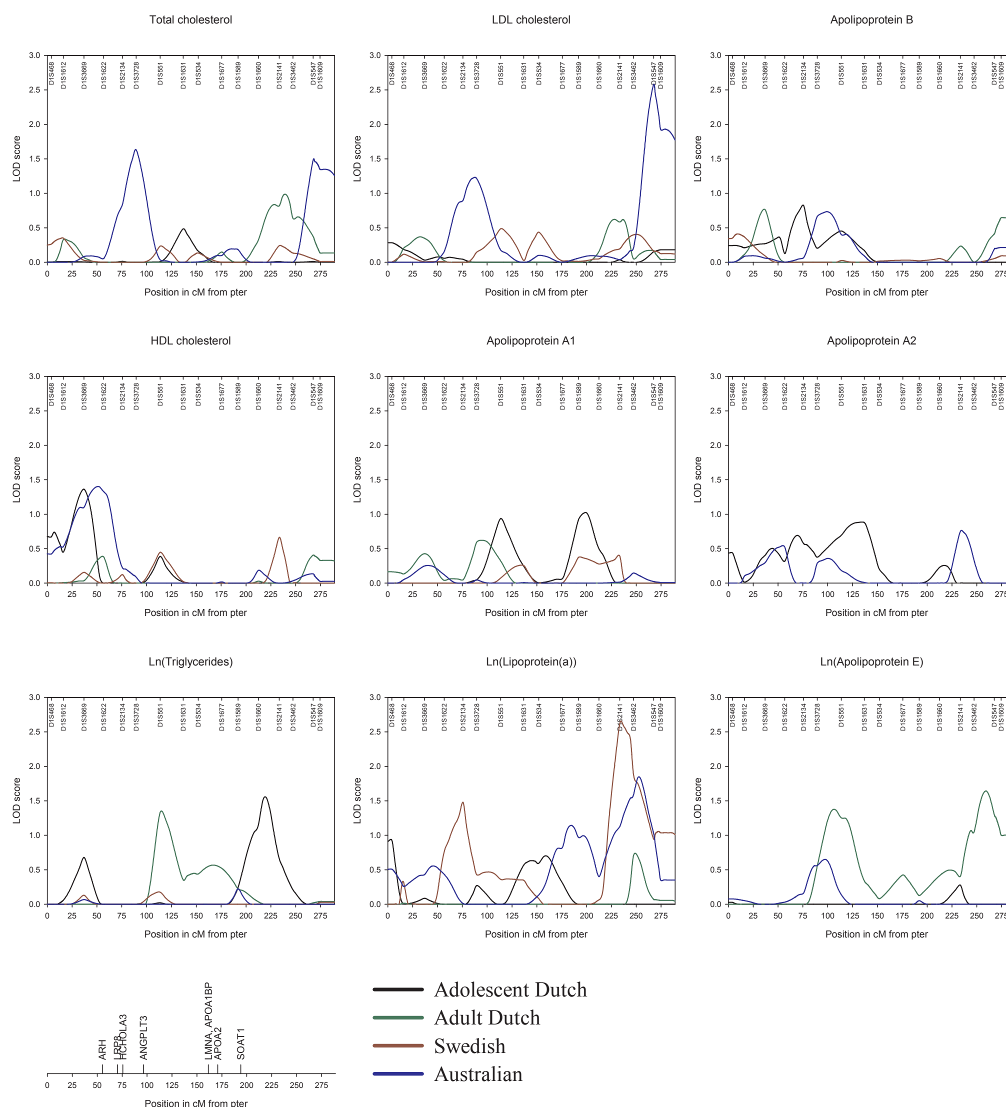
Gray background indicates the confirmed influence of the *LPA* gene. **Bold font** indicates MLS higher that 3.6, ***bold as well as italic font*** indicates MLS between 2.2 and 3.6, underlined font indicates linkage results with MLS higher than 2.2 in one twin sample and higher than 1.4 in an other twin samples within 20 cM.

Chromosome 17	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.21	126.5	0.71	82.0	0.25	116.9	1.11	100.0
LDL cholesterol	0.35	126.5	0.34	82.0	0.667	126.5	1.24	93.6
ApoB	0.03	66.9	0.51	86.4	0.72	126.5	0.55	89.3
HDL cholesterol	0.40	0.6	0.04	100.0	0.58	97.9	0.08	126.5
ApoA1	0.53	97.9	0.05	89.3	0.84	100.0	0.27	116.9
ApoA2	0.35	66.9	-	-	-	-	1.10	0.6
Ln(TG)	0.37	0.6	0.28	82.0	0.02	82.0	1.04	91.5
Ln(Lp(a)	0.00	126.5	0.34	86.4	0.29	100.0	0.01	10.7
Ln(apoE)	0.23	82.0	1.07	4.7	-	-	0.42	23.6

