

Genome-wide Linkage Study for Plasma HDL Cholesterol Level in an Isolated Population of Mongolia

Hansoo Park¹, Jong-II Kim^{1,2}, Sung-II Cho³,
Joonhoo Sung³, Hyung-Lae Kim⁴, Young Seok
Ju¹, Gombojav Bayasgalan⁵, Mi-Kyeong Lee²
and Jeong-Sun Seo^{1,6*}

¹Department of Biochemistry and Molecular Biology, Seoul National University College of Medicine, Seoul 110-799, Korea, ²Psoma Therapeutics Inc., Seoul 110-799, Korea, ³Seoul National University School of Public Health, Seoul 110-799, Korea, ⁴Ewha Women's University, Seoul 120-750, Korea, ⁵Yonsei University College of Medicine, ⁶Macrogen Inc., Seoul 153-023, Korea

Abstract

High-density lipoprotein (HDL) whose primary role is to transport cholesterol from peripheral tissues to the liver, is associated with the incidence of coronary heart disease. We analyzed HDL cholesterol levels in a genetically isolated population of extended Mongolian families. A total of 1002 individuals (54.5% women) from 95 families were enrolled. After genotyping by use of 1000 microsatellite markers, we performed a genome-wide linkage search with variance component analysis. The estimated heritability of HDL cholesterol was 0.45, revealing that HDL cholesterol was under significant genetic influence. We found peak evidence of linkage (LOD score=1.88) for HDL cholesterol level on chromosome 6 (nearest marker *D6S1660*) and potential evidences for linkage on chromosomes 1, 12 and 19 with the LOD scores of 1.32, 1.44 and 1.14, respectively. These results should pave the way for the discovery of the relevant genes by fine mapping and association analysis.

Keywords: chromosome, genome-wide linkage search, heritability, HDL cholesterol

Introduction

Cholesterol is a major part of cell membranes. Cholesterol is carried in the blood by chylomicrons, very low density lipoproteins (VLDL), high density lipoproteins (HDL) and low density lipoproteins (LDL) (Dastani *et al.* 2006). HDL cholesterol is reversely associated with car-

diovascular disease, and is more tightly controlled by genetic factors than the other lipoproteins such as LDL, VLDL and chylomicrons. Environmental factors including chronic alcoholism, estrogen replacement therapy, and exercise influence the levels of HDL cholesterol. Several families with strikingly elevated HDL cholesterol levels have been identified. HDL cholesterol levels are higher in blacks compared with whites and HDL cholesterol levels of females are higher than those of males (Barcat *et al.* 2006; Brousseau *et al.* 2004; Yamashita *et al.* 2000; Imperatore *et al.* 2000).

Candidate gene analysis using population-based case-control studies has been used to test the association between SNPs and HDL cholesterol levels. Among the candidate genes selected mainly from lipid metabolism pathways, *ApoA-I* gene is the one most intensively studied (Inazu *et al.* 1994; Kuivenhoven *et al.* 1997). By genome-wide linkage analysis, susceptibility genes can be identified although the genes are not candidates based on lipid metabolism. Genome-wide linkage scans are conducted by use of microsatellite markers to identify genetic determinants affecting the traits (Wang and Paigen 2005). Using HDL cholesterol levels as either discrete or quantitative trait, several linkage studies on genetic determinants of HDL cholesterol have been reported (Yancey *et al.* 2003). Genetic effects on the variations in HDL cholesterol were studied mainly in Caucasians and Africans thus far, and little attention has been focused in this regard on Asian populations. We found suggestive evidence for linkage for HDL cholesterol on chromosome 6, 1, 12 and 19, in studies conducted as part of GENDISCAN study, a large epidemiological study of Complex traits in geographically, culturally and genetically isolated large Mongolian families I in Dornod, Mongolia report.

Methods

Subjects, medical histories, genotyping, and measurement of HDL cholesterol

We analyzed data from 1002 Mongolian individuals from 95 large extended families. Informed consent was obtained from all subjects prior to participation and the protocol was approved by the Institutional Review Board at Seoul National University. Potentially confounding variables were assessed for each participant along with overall medical history. Information on age,

*Corresponding author: E-mail jeongsun@snu.ac.kr
Tel +82-2-764-9444, Fax +82-2-740-4534
Accepted 14 March 2008

gender and anthropometry (height, weight, waist circumference, hip circumference and body fat content) were obtained for each individual. Height in centimeter (cm) and weight in kilograms (kg) were measured using an automatic measuring instrument (IMI 1000, Immanuel Elec., Korea). Body mass index (BMI) was calculated in kg/m^2 . Waist circumference was measured to the nearest centimeter at the level of the umbilicus, and hip circumference was measured at the level of the maximal circumference of the gluteus. All other variables were collected through interviews performed by trained interviewers. Information about amount of alcohol and smoking was also obtained from all the participants.

