VNIVERSITAT DE VALÈNCIA

FACULTAT DE MEDICINA I ODONTOLOGIA

Departament de Medicina Preventiva i Salut Pública,

Ciències de l'Alimentació, Toxicologia i Medicina Legal

MODULACIÓN POR GENES CANDIDATOS DE FENOTIPOS INTERMEDIOS Y FINALES DE ENFERMEDAD CARDIOVASCULAR EN POBLACIÓN MEDITERRÁNEA Y ALEMANA. APROXIMACIÓN AL ESTUDIO DE LA INTERACCIÓN GEN-DIETA

MODULATION BY MEANS OF CANDIDATE GENES OF INTERMEDIATE AND FINAL PHENOTYPES OF CARDIOVASCULAR DISEASES IN MEDITERRANEAN AND GERMAN POPULATIONS. APPROACH TO THE STUDY OF GENE-DIET INTERACTIONS

TESIS DOCTORAL INTERNACIONAL

Presentada por: María Arregui Rementería

Dirigida por: Dolores Corella Piquer, Oscar Coltell Simón, Olga Portolés Reparaz

Programa de estudios de doctorado: 265E Salud Pública y Medicina Legal

Valencia, 6 de mayo de 2013

Thesis supervised by

Dr. Dolores Corella Piquer

Department of Preventive Medicine and CIBER Fisiopatología de la Obesidad y Nutrición, School of Medicine, Universitat de València. Blasco Ibáñez, 15, 46010, Valencia, Spain. Tel: (+34) 963864417; fax: (+34) 963864166; e-mail: dolores.corella@uv.es.

Dr. Oscar Coltell Simón

Department of Computing Languages and Systems, Universitat Jaume I. Avda Vicent Sos Baynat, s/n. 12071, Castellón, Spain. Tel: (+34) 964728314; fax:(+34) 964728435. e-mail: coltell@lsi.uji.es.

Dr. Olga Portolés Reparaz

Department of Preventive Medicine. School of Medicine. Universitat de València. Blasco Ibáñez, 15, 46010, Valencia, Spain. Tel: (+34) 963864417; fax: (+34) 963864166; e-mail: olga.portoles@uv.es

Universitat de València

Dolores Corella Piquer, Catedrática de Universidad, del Departament de Medicina Preventiva i Salut Pública, Ciències de l'Alimentació, Toxicologia i Medicina Legal de la Universitat de València.

Oscar Coltell Simón, Profesor Titular del Departamento de Lenguajes y Sistemas Informáticos, de la Escuela Superior de Tecnología y Ciencias Experimentales de la Universitat Jaume I.

Olga Portolés Reparaz, Profesor contratado doctor, del Departament de Medicina Preventiva i Salut Pública, Ciències de l'Alimentació, Toxicologia i Medicina Legal de la Universitat de València.

CERTIFICAN:

Que la presente tesis doctoral con el título: "*Modulación por genes candidatos de fenotipos intermedios y finales de enfermedad cardiovascular en población mediterránea y alemana. Aproximación al estudio de la interacción gen-dieta*", ha sido realizada por María Arregui Rementería bajo nuestra dirección, y reúne los méritos suficientes para que su autora obtenga el título de Doctora, en la modalidad de tesis doctoral internacional, por la Universitat de València.

Y para que así conste, firman el presente certificado en Valencia, a 20 de marzo de 2013

Dolores Corella Piquer Oscar Coltell Simon Olga Portolés Reparaz

AGRADECIMIENTOS

Quisiera agradecer su apoyo a todas aquellas personas que han contribuido a que este trabajo sea posible:

A mis directores de tesis Dolores Corella, Oscar Coltell y Olga Portolés, por su dirección, sus siempre acertadas recomendaciones y por proporcionarme todos los medios necesarios para que este trabajo salga adelante.

A mi querida Olga, por ser su amistad, compañía y apoyo incondicional. Por sus sabios consejos, honestidad, paciencia y cariño.

A mis compañeros del departamento de Lenguajes y Sistemas Informáticos de la UJI. Muy en especial a mis queridas "*xiquetes*" Anna y Zoe, con quien sé que seguiré compartiendo alegrías y peripecias; a los "*bioinfogenómicos*" Edu, Eli y Toni;-). También a Gus, mi compi de despacho; a Juanma y a Joaquín por su inestimable ayuda en mis primeros pasitos como docente. Muchas gracias también a Amparo, Irene, Ramón, Pablo, Óscar… por contribuir a que el ambiente en el departamento fuera siempre tan agradable.

A todos mis compañeros del departamento de Medicina Preventiva de la UV. En especial a Paula y Carolina por su amistad, sus ánimos en los momentos difíciles, su pacieeeeeeeencia en el laboratorio :-) y tantas risas juntas! También a Gema, Paco, Marisa, Manolo, Francesc, Javi, Eva… por su amabilidad, cercanía y siempre buen humor.

To my DIfE colleagues, especially professor Boeing, who gave me the great chance of becoming part of the Epi Abteilung; to Conny for the great opportunity, the unconditional support, encouragement, and scientific guidence. To Romina, my Thelma and torella :-) and to Krassi, the best office-mate and friend. Thank you also Hélène, Annika, Janine, Dagmar, Sven, Ellen, Madlen, Gaby… y mi querida Adela, for making of the epidemiology department a great place to work, learn and find friends.

A los participantes de los estudios PREDIMED-Valencia y EPIC-Potsdam y a todas las personas que los hacen posibles.

Finalmente gracias a mis queridísimos padres, Fernando y Rosa María y hermanos, Miguel y Carlos, para quienes las palabras de agradecimiento quedan muy escasas; and to Brian my happiness, support, strength and peace.

Dedicatoria

A mis padres, a quienes debo todo lo que soy

A mis hermanos, los mejores compañeros de vida

To Brian and our baby, the loves of my life

TABLE OF CONTENTS

ABBREVIATIONS

24HDRs 24 hour dietary recalls

ALT Alanine aminotransferase

AMPM Automated Multiple Pass Method

ANCOVA Analysis of covariance

ABCA1 ATP-binding cassette transporter

ATP III Adult treatment panel III

ApoA-I Apolipoprotein A-I

ApoB apolipoproteins B

BMI Body mass index

CAD Coronary artery disease

CAFE Compositional analyses from frequency estimates

CAPI Computer-assisted personal interviewing

CETP Cholesteryl ester transfer protein

C-24HDRs Computerized 24 h dietary recalls

C-FFQs Computerized food frequency questionnaires

CHARGE Cohorts for heart and aging research in genome epidemiology

CHD coronary heart disease

CI Confidence interval

Cm Centimetre

CSIC Consejo Superior de Investigaciones Científicas (Spanish National Research Council)

CVD Cardiovascular disease

D day

DASH Dietary approaches to stop hypertension

DBP Diastolic blood pressure

DGAT Diacylglycerol acyltransferase

DHQ Diet history questionnaire

DIfE German institute of human nutrition

DIDeA Dietary intervention for type 2 diabetes

dL Decilitre

DNA Deoxyribonucleic acid

DZD Deutsches zentrum für diabetesforschung

EEES Espacio Europeo de Educación Superior

EPIC European Prospective Investigation into Cancer and Nutrition

FCTs Food composition tables

FBQ Food behaviour questionnaire

FFQ Food frequency questionnaire

FIRSSt Food intake recording sw system

FNDDS Food and nutrient database for dietary studies

g Grams

GEP Good Epidemiologic Practice

GGT Gamma-glutamyltransferase

GWA Genome-wide association

H Hour

HbA_{1c} Glycated hemoglobin

HCE Historia clínica electrónica

HDL High density lipoprotein

HR Hazard ratio

hs-CRP High-sensitivity C-reactive protein

HWE Hardy–Weinberg equilibrium

ICD-10 International classification of diseases, injuries, and causes of death, tenth revision

IS Ischemic stroke

Kg Kilogram

LDL Low density lipoprotein

LPL Lipoprotein lipase

LCT Lactase

LP Lactase persistence

LNP Lactase non-persistent

MAF Minor allele frequency

MUFA Monounsaturated fatty acids

MDC Max Delbrück center for molecular medicine

MeSH Medical subject heading

MetS Metabolic syndrome

mg Miligram

MI Myocardial infarction

MTTP microsomal triglyceride transfer protein

N Number of participants

mRNA Messenger ribonucleic acid

MUFAs Monounsaturated fatty acids

NCBI National Center for Biotechnology Information

NCI National cancer institute

NDSR Nutrition data system for research

NHANES National Health and Nutrition Examination Survey

NuGO Nutrigenomics organisation

OR Odds ratio

PGH Proyecto genoma humano

PREDIMED Prevención con Dieta Mediterránea

SAS Statistical analysis system

SBP Systolic blood pressure

SCD1 Stearoyl-CoA desaturase-1

SE Standard errors

SFAs Saturated fatty acids

SGBD Sistemas de gestión de bases de datos

SI Synergy Index

SNAPTM Synchronized nutrition and activity program

SNP Single nucleotide polymorphism

P P-value, level of significance

PDAs Personal digital assistants

PDHQ Pictorial diet history questionnaire .

PI3K Phosphatidylinositol-3- kinase

PIPS Post-interview processing system

SD Standard deviation

SEM Means and standard error

SES Socio-economic status

SNP single nucleotide polymorphism

SREBP sterol regulatory element binding protein

SRE sterol response element

SVMs Support vector machines

T2DM type 2 diabetes mellitus

TADA Technology assisted dietary assessment

tag-SNPs Tagging single nucleotide polymorphism

TG Triglyceride

USDA US Department of Agriculture

VLDL Very low density lipoprotein

WC Waist circumference

Web-24HDR Web-based 24 h dietary recalls

Web-FFQs Web-based food frequency questionnaires

WHO World health organisation

WTCCC Welcome trust case control consortium

Wk week

WWEIA What we eat in america

XMCT X-ray computerised microtomography

y Year

YANA-C Young adolescent's nutrition assessment on computer

ABSTRACT

Cardiovascular diseases (CVD) are the main cause of death worldwide, and the long-term consequences of the non-fatal events also represent a great burden. The multifactorial etiology of these complex pathologies, involves the interaction between lifestyle, environmental, and genetic factors. The study of the influence of single nucleotide polymorphisms (SNPs) on intermediate and final CVD phenotypes, is helping recognize the molecular biological means of these diseases.