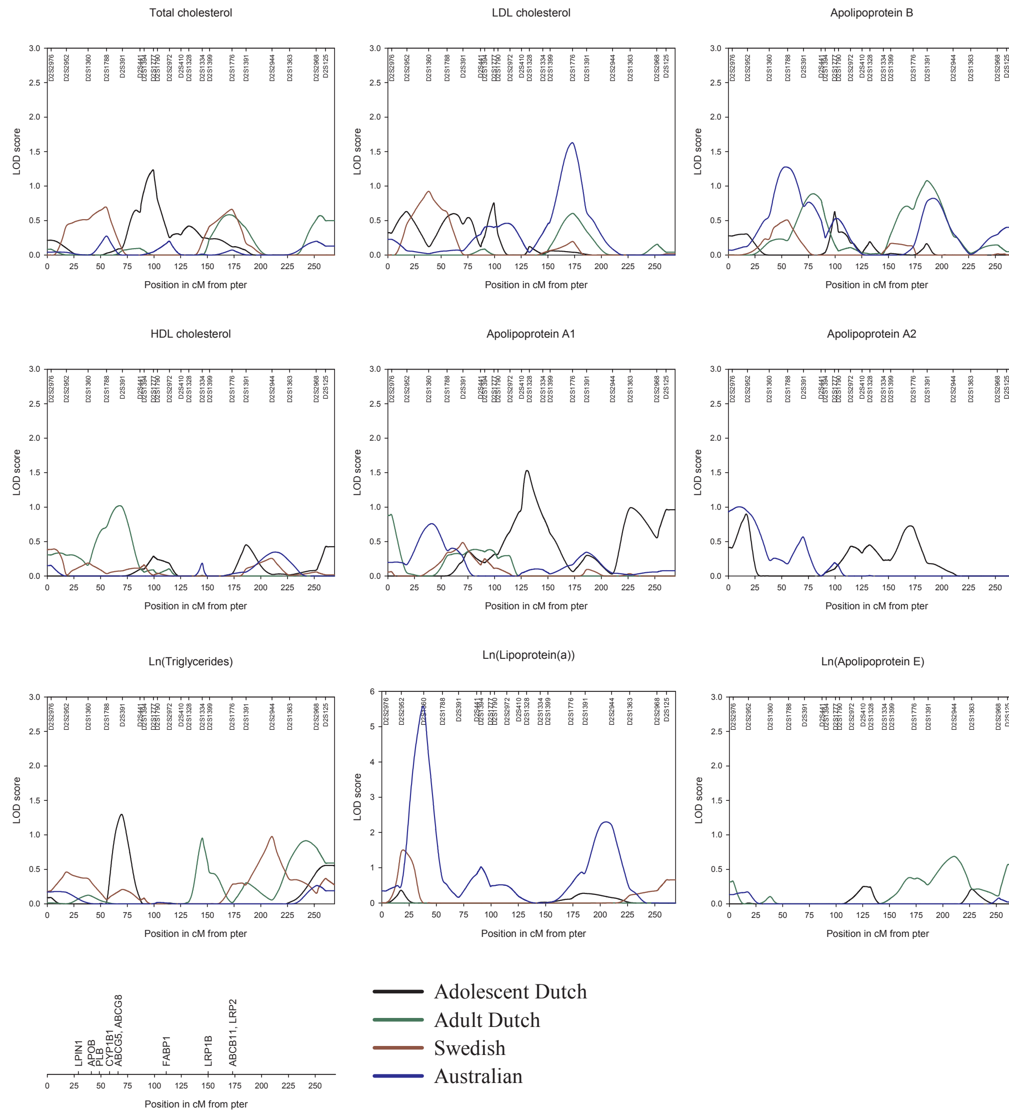
Chromosome 19	**Adolescent Dutch**		**Adult Dutch**		**Swedish**		**Australian**	
Trait	MLS	Position	MLS	Position	MLS	Position	MLS	Position
Total cholesterol	0.12	100.6	1.13	40.97	1.40	65.8	0.15	34.3
LDL cholesterol	0.48	92.8	2.19	54.5	1.28	65.8	1.73	34.3
ApoB	0.10	100.6	0.51	50.8	1.61	69.5	0.11	69.5
HDL cholesterol	0.05	100.6	0.07	29.4	0.20	34.3	1.43	50.8
ApoA1	0.04	87.7	0.33	100.6	0.78	34.3	1.26	50.8
ApoA2	0.11	62.0	-	-	-	-	0.79	50.8
Ln(TG)	0.67	50.8	0.33	50.8	0.10	40.9	0.06	87.7
Ln(Lp(a)	0.36	69.5	0.17	34.3	0.03	100.6	2.24	9.8
Ln(apoE)	1.00	65.8	0.55	65.8	-	-	0.09	100.6

Appendix 6

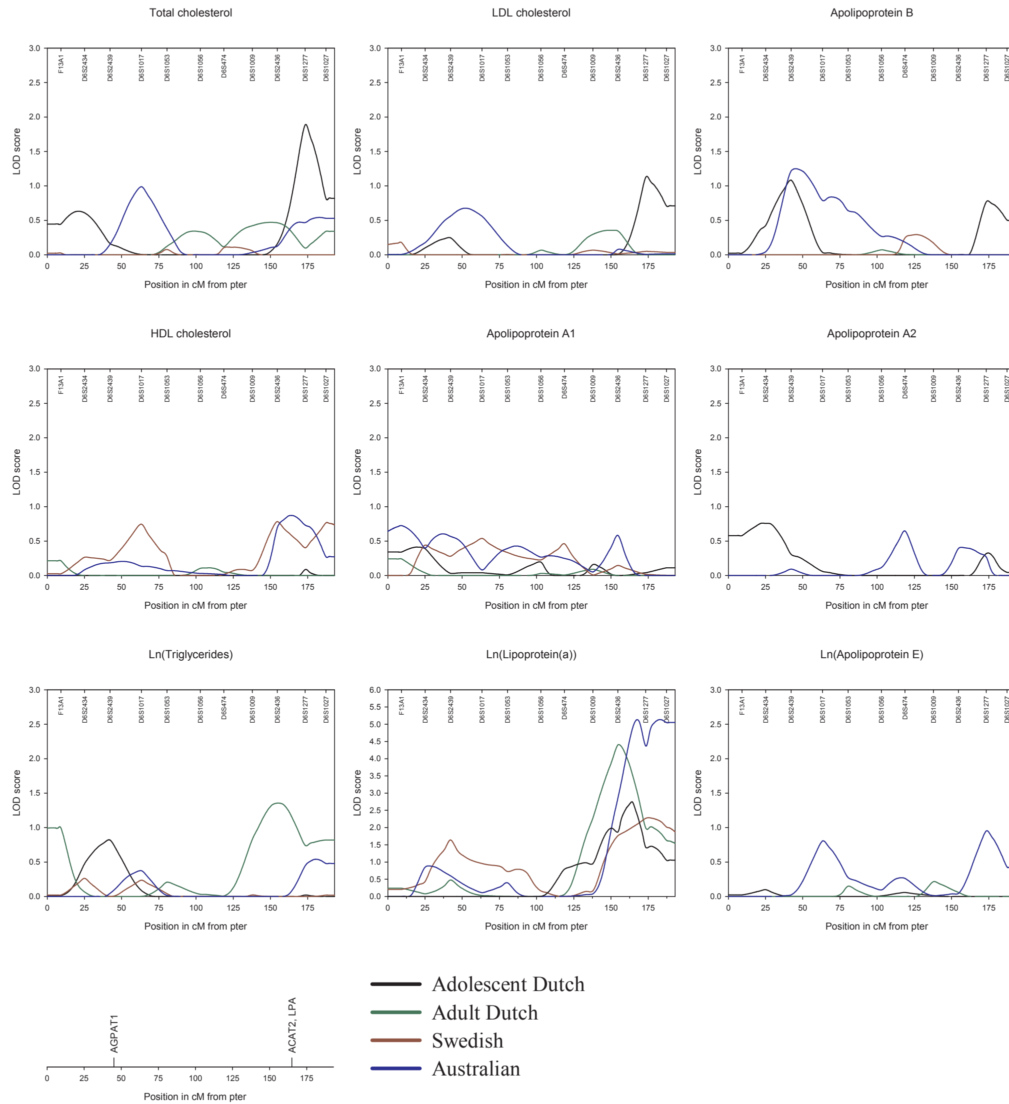
Linkage results of chromosome 1.