All the subjects were asked to fast for 12 hours before

their visit. Blood samples were collected from an ante-cubital vein into vacutainer tubes containing EDTA. Blood samples were centrifuged at 3000 rpm for 10 minutes and then stored at -70°C . DNA was isolated from lymphocytes for polymerase chain reaction (PCR) and automated genotyping. A 10 ml blood sample was collected from each participating individual for genomic DNA extraction. DNA was extracted from peripheral lymphocytes using the PUREGENE DNA Purification Kit for whole blood (Gentra Systems Inc, USA). For genotyping, a set of 1000 microsatellite markers deCODE mapping sets (deCODE genetics, USA) was used covering the genome at an average density of 3 centimorgans (cM).

HDL cholesterol was measured by the enzymatic method using Cholestest-N-HDL kit (DAICHI, JAPAN) and HITACHI 7600-210 & HITACHI 7180 instruments. Extensive quality control procedures ensured the validity and reproducibility of the measurements.

Table 1. Baseline demographic characteristics of the study population

Characteristics	Value
No. of families	95
Mean \pm SD no. of family members	15,68 \pm 12,44
No. of genotyped subjects	1,002
Mean \pm SD age (in years)	30,63 \pm 15,36
Percentage female (%)	54,5
Total cholesterol (mg/dl)	159,82 \pm 37,19
HDL (mg/dl)	55,19 \pm 11,95
LDL (mg/dl)	90,51 \pm 28,79
Triglyceride (mg/dl)	63,30 \pm 29,93
Height (cm)	155,78 \pm 11,11
Weight (kg)	57,95 \pm 15,55
Body mass index (kg/m^2)	23,35 \pm 4,47
Waist circumference	78,17 \pm 12,71
Hip circumference	92,73 \pm 11,25
Fasting glucose (mg/dl)	93,87 \pm 11,85
Insulin	8,54 \pm 4,10
Alcohol (%) [*]	151 (14,7)
Smoking (%) [†]	221 (21,19)

^{*}Habitual alcohol intake (in %; No=0 and Yes=1). [†]Current smoker (in %; No=0 and Yes=1).

Table 2. Pedigree characteristics of the study population

Characteristics	Value
No. of study pairs of :	
Parent-offspring	1,812
Full-sib pairs	734
Sisters	198
Brothers	167
Brother and sister	369
Half-sib pairs	395
Grandparent-grandchild pairs	1,202
Avuncular pairs	888
First-cousin pairs	598

Statistical analyses, heritability estimation, and variance component linkage analysis

Multiple linear regression analysis was used by PC SAS

Table 3. Relationships between baseline characteristics and HDL cholesterol

Characteristics	Pearson correlation coefficients [*]
Age	0,022 (0,46)
Systolic BP	0,008 (0,77)
Alcohol	-0,005 (0,86)
Mean values (SD)	
Gender (t-test) [*]	< ,0001
Women (544)	56,88 \pm 11,98
Men (458)	53,38 \pm 11,67
Smoking status (t-test) [†]	0,045
Yes	53,72 \pm 11,79
No	55,54 \pm 11,98

^{*}p values reflect test of subgroup differences (women *versus* men). [†]p values reflect test of subgroup differences (Smoking yes *versus* no).

Table 4. Heritability of HDL

Parameters	Value
Fully adjusted Heritability [*]	0,45
p-value	< ,0001
Kurtosis	0,59

^{*}Includes adjustment for age, gender, systolic BP, smoking and alcohol.

version 8,2 and PC SPSS version 12 to account for effect of confounding variables. Pedigree data was managed by PedSys (Southwest Foundation for Biomedical Research, San Antonio, Texas, USA). Nonpaternity was examined using PEDCHECK (Mcpeek and Sun 2000) and relationships other than paternity were checked using average IBD-based method by PREST. After correcting pedigree error and Mendelian errors, non-mendelian errors were examined and corrected using SimWalk. Identity by descent (IBD) matrix between every relationship pairs in family was calculated and IBD matrix for single marker was calculated by SOLAR (Sequential Oligogenic Linkage Analysis Routines software version 2.1.4). Multipoint IBD matrices were computed on every

1 cM distance using Markov chain Monte Carlo method by LOKI (Heath 1997).