From all of the modifiable factors of CVD, diet is of particular relevance, both because people eat and drink on a daily basis, and also due to its potential to interact with non-modifiable risk factors, such as SNPs. This thesis presents five studies: four genetic association studies in two different study populations, and a systematic review of informatic tools available to asses diet in epidemiological studies, which facilitate the study of diet as a risk factor for different chronic diseases including CVD.

The **first study** cross-sectionally investigated whether the common genetic variant rs4988235 (C/T), located -13910 pb upstream the gene coding for lactase (*LCT*), was associated with obesity in a Spanish population at high CVD risk, sampled in the Mediterranean area (PREDIMED-Valencia Study, n = 940). Further, its potential modulation by consumption of dairy products was examined. This SNP, which is strongly associated with lactase persistence, is an emerging candidate for obesity. Results from this study suggested that, despite dairy product consumption did not substantially differ by genotype, participants homozygote for the C allele had lower mean body mass index (BMI) (29.7±4.2 vs. 30.6±4.2 kg/m^2; P = 0.003) and waist circumference (101.1±1.8 vs. 103.5±11.5 cm; P = 0.005) than T-allele carriers. Obesity prevalence was also significantly higher in T-allele carriers than in CC individuals, however only among participants consuming moderate or high lactose intakes (>8 g/day).

The **second study** investigated whether the gene variant rs2943634 (C/A), located in a non-coding region of chromosome 2q36.3, and recently associated with coronary artery disease in two GWA studies, was associated with myocardial infarction and ischemic stroke in the Potsdam arm

of the European Prospective Investigation into Cancer and Nutrition (EPIC). A case-cohort design was used, with a subsample of 2500 persons randomly drawn from the total population of 27,548 middle-aged men and women, along with all incident cases of myocardial infarction (n = 211) and ischemic stroke (n = 144), occurring during a mean follow-up of 8.2 years. Because the mode of mechanism by which this SNP may increase CVD risk is unknown, associations of this SNP with 12 available intermediate risk phenotypes of CVD were also investigated. The minor allele of rs2943634 was associated in an additive fashion with lower risk of ischemic stroke after adjustment for age and sex (per-allele hazard ratio 0.66, 95% CI 0.50-0.87) but not with myocardial infarction (per-allele hazard ratio 1.02, 95% CI 0.82-1.28). Furthermore, the minor allele was related to slightly higher levels of plasma adiponectin (CC 6.94, CA 7.27, AA 7.86µg/ml, P=0.0002) and HDL-cholesterol (CC 52.08, CA 53.05 and AA 55.27mg/dl, P=0.002) in an additive fashion. Adjustment for adiponectin and HDL-cholesterol did, however, not attenuate the association between the SNP and ischemic stroke risk. In contrast, adjustment for adiponectin abolished the association between the SNP and HDL-cholesterol and adjustment for HDL-cholesterol attenuated the association between the SNP and adiponectin.

In the **third study**, the impact of common genetic variation in the Stearoyl-CoA desaturase-1 (*SCD1*) gene, captured by means of 7 tagging SNPs and 5 inferred haplotypes, on the modulation of 8 metabolic risk factors linked to the activity of SCD1 (triglycerides, body mass index, waist circumference, glycated haemoglobin, high-sensitivity C-reactive protein, gamma-glutamyltransferase, alanine aminotransferase and fetuin-A) was investigated. SCD1 is the rate limiting enzyme catalyzing the conversion of the endogenous and dietary saturated fatty acids palmitic and stearic into the monounsaturated palmitoleic and oleic, respectively. Its activity has been associated with traits of the metabolic syndrome in mice and humans, but also with the prevention of saturated fatty acids accumulation and subsequent inflammation, whereas for liver fat content inconsistent results have been reported. In the EPIC-Potsdam Study (n = 2157), however, no associations between common variants of *SCD1* or its inferred haplotypes and the investigated metabolic risk factors were observed.

The **fourth study** was conducted in EPIC-Potsdam following a case-cohort design (193 incident myocardial infarction, 131 incident ischemic stroke cases and 1978 non-cases). Further, the Heinz Nixdorf Recall Study (30 CVD cases and 1,188 controls) was used to replicate findings.

The microsomal triglyceride transfer protein (MTTP), encoded by the *MTTP* gene, plays an important role in the assembly and secretion of apolipoprotein-B containing lipoproteins as chylomicrons in the intestine, and of very low density lipoproteins in the liver. *MTTP* is regulated by cholesterol. In this study the investigated hypothesis was that, the -164T>C polymorphism, located in the promoter region of *MTTP*, could modify the risk of CVD, depending on cholesterol levels. This SNP had inconsistently been associated with CVD in previous studies. In the EPIC-Potsdam study, individuals with cholesterol levels <200 mg/dL showed a significant increased risk of CVD ($HR_{additve}$= 1.38, 95% CI: 1.07 to 1.78); while $HR_{additve}$ for subjects with cholesterol levels ≥200 mg/dL was 0.77 (0.58-1.03). $HR_{additive}$ for participants in the Heinz Nixdorf Recall Study were 1.06 (0.33-3.40) and 0.60 (0.29-1.25) respectively for participants with cholesterol levels <200 mg/dL and ≥200 mg/dL. These results suggest that risk allele carriers with low cholesterol levels may be predisposed to an increased risk of developing CVD, which seems to be abolished among risk allele carriers with high cholesterol levels.

In the last years, the tedious task of a assessing a person's diet has started to benefit from the fast development of the information and communication technologies. Thus, a variety of new tools now makes it possible to study diet as a risk factor for different chronic diseases in large epidemiological studies, without too much labour costs. In the **fifth study** of this thesis, a state-of-the-art review of applications for automating the most commonly used dietary surveys in nutritional research, that is food-frequency questionnaires and 24-hour dietary recalls, is provided.

RESUMEN

Las enfermedades cardiovasculares (ECV) constituyen la principal causa de mortalidad en todo el mundo. Su etiología es multifactorial y compleja e incluye la interacción de factores de estilo de vida, ambientales y genéticos. Mediante estudios de asociación acerca de la influencia de polimorfismos de un sólo nucleótido (SNP) sobre fenotipos intermedios y finales de ECV, es posible identificar rutas metabólicas implicadas en ECV. De todos los factores de riesgo modificables de ECV, la dieta es de particular relevancia, tanto porque la gente come y bebe a diario como por su potencial para interactuar con factores de riesgo no modificables, tales como factores genéticos. En esta tesis se presentan cinco estudios: cuatro de ellos de asociación genética en dos cohortes diferentes y una revisión sistemática sobre herramientas informáticas disponibles para la medida de la dieta en grandes estudios epidemiológicos, que facilitan su estudio como factor de riesgo de diversas enfermedades, incluidas las ECV.

El **primer estudio** se llevó a cabo en una muestra de población mediterránea española de alto riesgo cardiovascular (estudio PREDIMED-Valencia, n = 940). Mediante un diseño transversal, se investigó la relación con obesidad de la variante genética rs4988235 (C/T), localizada a 13910 pares de bases del gen que codifica la β-galactosidasa Lactasa (*LCT*). Así mismo de determinó su possible modulación por consumo de productos lácteos. Este polimorfismo, ha sido consistentemente relacionado con la persistencia del enzima lactasa y es además un candidato emergente de obesidad. A pesar de que el consumo de productos lácteos no varió sustancialmente de acuerdo a genotipo, los participantes homocigotos para el alelo C, presentaron menor índice de masa corporal (29.7±4.2 vs 30.6±4.2 kg/m^2; P = 0.003) y perimétro de de cintura (101.1±11.8 vs 103.5±11.5 cm, P = 0.005) que los portadores del alelo T. La prevalencia de obesidad fue también significativamente mayor en los portadores del alelo T, sin embargo, sólo entre aquellos participantes cuyo consumo de lactosa fue moderado o alto (>8g/día).

En el **segundo estudio,** desarrollado en la cohorte de Potsdam del Estudio Europeo Prospectivo sobre Cáncer y Nutrición (EPIC), se investigó la asociación de la variante genética rs2943634 (C/A) con incidencia de infarto de miocardio y accidente cerebrovascular isquémico. El

polimorfismo rs2943634 se localiza en una región no codificante del cromosoma 2q36.3, sin embargo, había sido asociado con enfermedad arterial coronaria en dos GWAS previos. Se utilizó un diseño caso-cohorte, que incluyó una submuestra aleatoria de 2500 participantes y todos los casos de infarto de miocardio (n = 211) y accidente cerebrovascular isquémico (n = 144) ocurridos en EPIC-Potsdam durante un periodo de seguimiento medio de 8.2 años. Dado que el mecanismo por el que rs2943634 podría modular el riesgo de ECV es desconocido, también se investigó de forma transversal su posible asociación con 12 fenotipos intermedios disponibles de ECV. El alelo C fue asociado de manera aditiva con menor riesgo de accidente cerebrovascular isquémico (ratio de riesgo (HR) por alelo 0.66, intervalo de confianza IC al 95%: 0.50-0.87), pero no con infarto de miocardio [1.02 (82-1.28)]. Además los portadores de este alelo, también presentaron mayores niveles plasmáticos de adiponectina (CC 6.94, CA 7.27, AA 7.86 µg/ml, P = 0.0002) y de cholesterol-HDL (CC 52.08, CA 53.05, AA 55.27mg/dl, P = 0.002). Sin embargo, el ajuste por adiponectina y cholesterol-HDL de los análisis de asociación del SNP con accidente cerebrovascular hisquémico, no atenuó la asociación.

El enzima Estearoil-CoA Desaturasa-1 (SCD1) cataliza la conversión de los ácidos grasos saturados palmítico y esteárico a los monoinsaturados palmitoléico y oléico, respectivamente. Su actividad se ha asociado con rasgos del síndrome metabólico en ratones y seres humanos, pero también con la prevención de la acumulación de ácidos grasos saturados y posterior inflamación, mientras que para contenido de grasa del hígado los resultados hasta ahora publicados son inconsistentes. En el **tercer estudio** de esta tesis, en una muestra aleatoria (n = 2500) de la cohorte del estudio EPIC-Potsdam, a través de 7 tagging-SNPs y 5 haplotipos inferidos, se investigó el efecto de la heterogeneidad genética común del gen que codifica la enzima SCD1 (*SCD1*), en la modulación de 8 factores de riesgo metabólico vinculados a la actividad de: SCD1: triglicéridos, índice de masa corporal, circunferencia de cintura, hemoglobina glicosilada, proteína-C-reactiva, gamma-glutamil transferasa, alanina aminotransferasa y fetuina-A. No se encontraron asociaciones estadísticamene significativas entre los genotipos y fenotipos estudiados.