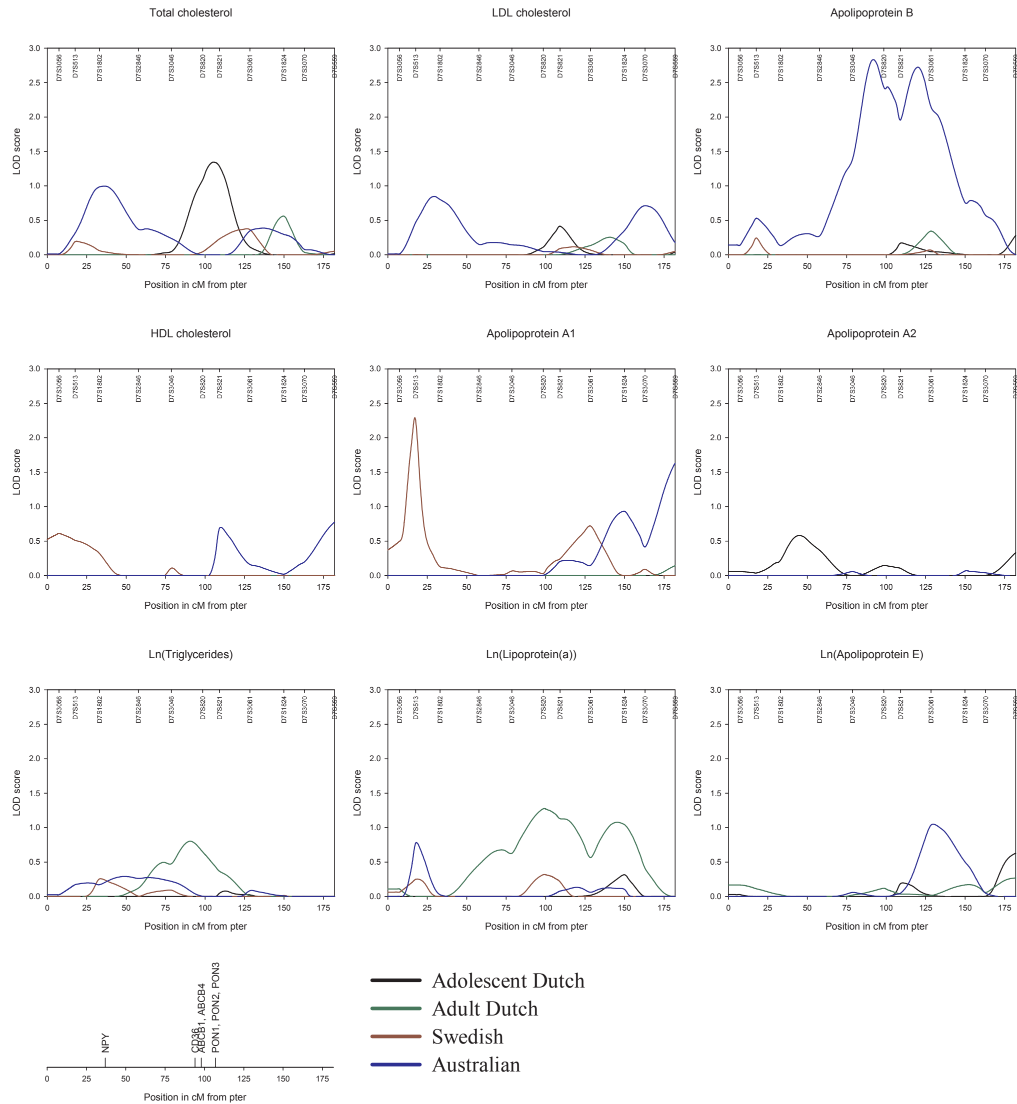
Linkage results of chromosome 2.



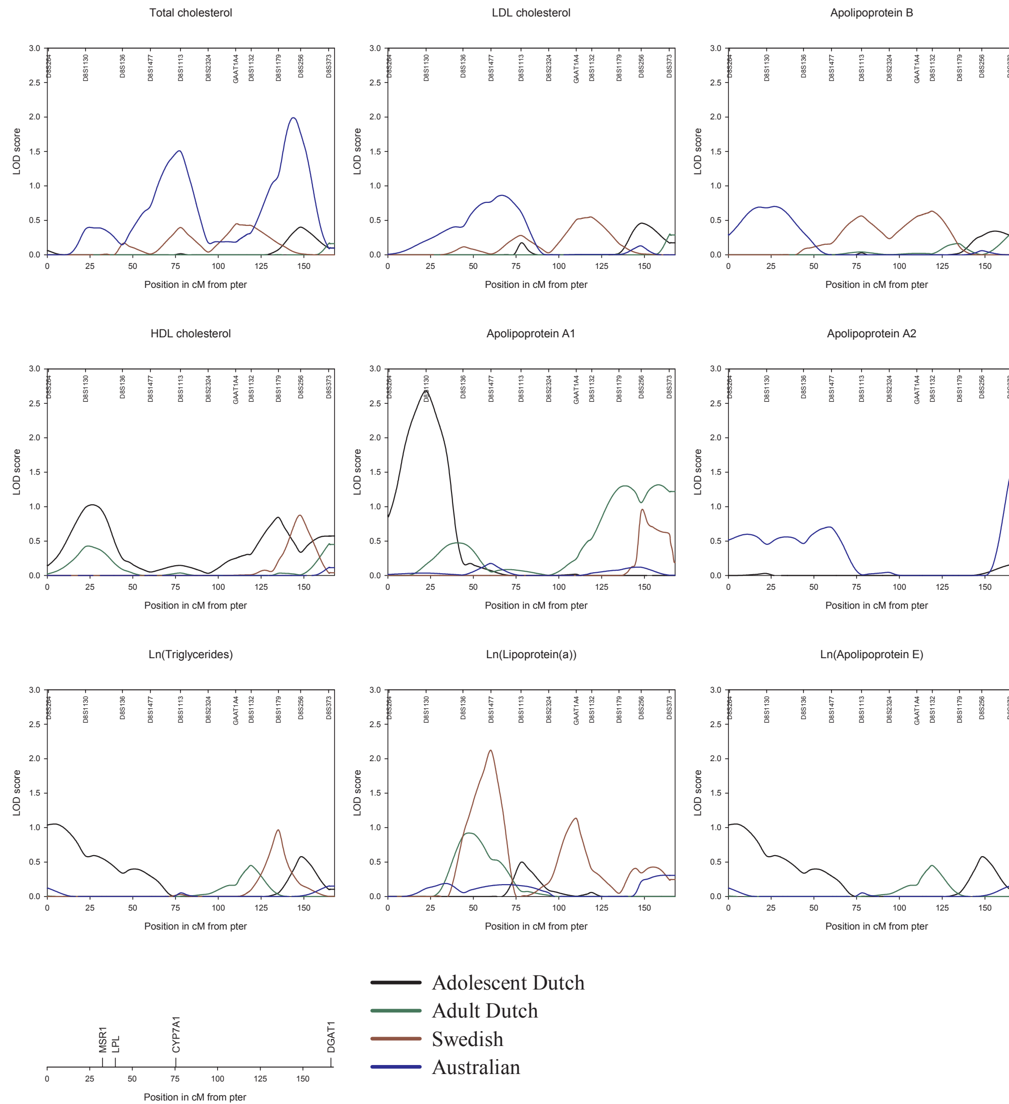
Linkage results of chromosome 6.