Genetic components of selected phenotypes were estimated in terms of heritability. Narrow sense heritability, defined as the proportion of total phenotypic variation due to additive genetic effects, was calculated. Heritability of HDL cholesterol adjusted for age, gender, age-square, product of age and gender, product of age-square and gender, systolic BP, smoking and alcohol was estimated and a variance component linkage analysis was carried out by SOLAR which uses maximum likelihood methods to estimate variance components for the polygenic genetic effect and random individual environmental effects.

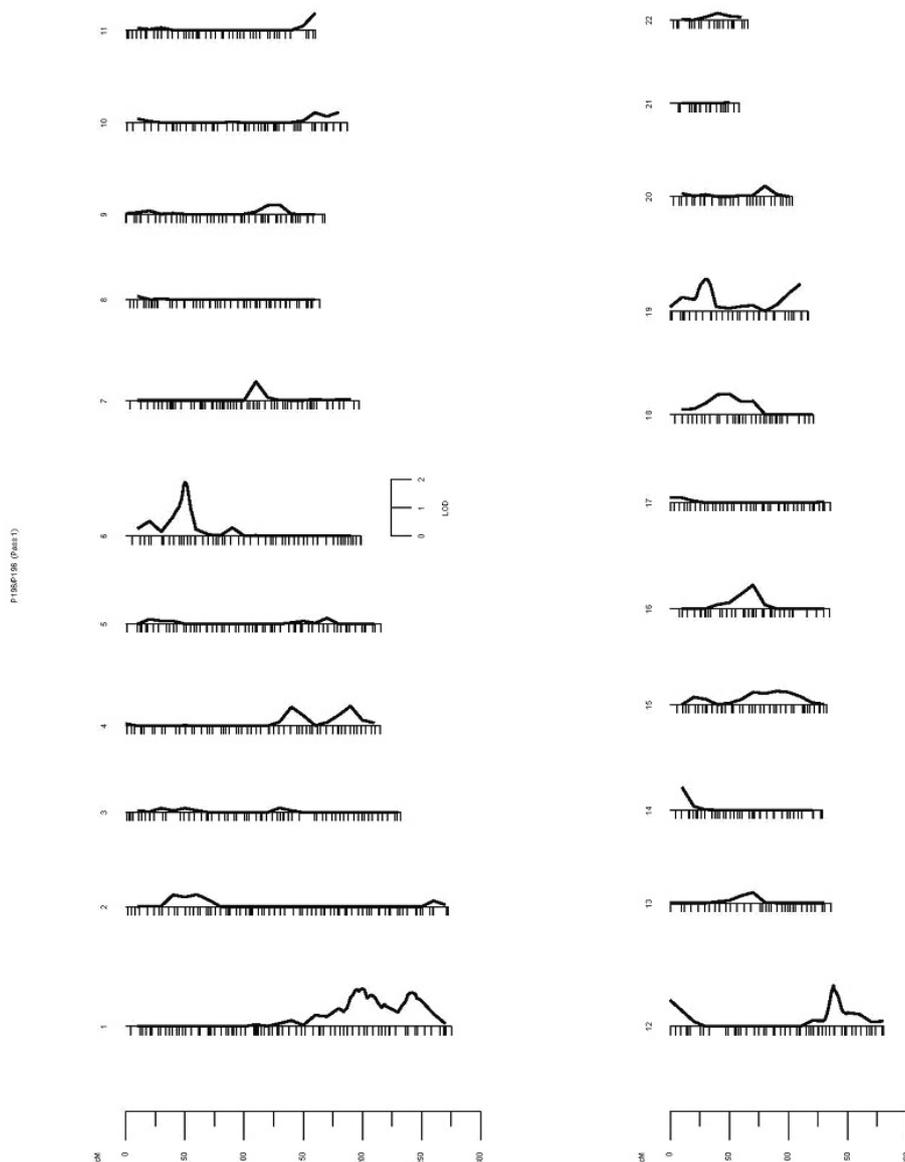


Fig. 1. Multipoint variance component linkage analyses for HDL cholesterol. The X axis represents the genetic location according to the deCODE genetic map and the Y axis represents the multipoint LOD score. HDL cholesterol was adjusted for age, gender, age-square, product of age and gender, product of age-square and gender, systolic BP, smoking and alcohol.