La proteína microsomal transferidora de triglicéridos (MTTP), participa en el ensamblaje y secreción de lipoproteínas con apolipoproteína-B. Esta proteína está codificada por el gen *MTTP*, que se regula mediante colesterol. Polimorfimos en la zona promotora de este gen, tales como el -

164 T>C, han sido inconsistentemente asociados con ECV en distintos estudios. En el **cuarto estudio** de esta tesis, utilizando un diseño caso-cohorte, se investigó la asociación del polimorfismo -164T> C con riesgo de padecer un evento cardiovascular, en función de los niveles de colesterol (subcohorte, n = 1978; casos de infarto de miocardio, n = 193; casos de ictus isquémico, n = 131). Los resultados obtenidos fueron replicados en la cohorte del estudio Heinz Nixdorf Recall (30 casos, 1188 controles). Los participantes de EPIC-Potsdam con niveles de colesterol inferiores a 200 mg/dL, mostraron un incremento significativo del riesgo cardiovascular [$HR_{aditivo}$ 1.38 (1.07-1.78)], mientras que para participantes con niveles de colesterol ≥200 mg/dL se observó una disminución del riesgo [$HR_{additvo}$ = 0.77 (0.58-1.03)]. En el estudio Heinz Nixdorf Recall, los respectivos HR para niveles de colesterol <200 mg/dL y colesterol ≥200 mg/dL fueron 1.06 (0.33-3.40) y 0.60 (0.29-1.25). Estos resultados sugieren una interacción entre el SNP investigado y colesterol, de manera que los portadores del alelo de riesgo, en presencia de niveles elevados de colesterol podrían evitar el riesgo conferido por el SNP.

En los últimos años, la tediosa tarea de evaluar la dieta de una persona, se ha beneficiado del rápido desarrollo de las tecnologías de la información y la comunicación. Con la nueva variedad de herramientas, ahora el estudio de la dieta como factor de riesgo para diferentes enfermedades crónicas en grandes estudios epidemiológicos es posible a costes inferiores. En el **quinto estudio** de esta tesis, se presenta una revisión de las aplicaciones existentes para la automatización de las encuestas más comúnmente utilizadas para la medida de la dieta en estudios de investigación, es decir, el cuestionario de frecuencia de consumo de alimentos y el recordatorios de 24 horas.

1

Introduction

Chapter 1: Introduction

1.1 CARDIOVASCULAR DISEASES

Cardiovascular diseases (CVD) are pathological conditions involving the cardiovascular system, i.e., the heart, blood vessels and pericardium. They have been described as a continuum that begins with the presence of risk factors, and proceeds via progressive vascular disease to target organ damage, end-organ failure, and death (1). CVD are the main cause of death worldwide (2) and the long-term consequences of the non-fatal events also represent a great burden (3). According to the 10th revision of the International Classification of Diseases (ICD-10) (4), CVD can be classified in 10 categories (I00-I99) of which, in terms of mortality, two subcategories, myocardial infarction (MI) (I21), followed by ischemic stroke (IS) (I63) are the most important with about 12.8 million people worldwide dying from them in 2010 (2).

MI and IS are complex multifactorial diseases that are thought to occur due to the interaction of lifestyle, environmental, and genetic factors. They mostly arise when arteries that supply blood to the heart or the brain are obstructed, which is usually instigated by the disruption of atherosclerotic plaques in the wall of the arteries, and subsequent activation of a clotting cascade (5).

1.2 RISK FACTORS FOR CARDIOVASCULAR DISEASES

Because CVD are such a large public health problem, many efforts have been taken worldwide to identify its modifiable risk factors. Regarding MI, a remarkable initiative is the INTERHEART study (6). In this standardized case-control study including 15152 acute MI cases and 14820 controls from 52 countries representing all inhabited continents, nine modifiable risk factors were identified that accounted for 90% of the population attributable risk to develop MI. Smoking and a raised ratio of the apolipoprotein B to A1 (respectively the major protein components of low-density lipoprotein (LDL) cholesterol and high-density lipoprotein (HDL) cholesterol) were the dominating factors. These were followed by, in different degrees of importance according to geographic location, psychosocial factors (which included depression, locus of control, perceived

stress, and life events), abdominal obesity (approximated by the waist-to-hip ratio), diabetes, history of hypertension, low consumption of fruit and vegetables, lack of regular physical activity, and consumption of alcohol less than three times per week (6).

Concerning stroke, a similar initiative, although smaller in size, was undertaken in the INTERSTROKE study (7). This case-control study of 3000 stroke cases (of which 78% ischemic stroke, the most common form of stroke) and 3000 controls without stroke, from 22 countries worldwide, identified ten modifiable risk factors for IS also accounting for 90% of the population attributable risk. Consistent with the fact that MI and IS share common pathological pathways due to atherosclerotic disease (8), all the nine risk factors identified for MI were also identified for IS, although with different relative importance. The list of risk factors for stroke was completed with cardiac causes, with atrial fibrillation as the most common source (7).

Apart from the INTERHEART and INTERSTROKE initiatives, which despite their large sample sizes were case-control studies and therefore sensitive to bias, results of a number of prospective studies also support a role for the following CVD risk factors: cigarette smoking (9-15); dyslipidemia (16-21), obesity (22-24); diabetes mellitus mellitus (25-31), hypertension (32-35), physical activity (36, 37); psychosocial factors (38-41); the consumption of fruit and vegetables (42-44); and alcohol intake (45-50). Additionally, results from interventional studies have confirmed the cardiovascular benefits of dietary modification (51, 52), antihypertensive treatments (34, 53, 54), correction of dyslipemia (55-60), moderate alcohol use (61) and also of anticoagulation treatments for stroke prevention in atrial fibrillation (62).

Further, the so-called metabolic syndrome (MetS), a clustering of abdominal obesity, high triglycerides (TG), low HDL-cholesterol, elevated blood pressure, and glucose intolerance (63), has been associated with a ~2-fold increased risk for CVD (64) (65), and it has been shown that the higher the number of individual MetS traits a person has, the greater the risk of developing an adverse outcome (66).

Other suggested risk factors that contribute to the development of CVD include, but are not limited to, the aging (67), chronic low-grade inflammation (68, 69), sleeping less than 6 or more than 9 hours (70-73) non-alcoholic fatty liver disease (74) or chronic kidney disease (75). Hypoadiponectinemia has also been suggested as a possible independent risk factor, given that

adiponectin exerts anti-inflammatory, antiatherogenic, antithrombotic and insulin-sensitizing effects (76, 77).

1.3 PARTICULAR RELEVANCE OF DIET AS A RISK FACTOR FOR CVD

From all the potentially modifiable established CVD risk factors, diet is of particular relevance given that it is the only factor everyone is exposed to from birth to death, and also for its potential to interact with non-modifiable risk factors such as genetic polymorphisms. Supporting the results from INTERHEART and INTERSTROKE, adherence to the traditional Mediterranean diet, characterised by high intakes of fruits, vegetables, legumes, nuts, unrefined cereals, olive oil, a moderate intake of alcohol (particularly in the form of red wine), and a low to moderate intake of dairy products, fish, meat, poultry, and saturated fats (78), has consistently been associated with reduced risk of developing CVD (79).

The assessment of a person's typical diet should consist of the analysis of his or her daily food intake throughout one or more years. In large epidemiological studies, however, this is not feasible and, in practice, a portion of it is captured and habitual dietary intake is extrapolated. Tools to collect dietary data in these settings, which can then be transformed into energy and nutrient intake by means of food composition tables, include: food-frequency questionnaires (FFQ), 24-hour dietary recalls (24HDRs), dietary records (DRs), and dietary histories. As the information and communication technologies gained importance in the last years, a great effort has been done to automate questionnaires used in epidemiological studies, with the aim to reduce random errors and save costs by accelerating the tasks of extraction and processing of data. First, computer programs were developed for the interviewer-administration or self-administration of FFQs, 24HDRs or combinations of both. Later, when access to the Internet became widespread, Web-based on-line FFQs, on-line 24HDRs, and combinations of both were introduced to substitute paper questionnaires. In Chapter 2 of this thesis a FFQ is used to estimate dairy intake in a population-based study. Characteristics of dietary surveys and methods developed for their automation in medical studies are reviewed in Chapter 5.

1.4 ROLE OF GENETICS IN THE ETIOLOGY OF CVD

Lifestyle factors are known to influence certain biological risk factors of CVD. It is thought, however, that biological risk factors also have a genetic component. This is supported by the observation that modifiable risk factors have a certain degree of heritability and genetic modulation (80). Also, parental history has been proposed as an independent risk factor for MI (81) and ischemic stroke (82, 83). Thus, in order to recognize persons at risk, it is also essential to identify genetic factors conferring susceptibility (80).

Except for rare forms that follow a Mendelian inheritance pattern (84), CVD are multifactorial diseases influenced by many genetic loci interacting with each other and with lifestyle and environmental factors (85). The increased risk can be driven by epigenetic changes, but also by mutations such as deletions, insertions, variable number tandem repeat or single nucleotide polymorphisms (SNPs). SNPs are the most common genetic variations, and occur on average once in every 300 nucleotides. Most of them have no direct effect on CVD, but some may modulate intermediate or final CVD phenotypes. Their influence on MI and IS is not fully understood. However, the development of studies testing for heterogeneity of SNPs across persons with different phenotypes, such as candidate-gene studies, gene-wide studies, and particularly genome-wide association (GWA) studies, is leading to the identification of several gene loci associated with an elevated risk of CVD, and thus helping recognize its molecular biological means (86).

Candidate gene studies are hypothesis-based. These studies test the hypothesis that heterogeneity of a given gene, typically encoding a protein that determines or influences an intermediate phenotype, may modulate its function, and eventually, the risk of clinical disease. A limitation of this approach is that only a few SNPs are examined (80), thereby missing other potential important SNPs. The gene-wide strategy overcomes this limitation by assessing the smallest number of SNPs (tagging-SNPs) necessary to extract the maximum amount of information for a given gene, taking into account allele frequencies, linkage disequilibrium, and position on the gene. Gene-wide studies also provide the possibility to predict haplotype structures, thereby capturing more genetic variation than assessing single SNPs (80).