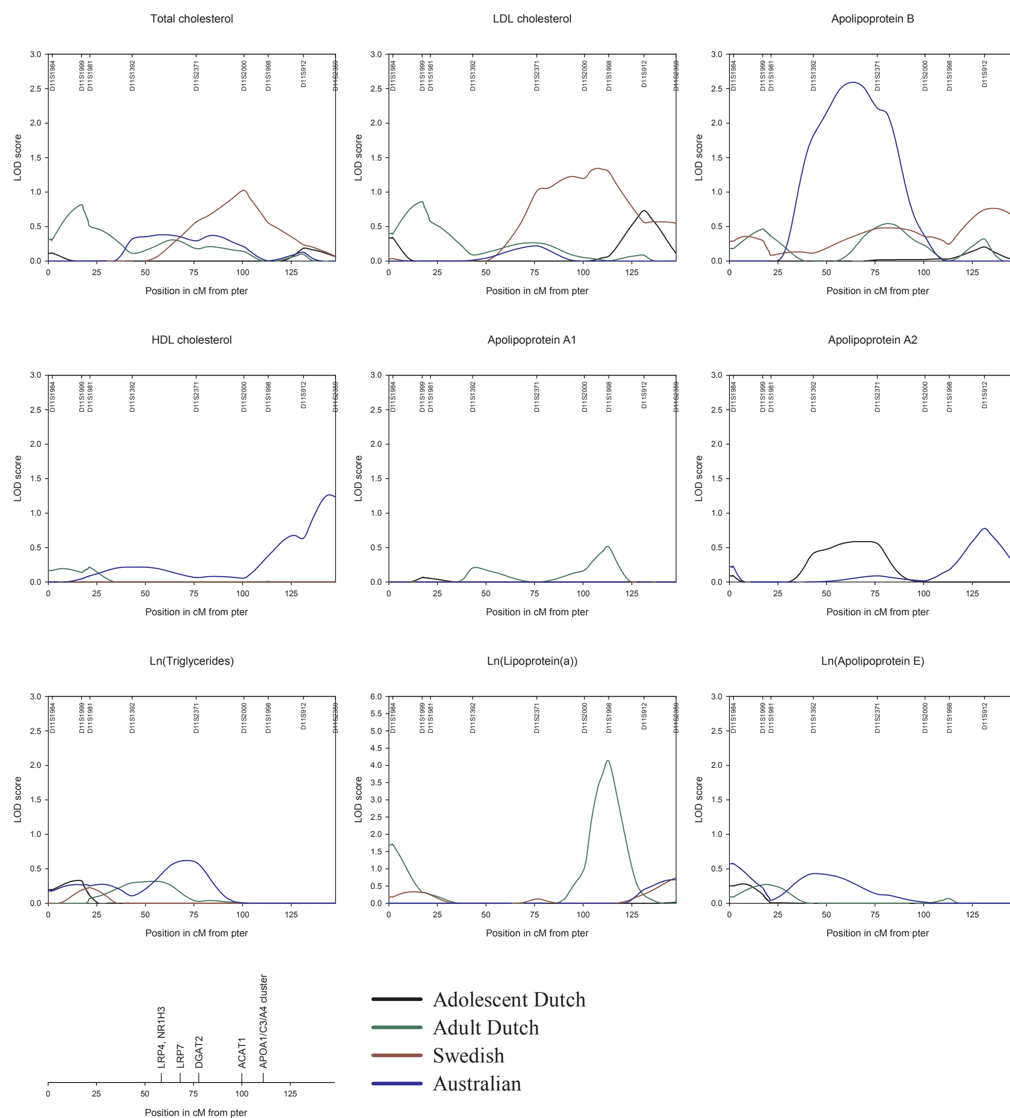
Linkage results of chromosome 7.



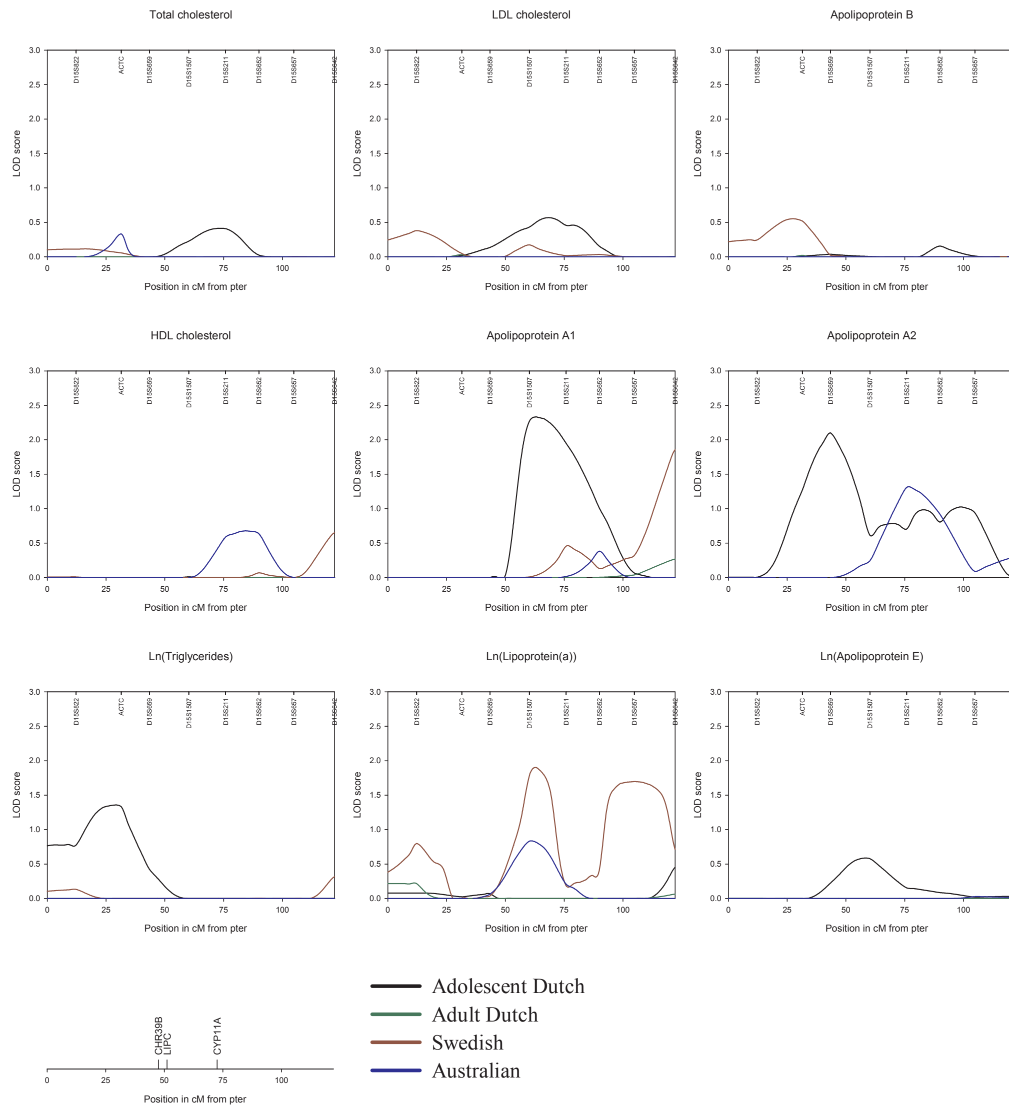
Linkage results of chromosome 8.