Results and Discussion

The mean age of the 1002 individuals was 31 years and 54.5% of them were female. Demographic and pedigree characteristics of the study sample are shown in Table 1. The family size had a mean of 16. Table 2 included information on 2546 pairs of first degree relatives (1812 parent-offspring pairs and 734 full-sib pairs), 2485 pairs of their second degree relatives (395 half-sibling pairs, 1202 grandparent-grandchild pairs, and 888 avuncular pairs), and 598 first-cousin pairs. Means of their total cholesterol, HDL cholesterol, LDL cholesterol, and triglyceride were 159.82 mg/dl, 55.19 mg/dl, 90.51 mg/dl,

and 63.30 mg/dl, respectively. Table 3 shows correlation between HDL cholesterol and covariates such as age, gender, systolic blood pressure, alcohol consumption status, and smoking status. These parameters were used as covariates in the variance component analysis which provided multivariable adjusted heritability estimates for HDL cholesterol of 0.45 (Table 4). The peak multipoint LOD score was 1.88 on 6p21 (nearest marker D6S1660) and a secondary peak (LOD score of 1.44) was found on 12q23 (nearest marker D12S354). We identified other potential evidence for linkage in the LOD score of 1.32 on 1q24 (nearest marker D1S412) and a LOD score of 1.14 at 19p13 (nearest marker D19S884) (Fig. 1, 2). Table 5 presents all LOD scores >1.0 for HDL cholesterol.

We identified potential evidence of linkage on several chromosomes. In other genome scan, a weak linkage signal for HDL cholesterol was observed for regions that overlapped slightly with the regions identified herein. Klos *et al.* reported the appearance of peak position in the chromosome 12q in European American population (Klos *et al.* 2001) (Table 6). We found evidence of link-

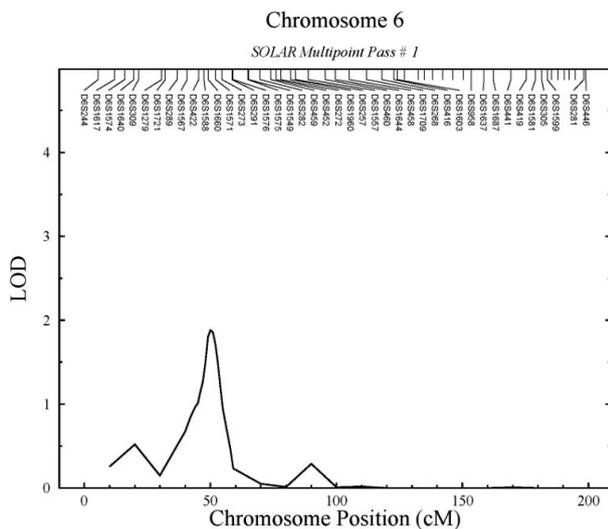


Fig. 2. Multipoint LOD scores for HDL cholesterol on chromosome 6p21. HDL cholesterol was adjusted for age, gender, age-square, product of age and gender, product of age-square and gender, systolic BP, smoking and alcohol.

Table 5. LOD scores, chromosomal locations, and nearest marker data for all LOD scores >1.00*

Trait	Marker at or Near Peak	Chromosome Location	Peak Position (cM)	Maximum LOD Score
HDL	D1S412	1q24	200	1.32
	D6S1660	6p21	50	1.88
	D12S354	12q23	138	1.44
	D19S884	19p13	30	1.14

*Includes adjustment for age, gender, systolic BP, smoking and alcohol.