To date, many candidate-gene and gene-wide studies examining the association of genetic factors with cardiovascular endpoints are underpowered (87), and few have been replicated (88). However, these studies have been successful for some intermediate phenotypes. Some examples include the identification of the genes encoding the apolipoproteins A, E and A5, which regulate levels of lipoprotein-A, cholesterol and triglycerides, respectively (87).

If the etiology of a disease is unknown, the candidate-gene and the gene-wide approaches are not useful. This limitation is overcome by the GWA approach, which allows examining genetic variants across the entire genome without assumptions about the genomic location of the causal variants. In fact, GWA studies have led to the discovery of unsuspected pathological pathways (80). However, although to date GWA studies are the most reliable type of genetic studies (80), they also often show the limitation of lack of reproducibility.

Regarding MI risk, eleven chromosomal regions identified in GWA studies have been successfully replicated (86), making the evidence for the associations more robust. These regions were summarized by Erdman et al. (86), and are reproduced in Table 1 of this chapter. Only two of them (1p13.3 and 6q26–27) are associated with traditional risk factors (LDL-cholesterol and lipoprotein-A). For the remainder, the associations with the risk of MI appear not to be mediated by traditional biological risk factors and their mechanisms are not yet understood (perhaps long-range regulation of gene expression). The strongest and most replicated genetic effect is that of a genetic variant in a non-coding region of chromosome 9p21.3 (86).

Furthermore, results from a recent meta-analysis including 8140 MI cases and 10522 controls of Caucasian, Asian, and African-American origin, from 30 case-control studies, suggested that the SNP rs1801133 of the gene coding for the enzyme methylenetetrahydrofolate reductase (MTHFR) is related to risk of MI. This enzyme catalyzes the conversion of 5,10-methylenetetrahydrofolate to 5-methyltetrahydrofolate, a co-substrate for homocysteine remethylation to methionine (89). Another meta-analysis including 9569 MI cases and 7264 controls of Asian and Caucasian origin from 18 case-control studies, indicated that, depending on the ethnic origin, SNPs rs3025058 and rs3918242 from the genes coding for matrix metalloproteinases 3 and 9 respectively, also confer susceptibility to develop MI. These enzymes weaken the arterial wall and contribute to the destabilization and rupture of atheromatous plaque,

by attacking the major components of the basal lamina around blood vessels (that is, type IV collagen, laminin and fibronectin) (90).

Concerning the genetic susceptibility to develop IS, the largest meta-analysis so far conducted, the METASTROKE collaboration, included 12389 IS cases and 62004 stroke-free controls of European ancestry from 15 case-control studies. The novel associations identified, were further replicated in the same study in an independent set of 13347 cases and 29083 controls from 18 independent case-control studies. This study identified loci near the genes coding the paired-like homeodomain transcription factor 2 (PITX2) and the Zinc finger homeobox protein 3 (ZFHX3) to be associated with cardioembolic stroke, and loci in 9p21 and histone deacetylase 9 (HDAC9) with large-vessel stroke (two of the most common IS subtypes) (91).

Another recent meta-analysis, including 2247 IS cases and 1813 controls of Chinese and Caucasian origin, from 10 case-control studies, pointed to the T833C genetic polymorphism of the Cystathionine β Synthase (CBS) gene as a risk factor for ischemic stroke (92). Finally, another meta-analysis including 4681 ischemic stroke cases and 8516 controls of Caucasian and Asian origin from 3 case-control studies suggested the Ser447Ter and Asn291Ser polymorphisms of the lipoprotein lipase gene to be associated with risk of ischemic stroke (93).

Table 1. Chromosomal regions associated with MI and identified by GWA studies and successfully replicated (taken from reference (86))

Chromosomic region	SNP	Risk allele Freq.	OR (95% CI)	Nearby Genes	Gene function
1p13.3	rs599839	77%	1.13 (1.08–1.19)	*PSCR1, CELSR2, SORT1, MYBPHL*	LDL increase
1q41	rs3008621	72%	1.10 (1.04–1.17)	*MIA3*	Collagen processing
2q33	rs6725887	14%	1.17 (1.11–1.23)	*WDR12*	Apoptosis
3q22.3	rs9818870	15%	1.15 (1.11–1.19)	*MRAS*	Adhesion signaling
6p24	rs12526453	65%	1.13 (1.08–1.17)	*PHACTR1*	Coronary calcification
6q26–27	rs2048327	18%	1.20 (1.13–1.28)	*SLC22A3, LPAL2, LPA*	Lp(a)
9p21.3	rs3127599	52%	1.36 (1.27–1.46)	*MTAP, CDKN2A, CDKN2B, ANRIL*	Unknown
10q11	rs7767084	84%	1.11 (1.05–1.18)	*SDF1*	EPC recruiting and inflammation
12q24	rs10755578	34%	1.14 (1.10–1.19)	*SH2B3*	Unknown
12q24.3	rs1333049	36%	1.08 (1.05–1.11)	*HNF1A, C12orf43*	Unknown
21q22	rs501120	13%	1.19 (1.14–1.27)	*SLC5A3, MRPS6, KCNE2*	Unknown

1.5 AIM AND OUTLINE OF THIS THESIS

This thesis presents four genetic association studies (Chapters 2-5) which aim to add further understanding of the role of selected SNPs on intermediate and final phenotypes of CVD. Also a systematic review on available methods for the automation of food questionnaires is presented (Chapter 6).

Chapter 2 presents a genetic association study based on a cohort of 940 persons at high cardiovascular risk from the Spanish Mediterranean area (PREDIMED-Valencia Study). In this study the candidate-gene approach was used to investigate whether the gene variant rs4988235, located nearby the lactase (*LCT*) gene, and strongly associated with lactase persistence in Europeans, was associated with obesity-related variables. Its potential modulation by consumption of dairy products was also examined.

Chapters 3-5 present three separate genetic association studies, conducted in the Potsdam (Germany) arm of the European Prospective Investigation into Cancer and Nutrition (EPIC-Potsdam), a prospective cohort study with baseline measurements on 27,548 participants between 1994-1998.

Chapter 3 reports on the association of the gene variant rs2943634, which is located in a non-coding region of chromosome 2q36.3 and recently has been associated with coronary artery disease in two GWA studies, with incident MI and IS. This association was assesed using a case-cohort design, including a representative subsample of the total EPIC-Potsdam cohort (n = 2500) and all incident cases of MI (n = 211) and IS (n = 144) that occurred during a mean follow-up period of 8.2 year. Furthermore, the association of this SNP with 12 available potential intermediate risk phenotypes of CVD was investigated in the representative subsample of 2500 participants.

The study on **Chapter 4** used the gene-wide approach to examine the impact of common genetic variation in the Stearoyl-CoA desaturase-1 (*SCD1*) gene, captured by means of 7 tag-SNPs and 5 inferred haplotypes, on the modulation of 8 metabolic risk factors linked to the activity of SCD1. SCD1 is an enzyme involved in lipid metabolism. Its activity has been associated with traits of the metabolic syndrome in mice and humans, but also with the prevention of saturated fatty acids accumulation and subsequent inflammation, whereas for liver fat content inconsistent results have been reported. This study was conducted in the sample of 2500 EPIC-Potsdam participants.

The study presented on **Chapter 5** of this thesis, reports on the differential association of the gene variant -164T>C with CVD, according to different cholesterol cutpoints. This polymorphism is located in the promoter region of the gene encoding for the microsomal triglyceride transfer protein (MTTP). MTTP plays an important role in the assembly and secretion of apolipoprotein-B containing lipoproteins as chylomicrons in the intestine, and of very low density lipoproteins in the liver. *MTTP* is regulated by cholesterol. The study was conducted following a case-cohort design within the EPIC-Potsdam cohort (193 incident myocardial infarction cases , 131 incident ischemic stroke cases and 1978 non-cases). Further, the Heinz Nixdorf Recall Study (30 CVD cases and 1,188 controls) was used to replicate findings. While the PhD candidate did not lead this study, she actively participated on it by

performing the genotyping work and providing critical revision of the article for important intellectual content.

In **Chapter 6** of this thesis, a state-of-the-art review of software applications for automating the most commonly used dietary surveys in nutritional research, that is FFQs and 24hDR, is presented.

Although results and conclusions from each study are discussed in the separate chapters, **Chapter 7** compiles the main findings of the work presented, and discusses some methodological issues. **Chapter 8** enumerates the mean conclusions derived in this thesis.

During the predoctoral training period, besides the studies presented in the compendium of publications that conforms this thesis, the PhD candidate had the opportunity to conduct additional genetic association studies and also to work in the field of biomedical informatics applied to nutrigenetic research. As a result, she obteined the "*Diploma de Estudios Avanzados*" degree, granted by the Universitat de València, Spain. Additionally she wrote several comunications for Spanish and international congresses. Although it is beyond the scope of this PhD thesis to provide a detailed description of this additional work, **Chapter 9** summarizes its main objectives, and the publicacions that derived from it.

Finally it is worth mentioning that during the development of this thesis, the PhD candidate had the opportunity to complete her education, by actively collaborating in field work tasks derived from epidemiological studies running at the University of Valencia, such as recruitment of participants, administration and processing of questionnaires (dietary, socio-demographic, life-style), collaboration in interventional studies, DNA extraction, genotyping, etc.

Refferences

1. Dzau VJ, Antman EM, Black HR, Hayes DL, Manson JE, Plutzky J, et al. The cardiovascular disease continuum validated: clinical evidence of improved patient outcomes: part I: Pathophysiology and clinical trial evidence (risk factors through stable coronary artery disease). Circulation. 2006;114(25):2850-70. Epub 2006/12/21.

2. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2013;380(9859):2095-128. Epub 2012/12/19.

3. Murray CJ, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, et al. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2197-223. Epub 2012/12/19.

4. ICD10. http://apps.who.int/classifications/icd10/browse/2010/en.

5. Ross R. The pathogenesis of atherosclerosis: a perspective for the 1990s. Nature. 1993;362(6423):801-9. Epub 1993/04/29.

6. Yusuf S, Hawken S, Ounpuu S, Dans T, Avezum A, Lanas F, et al. Effect of potentially modifiable risk factors associated with myocardial infarction in 52 countries (the INTERHEART study): case-control study. Lancet. 2004;364(9438):937-52. Epub 2004/09/15.