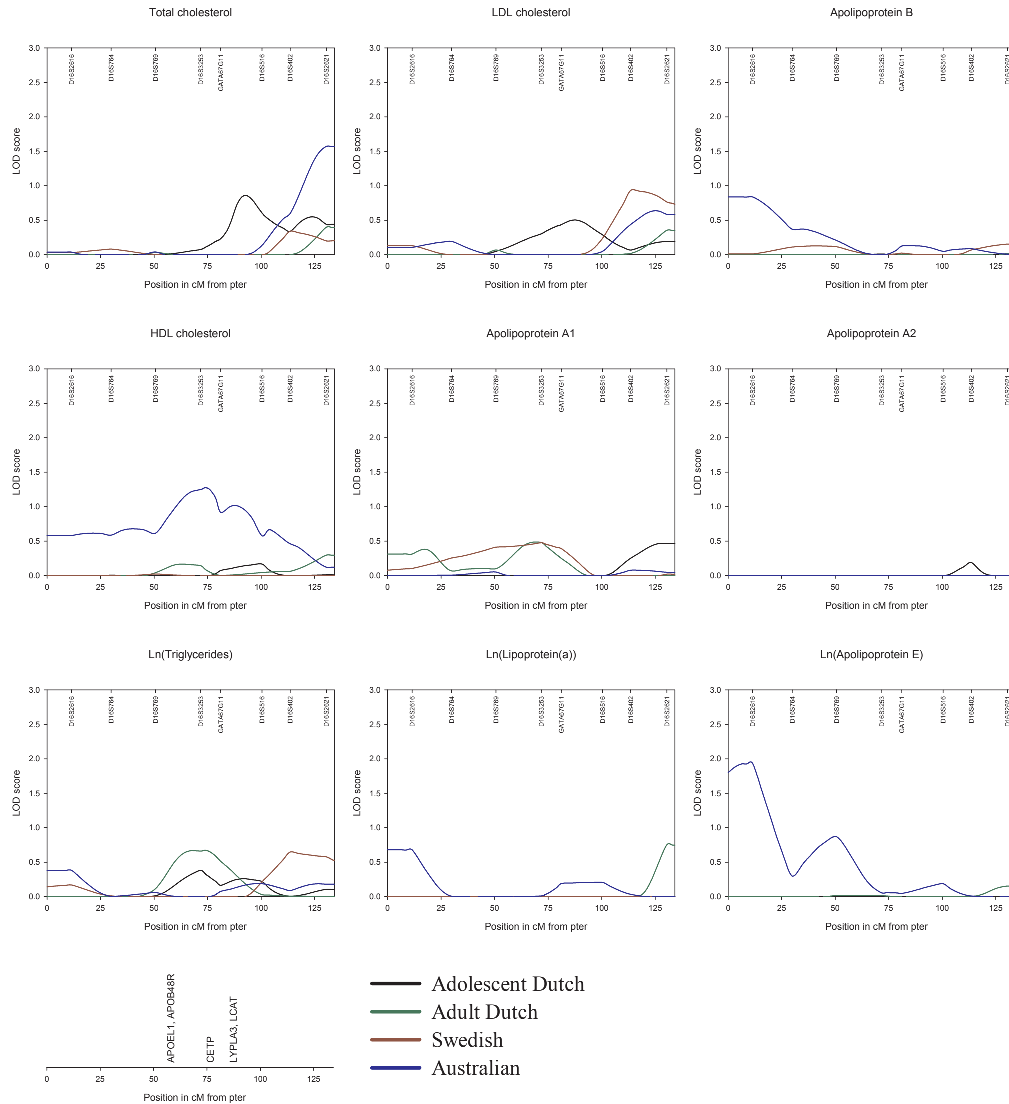
Linkage results of chromosome 11.