Table 6. Loci for HDL cholesterol identified by genome-wide linkage scan

Authors and year	Population	No. of families	No. of subjects	Selection criteria	Locus
Seo <i>et al</i>	Asian (GENDISCAN study)	95	1,002	None	Chr 1q24 Chr 6p21 Chr 12q23 Chr 19p13
Mahaney <i>et al</i> 2003	Mexican Americans	10	472	None	Chr 16q22-24
Klos <i>et al</i> 2001	Rochester, MN, USA	232	1,484	None	Chr 12, 17
Peacock <i>et al</i> 2001	FHS, UHFTS, Minneapolis, MN, USA, ARIC	101	1,027	None	Chr 5p13.1, 13q, 13.2
Almasy <i>et al</i> 1999	Mexican Americans	10	477	None	Chr 8, 15
Dastani <i>et al</i> 2006	French Canadian	13	362	Low HDL chol (<5%)	Chr 4q31.21

FHS, Framingham Heart Study; UHFTS, Utah Health Family Tree Study; ARIC, ARIC, Forsyth County, NC, field centers of the Atherosclerosis Risk in Communities Study.

age at similar region in the chromosome 12q23. This locus is reported to be associated with variation in plasma apoA-I level. Given the fact that the subjects of other studies were primarily of European and African origins, whereas those in the present study were of Asian origin, the loci on the chromosome 1, 6, and 19 appear to be potential susceptibility loci for the variation of HDL cholesterol in our Asian ethnic group, with likely smaller effects for European and African American compared with Asian (Peacock *et al.* 2001; Almasy *et al.* 1999).

Several candidate genes influencing HDL cholesterol variation, reside at chromosome 1q24 region. Microsomal glutathione S-transferase 3 catalyzes glutathione-dependent peroxidase activity towards lipid hydroperoxides (Mari and Cederbaum 2001). HDL cholesterol may be more rapidly oxidized than LDL *in vivo*. HDL is the principal vehicle for circulating plasma lipid hydroperoxides (Bowry, Stanley, and Stocker 1992). A specific gene called SRY (sex determining region Y)-box 4 is located at chromosome 6p21 region. This gene encodes a member of the SOX (SRY-related HMG-box) family of transcription factors. It plays a major role in the regulation of embryonic development and may function in tumorigenesis as well. (Forwood, Harley, and Jans 2001). TRAF and TNF receptor-associated protein encoded by *TTRAP* associates with CD40, tumor necrosis factor (TNF) receptor-75 and TNF receptor associated factors (TRAFs), and inhibit nuclear factor-kappa-B activation.

We have found four loci for HDL cholesterol, covering a large portion of the genome. The genomic regions identified in this study need further studies involving fine mapping of these regions. The region is very large, encompassing more than 3 cM of genetic length and containing between 3-5 megabases of DNA, with several genes. After genotyping at less than 1 cM intervals using SNP chip and identifying the polymorphisms that are likely to have functional significance for HDL cholesterol variation, we can confirm the polymorphisms by association study.

Several other genetic studies attempted to identify the genes contributing to HDL cholesterol variation. These studies focused on families with extreme HDL cholesterol levels (Shoulders, Jones, and Naoumova 2004; Aouizerat *et al.* 1999). Our study as did some others, examined families with members having serum HDL cholesterol within the normal range (Mahaney *et al.* 2003; Pajukanta *et al.* 2003; Arya *et al.* 2002). When homogeneous populations are studied, environmental variations might be lower and the genetic makeup of these populations less complex (Coon *et al.* 2001). GENDISCAN study is the first and largest family study of Asian populations. The uniqueness of this study is that study

subjects are members of large extended families in an isolated rural area, long exposed to similar environmental conditions. The component of the GENDISCAN study, we describe here, provides statistical support for linkage of the level of HDL cholesterol. It should provide the basis for identification of genetic polymorphism related with lipid metabolism in Asian population.

Acknowledgements

We would like to thank all the individuals and their families for participation in this study. Korean government approved and has supported the pilot study, and phase I study (Ministry of Science and Technology, Grant # M10306030002-03B4803-00210). Total financial support from the government in 2006 was about 800,000 U.S. Dollars.