7. O'Donnell MJ, Xavier D, Liu L, Zhang H, Chin SL, Rao-Melacini P, et al. Risk factors for ischaemic and intracerebral haemorrhagic stroke in 22 countries (the INTERSTROKE study): a case-control study. Lancet. 2010;376(9735):112-23. Epub 2010/06/22.

8. Faxon DP, Fuster V, Libby P, Beckman JA, Hiatt WR, Thompson RW, et al. Atherosclerotic Vascular Disease Conference: Writing Group III: pathophysiology. Circulation. 2004;109(21):2617-25.

9. Wilson K, Gibson N, Willan A, Cook D. Effect of smoking cessation on mortality after myocardial infarction: meta-analysis of cohort studies. Archives of Internal Medicine. 2000;160(7):939-44. Epub 2000/04/13.

10. Taylor BV, Oudit GY, Kalman PG, Liu P. Clinical and pathophysiological effects of active and passive smoking on the cardiovascular system. The Canadian journal of cardiology. 1998;14(9):1129-39. Epub 1998/10/21.

11. Rossi M, Negri E, La Vecchia C, Campos H. Smoking habits and the risk of non-fatal acute myocardial infarction in Costa Rica. European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology. 2011;18(3):467-74. Epub 2011/04/01.

12. Doll R, Peto R, Boreham J, Sutherland I. Mortality in relation to smoking: 50 years' observations on male British doctors. Bmj. 2004;328(7455):1519. Epub 2004/06/24.

13. Haheim LL, Holme I, Hjermann I, Leren P. Smoking habits and risk of fatal stroke: 18 years follow up of the Oslo Study. Journal of epidemiology and community health. 1996;50(6):621-4. Epub 1996/12/01.

14. Robbins AS, Manson JE, Lee IM, Satterfield S, Hennekens CH. Cigarette smoking and stroke in a cohort of U.S. male physicians. Annals of internal medicine. 1994;120(6):458-62. Epub 1994/03/15.

15. Qin R, Chen T, Lou Q, Yu D. Excess risk of mortality and cardiovascular events associated with smoking among patients with diabetes: Meta-analysis of observational prospective studies. International journal of cardiology. 2012. Epub 2012/01/19.

16. Chirovsky DR, Fedirko V, Cui Y, Sazonov V, Barter P. Prospective studies on the relationship between high-density lipoprotein cholesterol and cardiovascular risk: a systematic review. European journal of cardiovascular prevention and rehabilitation : official journal of the European Society of Cardiology, Working Groups on Epidemiology & Prevention and Cardiac Rehabilitation and Exercise Physiology. 2009;16(4):404-23. Epub 2009/05/26.

17. Sarwar N, Danesh J, Eiriksdottir G, Sigurdsson G, Wareham N, Bingham S, et al. Triglycerides and the risk of coronary heart disease: 10,158 incident cases among 262,525 participants in 29 Western prospective studies. Circulation. 2007;115(4):450-8. Epub 2006/12/28.

18. Ip S, Lichtenstein AH, Chung M, Lau J, Balk EM. Systematic review: association of low-density lipoprotein subfractions with cardiovascular outcomes. Annals of internal medicine. 2009;150(7):474-84. Epub 2009/04/08.

19. Patel A, Barzi F, Jamrozik K, Lam TH, Ueshima H, Whitlock G, et al. Serum triglycerides as a risk factor for cardiovascular diseases in the Asia-Pacific region. Circulation. 2004;110(17):2678-86. Epub 2004/10/20.

20. Labreuche J, Touboul PJ, Amarenco P. Plasma triglyceride levels and risk of stroke and carotid atherosclerosis: a systematic review of the epidemiological studies. Atherosclerosis. 2009;203(2):331-45. Epub 2008/10/29.

21. Amarenco P, Labreuche J, Touboul PJ. High-density lipoprotein-cholesterol and risk of stroke and carotid atherosclerosis: a systematic review. Atherosclerosis. 2008;196(2):489-96. Epub 2007/10/10.

22. Yatsuya H, Toyoshima H, Yamagishi K, Tamakoshi K, Taguri M, Harada A, et al. Body mass index and risk of stroke and myocardial infarction in a relatively lean population: meta-analysis of 16 Japanese cohorts using individual data. Circulation Cardiovascular quality and outcomes. 2010;3(5):498-505. Epub 2010/08/12.

23. Strazzullo P, D'Elia L, Cairella G, Garbagnati F, Cappuccio FP, Scalfi L. Excess body weight and incidence of stroke: meta-analysis of prospective studies with 2 million participants. Stroke; a journal of cerebral circulation. 2010;41(5):e418-26. Epub 2010/03/20.

24. Guh DP, Zhang W, Bansback N, Amarsi Z, Birmingham CL, Anis AH. The incidence of co-morbidities related to obesity and overweight: a systematic review and meta-analysis. BMC public health. 2009;9:88. Epub 2009/03/27.

25. Moss SE, Klein R, Klein BE, Meuer SM. The association of glycemia and cause-specific mortality in a diabetic population. Archives of internal medicine. 1994;154(21):2473-9. Epub 1994/11/14.

26. Moss SE, Klein R, Klein BE. Cause-specific mortality in a population-based study of diabetes. American journal of public health. 1991;81(9):1158-62. Epub 1991/09/01.

27. Dorman JS, Laporte RE, Kuller LH, Cruickshanks KJ, Orchard TJ, Wagener DK, et al. The Pittsburgh insulin-dependent diabetes mellitus (IDDM) morbidity and mortality study. Mortality results. Diabetes. 1984;33(3):271-6. Epub 1984/03/01.

28. Fox CS, Coady S, Sorlie PD, D'Agostino RB, Sr., Pencina MJ, Vasan RS, et al. Increasing cardiovascular disease burden due to diabetes mellitus: the Framingham Heart Study. Circulation. 2007;115(12):1544-50. Epub 2007/03/14.

29. Selvin E, Marinopoulos S, Berkenblit G, Rami T, Brancati FL, Powe NR, et al. Meta-analysis: glycosylated hemoglobin and cardiovascular disease in diabetes mellitus. Annals of internal medicine. 2004;141(6):421-31. Epub 2004/09/24.

30. Quinn TJ, Dawson J, Walters MR. Sugar and stroke: cerebrovascular disease and blood glucose control. Cardiovascular therapeutics. 2011;29(6):e31-42. Epub 2010/05/25.

31. Lee M, Saver JL, Hong KS, Song S, Chang KH, Ovbiagele B. Effect of pre-diabetes on future risk of stroke: meta-analysis. Bmj. 2012;344:e3564. Epub 2012/06/09.

32. Fagard RH, Celis H, Thijs L, Staessen JA, Clement DL, De Buyzere ML, et al. Daytime and nighttime blood pressure as predictors of death and cause-specific cardiovascular events in hypertension. Hypertension. 2008;51(1):55-61. Epub 2007/11/28.

33. Pearce KA, Furberg CD, Rushing J. Does antihypertensive treatment of the elderly prevent cardiovascular events or prolong life? A meta-analysis of hypertension treatment trials. Archives of family medicine. 1995;4(11):943-9; discussion 50. Epub 1995/11/01.

34. Kraja AT, Hunt SC, Rao DC, Davila-Roman VG, Arnett DK, Province MA. Genetics of hypertension and cardiovascular disease and their interconnected pathways: lessons from large studies. Current hypertension reports. 2011;13(1):46-54. Epub 2010/12/04.

35. Lawes CM, Bennett DA, Feigin VL, Rodgers A. Blood pressure and stroke: an overview of published reviews. Stroke; a journal of cerebral circulation. 2004;35(4):1024. Epub 2004/04/01.

36. Li J, Siegrist J. Physical activity and risk of cardiovascular disease--a meta-analysis of prospective cohort studies. International journal of environmental research and public health. 2012;9(2):391-407. Epub 2012/04/04.

37. Reimers CD, Knapp G, Reimers AK. Exercise as stroke prophylaxis. Deutsches Arzteblatt international. 2009;106(44):715-21. Epub 2009/12/10.

38. Hamer M, Malan L. Psychophysiological risk markers of cardiovascular disease. Neuroscience and biobehavioral reviews. 2010;35(1):76-83. Epub 2009/11/17.

39. Roest AM, Martens EJ, de Jonge P, Denollet J. Anxiety and risk of incident coronary heart disease: a meta-analysis. Journal of the American College of Cardiology. 2010;56(1):38-46. Epub 2010/07/14.

40. Serrano CV, Jr., Setani KT, Sakamoto E, Andrei AM, Fraguas R. Association between depression and development of coronary artery disease: pathophysiologic and diagnostic implications. Vascular health and risk management. 2011;7:159-64. Epub 2011/04/15.

41. Dong JY, Zhang YH, Tong J, Qin LQ. Depression and risk of stroke: a meta-analysis of prospective studies. Stroke; a journal of cerebral circulation. 2012;43(1):32-7. Epub 2011/10/25.

42. Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. The Journal of nutrition. 2006;136(10):2588-93. Epub 2006/09/22.

43. He FJ, Nowson CA, Lucas M, MacGregor GA. Increased consumption of fruit and vegetables is related to a reduced risk of coronary heart disease: meta-analysis of cohort studies. Journal of human hypertension. 2007;21(9):717-28. Epub 2007/04/20.

44. Dauchet L, Amouyel P, Dallongeville J. Fruit and vegetable consumption and risk of stroke: a meta-analysis of cohort studies. Neurology. 2005;65(8):1193-7. Epub 2005/10/26.

45. Ronksley PE, Brien SE, Turner BJ, Mukamal KJ, Ghali WA. Association of alcohol consumption with selected cardiovascular disease outcomes: a systematic review and meta-analysis. Bmj. 2011;342:d671. Epub 2011/02/24.

46. Liu PM, Dosieah S, Zheng HS, Huang ZB, Lin YQ, Wang JF. [Alcohol consumption and coronary heart disease in Eastern Asian men: a meta-analysis of prospective cohort studies]. Zhonghua xin xue guan bing za zhi. 2010;38(11):1038-44. Epub 2011/01/11.

47. Liu PM, Dosieah S, Luo NS, Huang ZB, Lin YQ, Wang JF. [Alcohol intake and stroke in Eastern Asian men:a systemic review and meta-analysis of 17 prospective cohort studies]. Zhonghua yi xue za zhi. 2010;90(40):2834-8. Epub 2010/12/18.