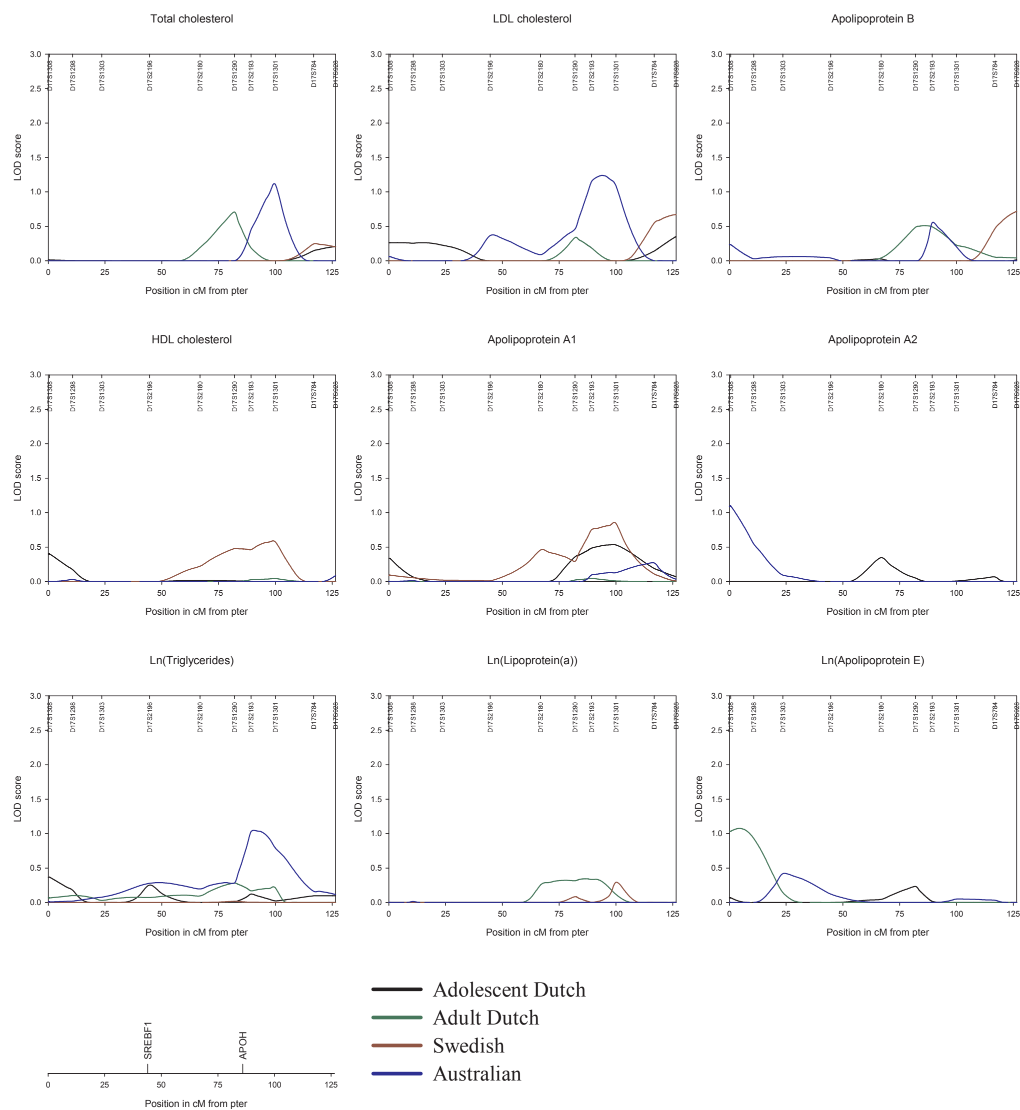
Linkage results of chromosome 15.



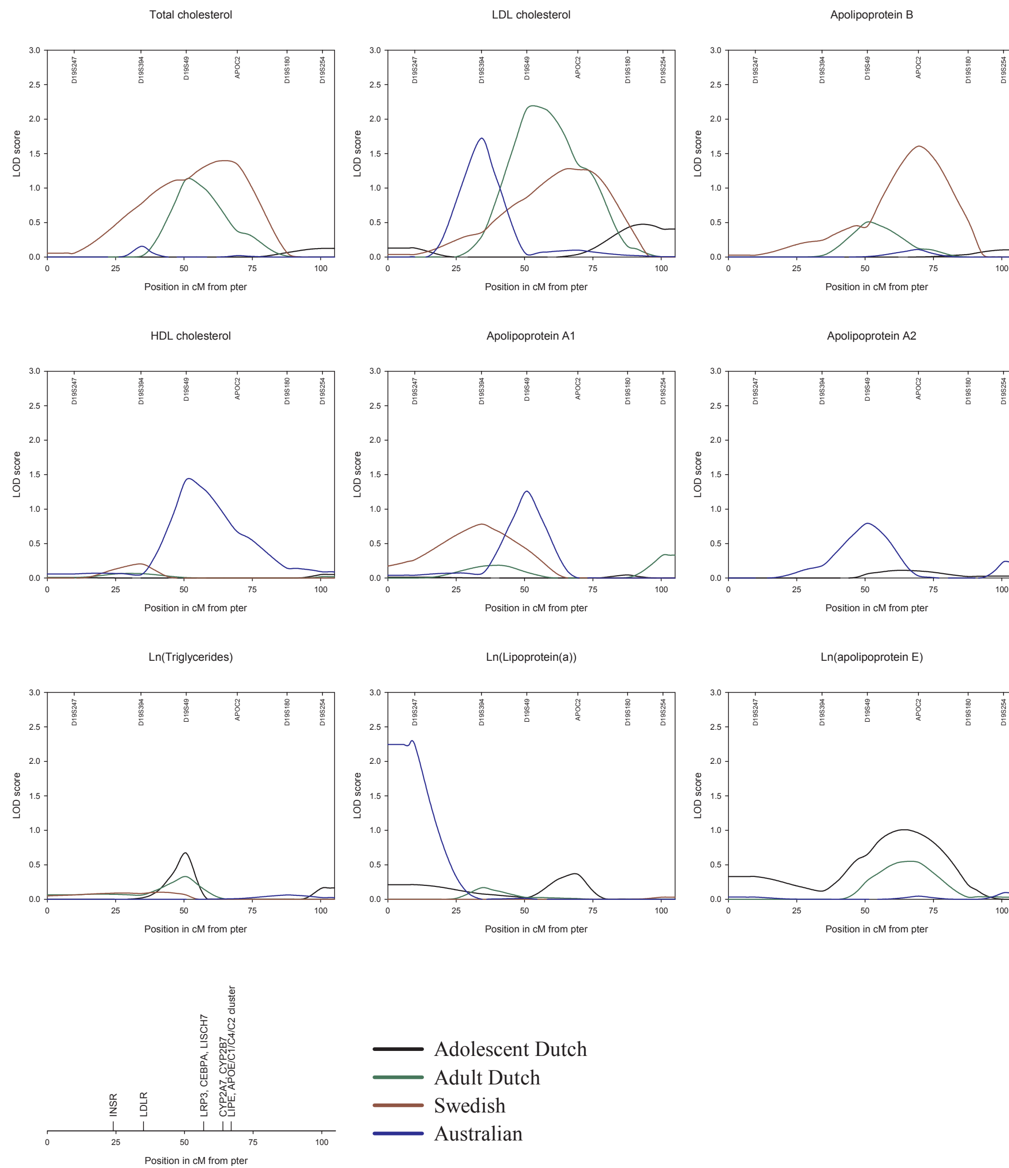
Linkage results of chromosome 16.



Linkage results of chromosome 17.