References

- Almasy, L., *et al.* (1999). Human pedigree-based quantitative-trait-locus mapping: localization of two genes influencing HDL-cholesterol metabolism. *Am. J. Human Genetics* 64, 1686-1693.
- Aouizerat, B.E., *et al.* (1999). A genome scan for familial combined hyperlipidemia reveals evidence of linkage with a locus on chromosome 11. *Am. J. Human Genetics* 65, 397-412.
- Arya, R., *et al.* (2002). Linkage of high-density lipoprotein-cholesterol concentrations to a locus on chromosome 9p in Mexican Americans. *Nature Genetics* 30, 102-105.
- Barcat, D., *et al.* (2006). Combined hyperlipidemia/hyperalphalipoproteinemia associated with premature spontaneous atherosclerosis in mice lacking hepatic lipase and low density lipoprotein receptor. *Atherosclerosis* 188, 347-355.
- Bowry, V.W., Stanley, K.K., and Stocker, R. (1992). high-density-lipoprotein is the major carrier of lipid hydroperoxides in human blood-plasma from fasting donors. *Proceedings of the National Academy of Sciences of the United States of America* 89, 10316-10320.
- Brousseau, M.E., *et al.* (2004). Effects of an inhibitor of cholesteryl ester transfer protein on HDL cholesterol. *N. Engl. J. Med.* 350, 1505-1515.
- Coon, H., *et al.* (2001). Genome-wide linkage analysis of Hypertension Genetic Epidemiology Network (HyperGEN) Blood Pressure study. *Arteriosclerosis Thrombosis and Vascular Biology* 21, 1969-1976.
- Dastani, Z., *et al.* (2006). Genetics of high-density lipoproteins. *Current Opinion in Cardiology* 21, 329- 335.
- Forwood, J.K., Harley, V., and Jans, D.A. (2001). The C-terminal nuclear localization signal of the sex determining region Y (SRY) high mobility group domain mediates nuclear import through beta 1. *Journal of Biological Chemistry* 276, 46575-46582.

- Heath, S.C. (1997). Markov chain Monte Carlo segregation and linkage analysis for oligogenic models. *American Journal of Human Genetics* 61, 748-760.
- Imperatore, G., *et al.* (2000). A locus influencing total serum cholesterol on chromosome 19p - Results from an autosomal genomic scan of serum lipid concentrations in Pima Indians. *Arteriosclerosis Thrombosis and Vascular Biology* 20, 2651-2656.
- Inazu, A., *et al.* (1994). Genetic cholesteryl ester transfer protein-deficiency caused by 2 prevalent mutations as a major determinant of increased levels of high-density lipoprotein cholesterol. *Journal of Clinical Investigation* 94, 1872-1882.
- Klos, K.L., *et al.* (2001). 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. *Arteriosclerosis Thrombosis and Vascular Biology* 21, 971-978.
- Kuivenhoven, J.A., *et al.* (1997). Heterogeneity at the CETP gene locus - Influence on plasma CETP concentrations and HDL cholesterol levels. *Arteriosclerosis Thrombosis and Vascular Biology* 17, 560-568.
- Mahaney, M.C., *et al.* (2003). A quantitative trait locus on chromosome 16q influences variation in plasma HDL-C levels in Mexican Americans. *Arteriosclerosis Thrombosis and Vascular Biology* 23, 339-345.
- Mari, M., and Cederbaum, A.I. (2001). Induction of catalase, alpha, and microsomal glutathione S-transferase in CYP2E1 overexpressing HepG2 cells and protection against short-term oxidative stress. *Hepatology* 33, 652-661.
- Mcpeck, M.S., and Sun, L. (2000). Statistical tests for detection of misspecified relationships by use of genome-screen data. *American Journal of Human Genetics* 66, 1076-1094.
- Pajukanta, P., *et al.* (2003). Combined analysis of genome scans of Dutch and Finnish families reveals a susceptibility locus for high-density lipoprotein cholesterol on chromosome 16q. *American Journal of Human Genetics* 72, 903-917.
- Peacock, J.M., *et al.* (2001). Genome scan for quantitative trait loci linked to high-density lipoprotein cholesterol - The NHLBI family heart study. *Arteriosclerosis Thrombosis and Vascular Biology* 21, 1823-1828.
- Shoulders, C.C., Jones, E.L., and Naoumova, R.P. (2004). Genetics of familial combined hyperlipidemia and risk of coronary heart disease. *Human Molecular Genetics* 13, R149-R160.
- Wang, X.S., and Paigen, B. (2005). Genome-wide search for new genes controlling plasma lipid concentrations in mice and humans. *Current Opinion in Lipidology* 16, 127-137.
- Yamashita, S., *et al.* (2000). Molecular mechanisms, lipoprotein abnormalities and atherogenicity of hyperalphalipoproteinemia. *Atherosclerosis* 152, 271-285.
- Yancey, P.G., *et al.* (2003). Importance of different pathways of cellular cholesterol efflux. *Arteriosclerosis Thrombosis and Vascular Biology* 23, 712-719.