48. Patra J, Taylor B, Irving H, Roerecke M, Baliunas D, Mohapatra S, et al. Alcohol consumption and the risk of morbidity and mortality for different stroke types--a systematic review and meta-analysis. BMC public health. 2010;10:258. Epub 2010/05/21.

49. Elkind MS, Sciacca R, Boden-Albala B, Rundek T, Paik MC, Sacco RL. Moderate alcohol consumption reduces risk of ischemic stroke: the Northern Manhattan Study. Stroke; a journal of cerebral circulation. 2006;37(1):13-9. Epub 2005/11/25.

50. Jimenez M, Chiuve SE, Glynn RJ, Stampfer MJ, Camargo CA, Jr., Willett WC, et al. Alcohol consumption and risk of stroke in women. Stroke; a journal of cerebral circulation. 2012;43(4):939-45. Epub 2012/03/10.

51. Johnsen SP. Intake of fruit and vegetables and risk of stroke: an overview. Current opinion in clinical nutrition and metabolic care. 2004;7(6):665-70. Epub 2004/11/10.

52. de Lorgeril M, Salen P, Martin JL, Monjaud I, Delaye J, Mamelle N. Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. Circulation. 1999;99(6):779-85. Epub 1999/02/17.

53. Talbert RL. Role of antihypertensive therapy with angiotensin-converting enzyme inhibitors or angiotensin II receptor blockers in combination with calcium channel blockers for stroke prevention. Journal of the American Pharmacists Association : JAPhA. 2010;50(5):e116-25. Epub 2010/09/14.

54. Turnbull F, Blood Pressure Lowering Treatment Trialists C. Effects of different blood-pressure-lowering regimens on major cardiovascular events: results of prospectively-designed overviews of randomised trials. Lancet. 2003;362(9395):1527-35. Epub 2003/11/15.

55. Musunuru K. Atherogenic dyslipidemia: cardiovascular risk and dietary intervention. Lipids. 2010;45(10):907-14. Epub 2010/06/05.

56. Martin SS, Blumenthal RS, Miller M. LDL cholesterol: the lower the better. The Medical clinics of North America. 2012;96(1):13-26. Epub 2012/03/07.

57. Collins R, Armitage J, Parish S, Sleigh P, Peto R, Heart Protection Study Collaborative G. MRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial. Lancet. 2003;361(9374):2005-16. Epub 2003/06/20.

58. Amarenco P, Labreuche J. Lipid management in the prevention of stroke: review and updated meta-analysis of statins for stroke prevention. Lancet neurology. 2009;8(5):453-63. Epub 2009/04/21.

59. Cholesterol Treatment Trialists C, Kearney PM, Blackwell L, Collins R, Keech A, Simes J, et al. Efficacy of cholesterol-lowering therapy in 18,686 people with diabetes in 14 randomised trials of statins: a meta-analysis. Lancet. 2008;371(9607):117-25. Epub 2008/01/15.

60. Baigent C, Keech A, Kearney PM, Blackwell L, Buck G, Pollicino C, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056

participants in 14 randomised trials of statins. Lancet. 2005;366(9493):1267-78. Epub 2005/10/11.

61. Brien SE, Ronksley PE, Turner BJ, Mukamal KJ, Ghali WA. Effect of alcohol consumption on biological markers associated with risk of coronary heart disease: systematic review and meta-analysis of interventional studies. Bmj. 2011;342:d636. Epub 2011/02/24.

62. Potpara TS, Lip GY, Apostolakis S. New anticoagulant treatments to protect against stroke in atrial fibrillation. Heart. 2012;98(18):1341-7. Epub 2012/06/26.

63. Simmons RK, Alberti KG, Gale EA, Colagiuri S, Tuomilehto J, Qiao Q, et al. The metabolic syndrome: useful concept or clinical tool? Report of a WHO Expert Consultation. Diabetologia. 2010;53(4):600-5.

64. Mottillo S, Filion KB, Genest J, Joseph L, Pilote L, Poirier P, et al. The metabolic syndrome and cardiovascular risk a systematic review and meta-analysis. Journal of the American College of Cardiology. 2010;56(14):1113-32. Epub 2010/09/25.

65. Li W, Ma D, Liu M, Liu H, Feng S, Hao Z, et al. Association between metabolic syndrome and risk of stroke: a meta-analysis of cohort studies. Cerebrovascular diseases. 2008;25(6):539-47. Epub 2008/05/16.

66. Reaven GM. The metabolic syndrome: time to get off the merry-go-round? J Intern Med. 2011;269(2):127-36.

67. Wang JC, Bennett M. Aging and atherosclerosis: mechanisms, functional consequences, and potential therapeutics for cellular senescence. Circulation research. 2012;111(2):245-59. Epub 2012/07/10.

68. Godsland IF, North BV, Johnston DG. Simple indices of inflammation as predictors of death from cancer or cardiovascular disease in a prospective cohort after two decades of follow-up. QJM : monthly journal of the Association of Physicians. 2011;104(5):387-94. Epub 2010/11/26.

69. Emerging Risk Factors C, Kaptoge S, Di Angelantonio E, Lowe G, Pepys MB, Thompson SG, et al. C-reactive protein concentration and risk of coronary heart disease, stroke, and mortality: an individual participant meta-analysis. Lancet. 2010;375(9709):132-40. Epub 2009/12/25.

70. Gottlieb DJ, Redline S, Nieto FJ, Baldwin CM, Newman AB, Resnick HE, et al. Association of usual sleep duration with hypertension: the Sleep Heart Health Study. Sleep. 2006;29(8):1009-14. Epub 2006/09/02.

71. Gangwisch JE, Heymsfield SB, Boden-Albala B, Buijs RM, Kreier F, Pickering TG, et al. Short sleep duration as a risk factor for hypertension: analyses of the first National Health and Nutrition Examination Survey. Hypertension. 2006;47(5):833-9. Epub 2006/04/06.

72. Wolk R, Somers VK. Sleep and the metabolic syndrome. Experimental physiology. 2007;92(1):67-78. Epub 2006/11/07.

73. Cappuccio FP, Cooper D, D'Elia L, Strazzullo P, Miller MA. Sleep duration predicts cardiovascular outcomes: a systematic review and meta-analysis of prospective studies. European heart journal. 2011;32(12):1484-92. Epub 2011/02/09.

74. Targher G, Day CP, Bonora E. Risk of cardiovascular disease in patients with nonalcoholic fatty liver disease. The New England journal of medicine. 2010;363(14):1341-50. Epub 2010/10/01.

75. Weiner DE, Tighiouart H, Amin MG, Stark PC, MacLeod B, Griffith JL, et al. Chronic kidney disease as a risk factor for cardiovascular disease and all-cause mortality: a pooled analysis of community-based studies. Journal of the American Society of Nephrology : JASN. 2004;15(5):1307-15. Epub 2004/04/22.

76. Lo MM, Mitsnefes M. Adiponectin, cardiovascular disease, chronic kidney disease: emerging data on complex interactions. Pediatric nephrology. 2012;27(4):521-7. Epub 2011/02/22.

77. Kato H, Kashiwagi H, Shiraga M, Tadokoro S, Kamae T, Ujiie H, et al. Adiponectin acts as an endogenous antithrombotic factor. Arteriosclerosis, thrombosis, and vascular biology. 2006;26(1):224-30. Epub 2005/11/05.

78. Willett WC, Sacks F, Trichopoulou A, Drescher G, Ferro-Luzzi A, Helsing E, et al. Mediterranean diet pyramid: a cultural model for healthy eating. The American journal of clinical nutrition. 1995;61(6 Suppl):1402S-6S. Epub 1995/06/01.

79. Sofi F, Abbate R, Gensini GF, Casini A. Accruing evidence on benefits of adherence to the Mediterranean diet on health: an updated systematic review and meta-analysis. The American journal of clinical nutrition. 2010;92(5):1189-96. Epub 2010/09/03.

80. Gianfagna F, Cugino D, Santimone I, Iacoviello L. From candidate gene to genome-wide association studies in cardiovascular disease. Thrombosis research. 2012;129(3):320-4. Epub 2011/12/14.

81. Chow CK, Islam S, Bautista L, Rumboldt Z, Yusufali A, Xie C, et al. Parental history and myocardial infarction risk across the world: the INTERHEART Study. Journal of the American College of Cardiology. 2011;57(5):619-27. Epub 2011/01/29.

82. Knottnerus IL, Gielen M, Lodder J, Rouhl RP, Staals J, Vlietinck R, et al. Family history of stroke is an independent risk factor for lacunar stroke subtype with asymptomatic lacunar infarcts at younger ages. Stroke; a journal of cerebral circulation. 2011;42(5):1196-200. Epub 2011/03/29.

83. Mvundura M, McGruder H, Khoury MJ, Valdez R, Yoon PW. Family History as a Risk Factor for Early-Onset Stroke/Transient Ischemic Attack among Adults in the United States. Public health genomics. 2010;13(1):13-20. Epub 2009/03/25.

84. Kelly M, Semsarian C. Multiple mutations in genetic cardiovascular disease: a marker of disease severity? Circulation Cardiovascular genetics. 2009;2(2):182-90. Epub 2009/12/25.

85. Banerjee A. A review of family history of cardiovascular disease: risk factor and research tool. International journal of clinical practice. 2012;66(6):536-43. Epub 2012/05/23.

86. Erdmann J, Linsel-Nitschke P, Schunkert H. Genetic causes of myocardial infarction: new insights from genome-wide association studies. Deutsches Arzteblatt international. 2010;107(40):694-9. Epub 2010/10/30.

87. Lusis AJ. Genetics of atherosclerosis. Trends in genetics : TIG. 2012;28(6):267-75. Epub 2012/04/07.

88. Altshuler D, Daly MJ, Lander ES. Genetic mapping in human disease. Science. 2008;322(5903):881-8. Epub 2008/11/08.

89. Xuan C, Bai XY, Gao G, Yang Q, He GW. Association between polymorphism of methylenetetrahydrofolate reductase (MTHFR) C677T and risk of myocardial infarction: a meta-analysis for 8,140 cases and 10,522 controls. Archives of medical research. 2011;42(8):677-85. Epub 2011/12/14.