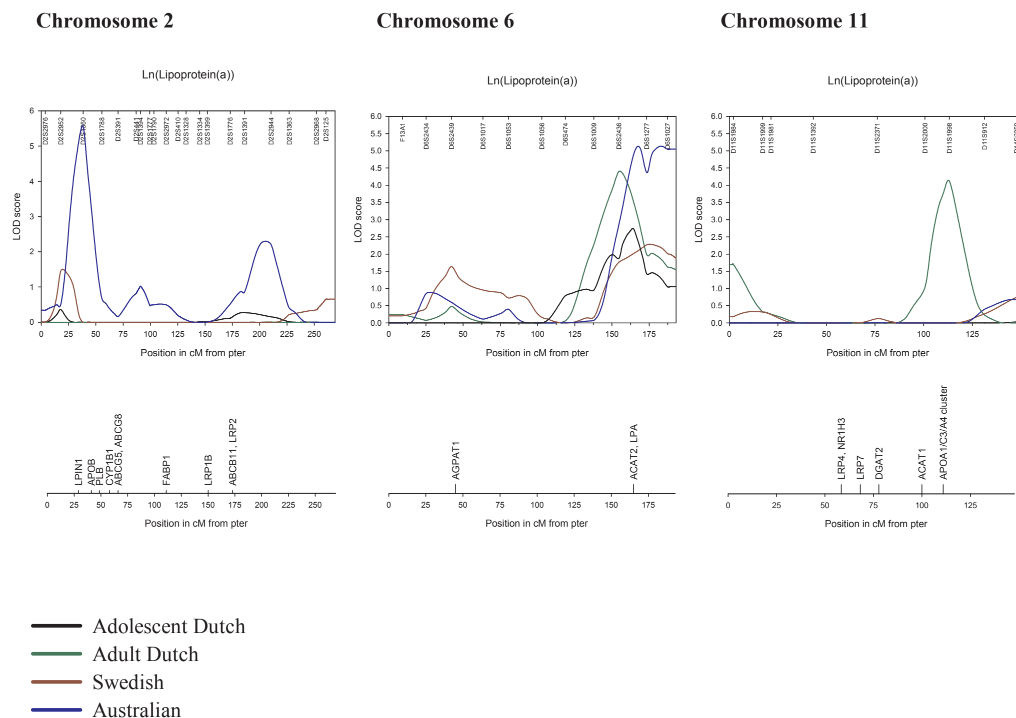
Linkage results chromosome 19.



Appendix 7

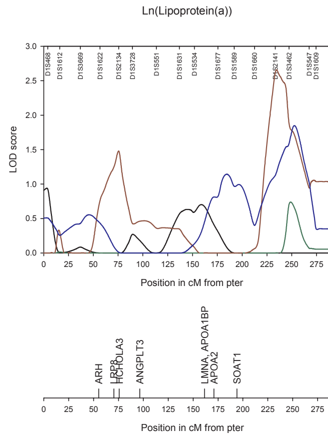
Linkage graphs of results shown in table 5, 6 and 7 of Chapter 4.

Significant linkage results (MLS ≥ 3.6) shown in table 5 of Chapter 4.

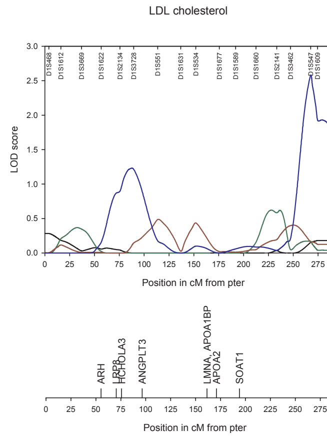


Suggestive linkage results (2.2 ≤ MLS ≤ 3.6) shown in table 6 of Chapter 4.