90. Wang J, Xu D, Wu X, Zhou C, Wang H, Guo Y, et al. Polymorphisms of matrix metalloproteinases in myocardial infarction: a meta-analysis. Heart. 2011;97(19):1542-6. Epub 2011/09/09.

91. Traylor M, Farrall M, Holliday EG, Sudlow C, Hopewell JC, Cheng YC, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE Collaboration): a meta-analysis of genome-wide association studies. Lancet neurology. 2012;11(11):951-62. Epub 2012/10/09.

92. Ding R, Lin S, Chen D. The association of cystathionine beta synthase (CBS) T833C polymorphism and the risk of stroke: a meta-analysis. Journal of the neurological sciences. 2012;312(1-2):26-30. Epub 2011/09/16.

93. Wang C, Sun T, Li H, Bai J, Li Y. Lipoprotein lipase Ser447Ter polymorphism associated with the risk of ischemic stroke: a meta-analysis. Thrombosis research. 2011;128(5):e107-12. Epub 2011/08/06.

2

Association of the rs4988235 at -13910 bp upstream from the lactase gene with obesity and its modulation by dairy product consumption in an elderly Mediterranean population

Chapter 2

Association of the rs4988235 at -13910 bp upstream from the lactase gene with obesity and its modulation by dairy product consumption in an elderly Mediterranean population

Corella D[1,2,3*], **Arregui M[1,2,4*]**, Coltell O[3,4], Portolés O[1,2], Guillem-Sáiz P[1,2], Carrasco P[1,2], Sorlí JV[1,2,5], Ortega-Azorín C[1,2], González JI[1,2], Ordovás JM[2,3,6]

*: These authors contributed equally to this work.

1. Genetic and Molecular Epidemiology Unit School of Medicine. University of Valencia, Valencia, Spain.

2. CIBER Fisiopatología de la Obesidad y Nutrición, Instituto de Salud Carlos III, Madrid, Spain.

3. Nutrition and Genomics Laboratory, JM-USDA Human Nutrition Research Center on Aging at Tufts University, Boston, MA.

4. Department of Computing Languages and Systems. University Jaume I, Castellón, Spain.

5. Primary Care Health Center. Generalitat Valenciana. Xirivella, Valencia, Spain

6. Department of Cardiovascular Epidemiology and Population Genetics, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain.

Obesity (Silver Spring). 2011 Aug;19(8):1707-14

Impact factor 2011: 4.284

ABSTRACT

The single nucleotide polymorphism (SNP) upstream (-13910C>T; rs4988235) from the lactase (*LCT*) gene, strongly associated with lactase persistence (LP) in Europeans, is emerging as new candidate for obesity. Our aim was to analyze the association of the rs4988235 with obesity-related variables as well as its modulation by dairy product intake in an elderly population. We studied 940 white subjects (aged 67+/-7 years) from the Spanish Mediterranean population. Dairy product consumption was assessed by a validated questionnaire. Anthropometric variables were directly measured, and metabolic-syndrome (MetS) related variables were obtained. Prevalence of the CC genotype [lactase non-persistent (LNP)] was high (38.0%). CT (LP) and TT (LP) were 45.7% and 16.3%, respectively. The CC genotype was not associated with lower milk or dairy product consumption in the population as a whole. Only in women was dairy product consumption significantly lower in CC. The most important association was obtained with anthropometric measurements. CC individuals presented lower weight (P=0.032), lower BMI (29.7+/-4.2 vs. 30.6+/-4.2 Kg/m2; P=0.003) and lower waist circumference (101.1+/-11.8 vs. 103.5+/-11.5 cm; P=0.005) than T-allele carriers. Obesity risk was also significantly higher in T-allele carriers than in CC individuals (OR: 1.38; 95% CI: 1.05-1.81; P=0.01) and remained significant after adjustment for sex, age, diabetes, physical activity and energy. Dairy lactose intake modulated these associations, being higher with higher lactose intake. No significant associations with lipids, blood pressure or MetS were obtained. In conclusion, despite not finding marked differences in dairy product consumption, the CC genotype was strongly associated with obesity in this Mediterranean population.

Key words: Lactase, obesity, gene, dairy products, metabolic syndrome, Mediterranean

INTRODUCTION

The association of dairy food consumption with obesity and other cardiovascular risk factors has been investigated in several studies, but with contradictory results (1-6). A beneficial effect of dairy consumption on the incidence of various metabolic syndrome components (including obesity, glucose intolerance, hypertension, and dyslipidemia) was reported by Pereira et al (1) in the CARDIA study and replicated in some (2-4), but not all (5,6) subsequent studies. Some meta-analyses carried out for this purpose reflect the inconsistency of results and underline the need to analyse the different factors involved in greater depth (7-9). One of the potential factors that may affect the quantity of milk consumed as well as the effects of dairy products on obesity and obesity-related variables in adults is lactose intolerance or lactase non-persistence (LNP). Lactose intolerance is the syndrome of diarrhea, abdominal pain or flatulence, occurring after lactose ingestion (10). These symptoms, caused by a decreased ability to hydrolyze lactose due to a deficiency in the enzyme lactase, may have an influence in the amount of dairy product consumed. On the other hand, if there is no restriction of dairy products in LNP subjects, the undigested lactose may have several metabolic effects that may be related to obesity.

Lactase is coded by the lactase gene (*LCT*), and LCT activity remains high until weaning, then it fades away in most of the adult population (adult-type hypolactasia or LNP). A single nucleotide polymorphism (SNP) (rs4988235), located at -13910 bp upstream from the *LCT* gene (-13910C > T) within intron 13 of the adjacent minichromosome maintenance 6 (MCM6) gene was found to be associated with LNP (11). Various studies (11-13) have demonstrated that the -13910C > T SNP is functional and is associated with changes in LCT gene expression. Individuals homozygous for the C allele (LNP) have almost undetectable levels of intestinal lactase production compared to TC or TT individuals [lactase persistent (LP), following a codominant model] (11). Pohl et al (14) found an excellent agreement between the lactose hydrogen test (10) and the genetic test based on this SNP for LNP in Europeans. The frequency of LP is high in northern European populations, decreases across southern Europe and more than half of the world's population is LNP (15). Although some studies have associated the CC genotype with a lower consumption of milk (16-18), this association is not always observed (19-21).

Interestingly, the *LCT* gene is emerging as a new candidate gene related with obesity and other anthropometric measurements. Hence, in a recent Genome-Wide Association Study (GWAs) carried out by CHARGE (Cohorts for Heart and Aging Research in Genome Epidemiology), the Consortium (22) found a strong association between various SNPs in the *LCT* gene and waist circumference. Likewise, Kettunen et al (23), undertook a meta-analysis on eight European cohorts (in which the Mediterranean population was not included) finding a strong association between the LP variant (T allele, rs4988235) and higher body mass index (BMI). However, none of the published studies has analyzed the joint influence of the LP genotype and dairy product consumption on obesity. Therefore, our aim was to study the association of the -13910C > T SNP with obesity and obesity-related variables as well as its modulation by lactose intake in an elderly Mediterranean population.

MATERIAL AND METHODS

Subjects and study design

We included 940 unrelated White individuals (338 men and 602 women), mean age 67.3+/-6.5 years, who participated in the PREDIMED (Prevención con Dieta MEDiterránea) study, were consecutively recruited in the Valencia Region (on the East Mediterranean coast of Spain) from October 2003 to December 2008 and had the lactase genotype determined. All participants gave their informed consent. The ethics committee of the University of Valencia approved the study. Details of this study have been previously reported (24). Briefly, high cardiovascular risk subjects were selected by physicians in Primary Care Centers participating in the study. Eligible subjects were community-dwelling people (55-80 years of age for men; 60-80 years of age for women) who fulfilled at least one of two criteria: type 2 diabetes; 3 or more cardiovascular risk factors (current smoking, hypertension, dyslipidemia, overweight, or a family history of premature cardiovascular disease).

Demographic, anthropometric and clinical measurements

The baseline examination included assessment of standard socio-demographic factors, clinical, biochemical and lifestyle variables, as previously detailed (24). Anthropometric variables were directly measured by trained nurses by standard techniques at baseline (24). Height and weight

were measured with light clothing and no shoes. Obesity was defined as BMI>=30kg/m2. Waist circumference was measured midway between the lowest rib and the iliac crest using an anthropometric tape. Trained personnel measured blood pressure with a validated semi-automatic sphygmomanometer (Omron HEM-705CP, The Netherlands) in a seated position after a 5-min rest. Physical activity was estimated by the Minnesota Leisure Time Physical Activity as previously reported (24). Blood samples were obtained for each participant after an overnight fast and were frozen at -80ºC. Fasting glucose, total cholesterol, triglycerides, HDL-C and LDL-C were determined as previously reported (24). The metabolic syndrome was defined according to updated ATP III criteria (25), which require that 3 or more of the following conditions be met: abdominal obesity (waist circumference >102 cm in men and >88 cm in women), hypertriglyceridemia (triglycerides level 150 mg/dL), low HDL cholesterol level (<40 mg/dL in men and <50 mg/dL in women), elevated fasting blood glucose level (100 mg/dL), and elevated blood pressure (systolic 130 mm Hg, diastolic 85 mm Hg, or taking antihypertensive medication). Participants who were being treated with antidiabetic, antihypertensive, or triglyceride-lowering medications were considered to be diabetic, hypertensive, or hypertriglyceridemic, respectively.

Dietary measurements

Food consumption was determined by a validated (26, 27) semi-quantitative 137-item food frequency questionnaire (FFQ). Energy and nutrient intake were calculated from Spanish food composition tables (28). Lactose content was not available in the Spanish tables and so foreign food composition tables were used (Fineli Food Composition Database, Finland, release 2010). The questionnaire was based on the typical portion sizes that were multiplied by the consumption frequency for each food. Information about dairy products was assessed in fifteen items of the semi-quantitative FFQ (whole-fat milk, partially skimmed milk, skimmed milk, condensed milk, whipped cream, yoghurt, skimmed yoghurt, milkshake, ricotta cheese or junket, petit Suisse cheese, spreadable cheese wedges, cottage cheese, other cheese, custard, and ice cream). We calculated total dairy products intake (in g/d) for each individual on the basis of the type and amount consumed.