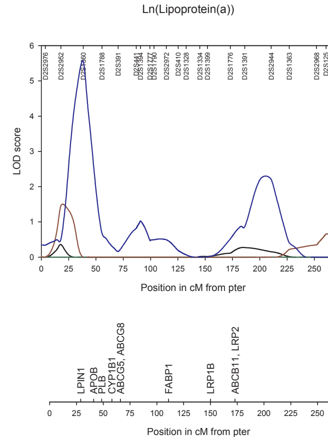
Chromosome 1



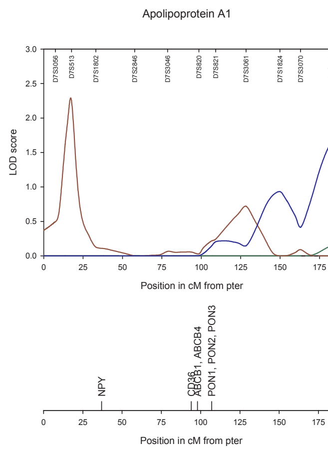
Chromosome 1



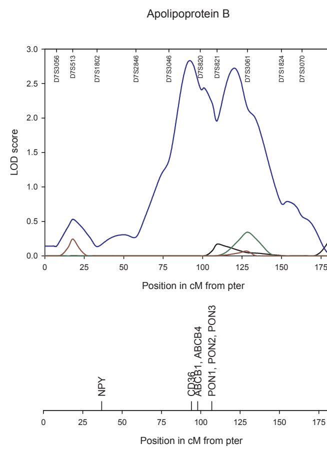
Chromosome 2



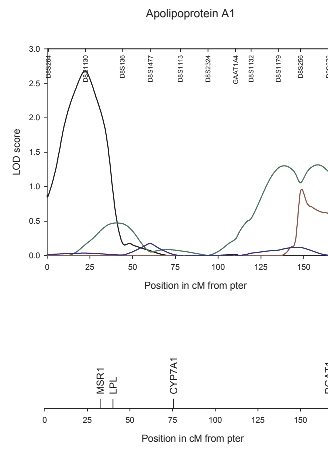
Chromosome 7



Chromosome 7

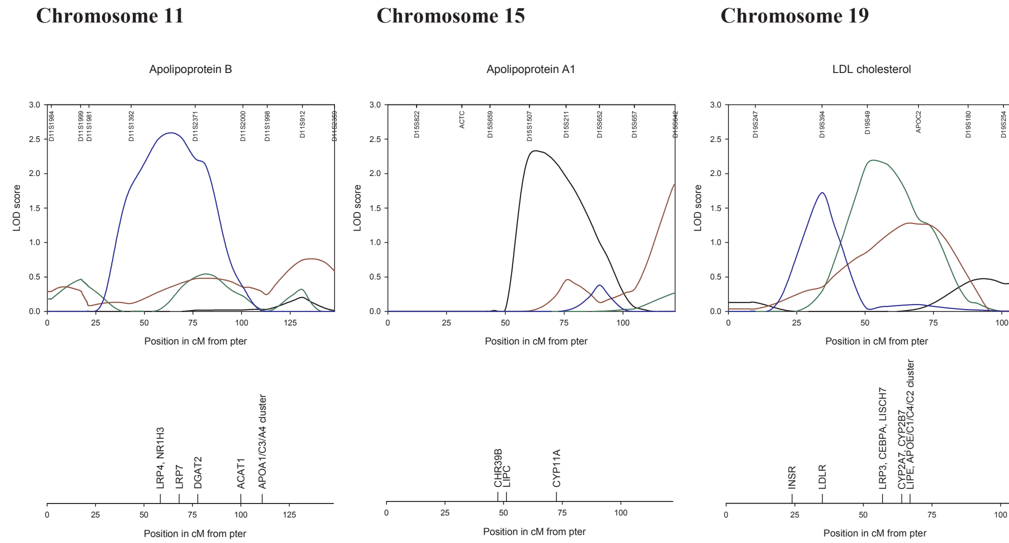


Chromosome 8

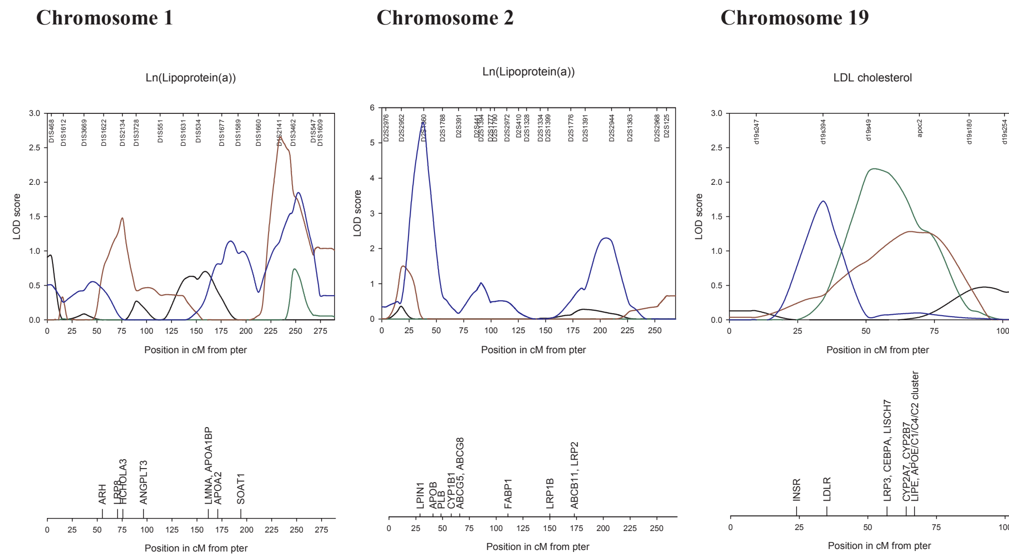


Adolescent Dutch
Adult Dutch
Swedish
Australian

Suggestive linkage results (2.2 ≤ MLS ≤ 3.6) shown in table 6 of Chapter 4-continued.



Replicated linkage results with MLS higher than 2.2 in one twin sample and MLS higher than 1.4 in a second twin sample within 20 cM show in table 7 of Chapter 4.



Adolescent Dutch
Adult Dutch
Swedish
Australian

Nawoord

De voltooiing van dit proefschrift is de verdienste geweest van een groot aantal mensen. Het onderzoek beschreven in dit proefschrift betrof een groot samenwerkingsverband tussen het *Center for Developmental and Health Genetics*, Pennsylvania (Verenigde Staten van America), *Karolinska Institutet*, Stockholm (Zweden), *Queensland Institute for Medical Research*, Brisbane (Australië), TNO Preventie en Gezondheid, Leiden (Nederland), Leids Universitair Medisch Centrum, Leiden (Nederland) en de Vrije Universiteit van Amsterdam (Nederland) en werd gefinancierd door *National Institutes of Health* (USA). De praktische uitvoering van het onderzoek werd gecoördineerd vanuit het Gaubius Laboratorium van TNO Preventie en Gezondheid (prof. dr. P. E. Slagboom) en de genetische analyses vanuit de Vrije Universiteit van Amsterdam (prof. dr. D. I. Boomsma).

Veel dank ben ik verschuldigd aan alle tweelingen die geheel belangeloos hebben meegewerkt aan dit onderzoek. Zonder hun medewerking had deze studie nooit plaats gevonden. Natuurlijk wil ik hier ook graag iedereen bedanken die heeft bijgedragen aan het verzamelen van deze tweelingparen en het bepalen van de fenotypen.

Bij TNO Preventie en Gezondheid zijn alle genotyperingen bepaald die voor het onderzoek nodig waren. Uit de goede samenwerking met Nico was een prachtig strak werkschema ontstaan, zodat 3000 genotyperingen per week konden worden voltooid. Van zijn doorzettingsvermogen en gevoel voor plannen heb ik veel geleerd. Bij het fijn karteren van chromosoom 19 heeft Eka een zeer belangrijke rol gespeeld en met name voor het bepalen van de *APOEε2/ε3/ε4* isoform in de gehele tweelingpopulatie ben ik haar zeer erkentelijk. Dennis typeerde een groot deel van de markers op de chromosomen 2, 15 en 17 in het kader van zijn stageopdracht. Samen met deze mensen heb ik de genotyperingen verzameld die het startpunt vormden voor de genetische analyses. Caroline heeft mij goed geholpen met de heritability analyses, Ingrid heeft veel bijgedragen aan het opstarten van Genehunter en de hulp van Daniëlle bij de analyses in Mx is van groot belang geweest. Bovendien zijn de inspiratie en stimulatie van Bas, Dorret en Eline onmisbaar geweest.

In december 2001 is onze groep verhuisd van TNO naar het LUMC. Uit de periode die we in het Gaubius Laboratorium werkten, wil ik graag Simone en Leonie noemen, waar ik jaren plezierig mee heb samengewerkt. Erik was als kamergenoot een goede afleiding. Onze huidige sectie Moleculaire Epidemiologie bestaat uit Bas, Bernd, Dennis, Eka, Eline, Esther, Greet, Inge,

Ingrid, Josine, Kate, Kim, Marja, Michiel, Nico, Rita, Ruud, en Suzan. Zij hebben mij de afgelopen jaren enorm veel steun en werkplezier gegeven.
Mijn vriendinnen Anneke, Daphne, Debbie, Ingrid, Martine, Natasja en Thessa, bij wie ik altijd mijn hart kan luchten onder het genot van een weldadige kaasfondue, wisten mij altijd weer op te peppen in moeilijke tijden. Mijn familie en schoonfamilie wil ik graag bedanken voor hun begrip en steun tijdens het schrijven van dit proefschrift. In het bijzonder wil ik graag mijn lieve ouders bedanken voor hun liefde en het feit dat ze altijd achter mij staan.
Het moge bekend zijn, lieve Marco, dat het zonder jouw liefde en enorme geduld allemaal nooit was gelukt.
Marian