DNA extraction and genotyping

Genomic DNA was isolated from blood. The rs4988235 *LCT* SNP was determined using a 7900HT Sequence Detection system (Applied Biosystems) and a customized fluorescent allelic discrimination TaqMan assay by standard procedures. For quality control purposes, 50% of randomly selected samples were also genotyped by restriction fragment length polymorphism analysis. Concordance between techniques was higher than 95%. Discrepant samples were sequenced. Information on probes and polymerase chain reaction conditions for genotyped single nucleotide polymorphisms can be obtained from the authors upon request.

Statistical analysis

X^2 tests were used to test differences in percentages. Taking into account that the genetically defined LP is considered to follow a dominant model, CT and TT subjects (LP) were grouped and compared with CC subjects for the statistical analysis after having checked that this dominant model is observed in this Mediterranean population. We applied the t tests to compare crude means for normally distributed variables. Alcohol and dairy product consumption did not follow a normal distribution and we applied the non-parametric Mann-Whitney U test. For continuous anthropometric variables, multivariate adjustment was carried out by linear regression analysis. Model were adjusted for sex, age (as continuous), diabetes, total energy intake (as continuous) and physical activity (as continuous). Additional adjustment for dairy product consumption was also done. Multivariate adjustment of plasma lipids, fasting glucose and blood pressure was also carried out by linear regression. Regression coefficients and adjusted means for each predictor were estimated from the multivariate models. Regression models with interaction terms and as well as stratified analysis were applied to test the homogeneity of effects by gender and lactose intake. Logistic regression models were fitted to estimate the odds ratio (OR) and 95% confidence interval (CI) of obesity and obesity-related variables associated with the LP genotype compared with LNP. Multiple logistic regression models were also fitted to check for the effect of covariates and effect modifiers as well as to test the interaction between the SNP and selected variables. Analyses were performed using the SPSS statistical software, version 17.0 (SPSS Inc, Chicago, Illinois).

RESULTS

Table 1 shows general characteristics of the study subjects by gender. Prevalence of obesity, diabetes and metabolic syndrome was high given that this study involved a population that was selected for being elderly and with a high cardiovascular risk. Total dairy product consumption was higher in women than in men (395.2+/-229.6 g/d vs. 322.3+/-193.9 g/d, respectively; P<0.001). Men consumed a greater amount of whole-fat milk (43.1+/-123.1 g/d vs. 35.3+/120.3 g/d; P=0.040), whereas women consumed more skimmed milk (114.8+/-191.0 g/d vs. 77.3+/-145.0 g/d; P=0.035). Women consumed a significantly greater amount of skimmed yoghurt (P<0.001), and there were no significant differences between men and women in the amount of whole-fat yoghurt consumed. Likewise, total cheese intake did not differ between men and women. The amount of lactose intake derived from dairy products was also significantly higher in women than men (P=0.01). However, there were no significant differences in the percentage of men and women who claim never to consume milk (14.2% vs. 14%; P=0.994).

Prevalence of the *LCT* -13910 C>T genotypes were: CC (LNP) 38.0% (n=357), CT 45.7% (n=430) and TT 16.3% (n=153). Carriers of the T allele were the genetically determined LP subjects. This distribution was in Hardy-Weinberg equilibrium (P=0.221) and did not differ between men and women (P=0.577).

Association between the rs4988235 SNP and dairy product intake

Table 2 shows mean intake of milk and dairy products (total and by gender) depending on the *LCT* -13910 C>T genotypes. The results are shown grouping the T carriers together (LP) and comparing them with CC subjects (LNP). Total energy intake did not differ between CC and subjects carrying the T- allele. Likewise, we did not find significant differences in physical activity depending on the *LCT* genotype (not shown). On analysing the results for men and women jointly, it is observed that although diary product consumption tended to be lower in CC subjects, the differences did not reach statistical significance. Neither was the total consumption of milk or the contribution of lactose or calcium through dairy products lower. Statistically significant differences were only reached in the consumption of skimmed yoghurt, which was lower in CC subjects.

On analysing the results per gender, it can be observed that in men the differences in milk and dairy product intake depending on genotype were minimal and did not reach statistical significance for any comparison. In women, these differences were more accentuated; reaching statistical significance in the consumption of skimmed yoghurt (lower in CC subjects) and when the consumption of skimmed yoghurt, skimmed milk and partially skimmed milk were analyzed together (265.1 +/- 207.9 g/d vs. 317.6+/-239.2 g/d in CC vs. CT+TT; P=0.014). Likewise, the total consumption of dairy products also reached statistically significant differences in women depending on the *LCT* genotype (P=0.045). No significant differences of lactose intake were observed.

Association between the rs4988235 SNP with anthropometric variables

On studying the relationship between the *LCT* genotype and anthropometric measurements (Table 3), we observed that this SNP presented a strong association with BMI and with waist circumference. CC individuals, although they do not differ in height from the other genotypes, presented less weight, a lower BMI and less waist-circumference than T-allele carriers. These differences remained statistically significant when the models were adjusted for gender and age, and even after additional adjustment for diabetes, physical activity and total energy intake. These associations are homogeneous by gender, and both in men and women CC subjects have lower means of anthropometric measurements than T-allele carriers (P for interaction *LCT* genotype x gender >0.05 for all the anthropometric variables).

Next, we analyzed the association of the *LCT* genotype with obesity (Table 3). Considering CC individuals as the reference category, we observed that T-allele carriers have a greater risk (OR) of obesity, both unadjusted (OR: 1.39; 95% CI: 1.07-1.82; P=0.014) and after adjustment for gender, age, diabetes, physical activity and total energy intake (OR: 1.37; 95% CI: 1.03-1.81; P=0.029). Homogeneity by gender was also observed between men and women in this association (P for interaction *LCT* x gender>0.05). Subsequent adjustments for dairy product intake do not modify the statistical significance of the associations of the *LCT* -13910 C>T SNP with the anthropometric variables (not shown).

Modulation of the association between the rs4988235 SNP with anthropometric variables by lactose intake

Considering that CC subjects may tolerate low amounts of lactose intake without gastrointestinal symptoms, we hypothesized that a higher lactose intake may modulate the effects of the rs4988235 SNP on anthropometric variables. We first tested the interaction effect between the rs4988235 SNP and dairy lactose intake as continuous. Taking into account that lactose intake was not normally distributed, eight identified outliers (corresponding to 8 TC+TT subjects with lactose intake higher that 50g/d) were removed to improve normality for this linear regression analysis. A statistically significant interaction term between lactose intake and the *LCT* -13910 C>T SNP in determining waist circumference was found (P=0.044 after adjustment for sex, age, diabetes, total energy intake and physical activity). According to this interaction, a higher dairy lactose intake increases the differences in waist-circumference between CC and CT+TT individuals (Figure 1 A). We also tested this modulation by lactose intake as categorical variable. Three categories of lactose intake based on habitual equivalents of milk consumption were considered (Figure 1B). If lactose intake was low [less than one small cup per day (<=8 g lactose/d)], we did not find significant differences in waist-circumference between CC and T-allele carriers (P=0.808). When lactose intake was higher [between 1 and 2 small or large cups of milk per day (8-24 g lactose/d)], significant differences in waist circumference between *LCT* genotypes were detected (P=0.012). These differences increased in magnitude when higher intakes of lactose were observed [more than 2 large cups of milk per day (>24 g lactose/d)].

In terms of obesity risk, in subjects with a low lactose intake (<=8 g/d) we did not find significant association between the *LCT* -13910 C>T SNP and obesity in the unadjusted model (OR=1.03, 95%CI: 0.55-1.912; P=0.910) or the model adjusted for sex, age, diabetes, physical activity and total energy intake. However, when lactose intake was higher (>8 g/d), we did observe a significant association of the CT+TT genotype with higher obesity risk (OR: 1.50, 95%CI: 1.10-2.03; P=0.012).

Association between the rs4988235 SNP with the metabolic syndrome related variables

Finally, we studied the association of the *LCT* genotype with biochemical parameters (fasting glucose, and plasma lipids), and blood pressure (Table 4) and observed that, after

adjustment for BMI, there were no statistically significant differences in total cholesterol concentrations, LDL-C, HDL-C, TG, fasting glucose and blood pressure between CC subjects and T-allele carriers. Likewise, when we analyzed the association of the *LCT* genotype with the Metabolic syndrome, taking CC individuals as the reference category, although the magnitude of the OR in T-allele carriers was greater than 1, it did not reach statistical significance, either in the crude model, or after adjustment for gender and age (OR: 1.26; 95%CI: 0.94-1.67).

DISCUSSION

In this study carried out on an elderly Mediterranean population, we have detected a high prevalence of CC individuals (LNP) in contrast to that found in northern European populations, where the opposite situation is reported (15,29). Although the effect of this genotype on the differences of milk and dairy produce consumption is not very high in the population as a whole, the association that we have found with anthropometric measurements is particularly relevant. In this Mediterranean population, CC individuals present lower BMI, lesser waist circumference and lower risk of obesity than T-allele carriers even after adjustment for sex, age, diabetes, physical activity and total energy intake. These results are in line with the results obtained in a recently published meta-analysis (23) that reported novel evidence of association between genetically defined LP and BMI (P = 7.9*10-5) in 31,720 individuals from four European populations. 8 cohorts were included in this study, 5 of which were of Finnish origin, and the others from Holland and England, without including a Mediterranean population. They observed that the CC genotype was associated with decreased BMI compared to CT/TT genotypes in the meta-analysis, and that this effect was observed in the same direction for both men and women. They also observed that this effect was due to the influence of the genotype on weight and not on height and discarded population stratification as being responsible for these effects. In our study, the probability of the influence of population stratification is very low, as we carried it out on White subjects recruited in a single region of Spain with a homogenous ethnic background. Moreover, another study undertaken on a sample recruited in another Spanish region (Catalonia) obtained a similar frequency of genotypes (30). In that study, aimed to investigate the association of the *LCT* SNP with osteoporosis phenotypes in 944 postmenopausal Spanish women, the researchers also found a significant association of the *LCT* SNP with weight. Hence, TT women were 1.91 kg heavier than CT+CC women, adding consistency to the association found in our

study. However, in that same study, the researchers did not analyse the possible differences in waist circumference or obesity risk associated with the *LCT* SNP, our study being the first to report such associations in a Mediterranean population.

Furthermore, our results are also the first to replicate the observations of the CHARGE Consortium, where strong associations were found between various SNPs in the *LCT* gene and waist-circumference (22). In line with those observations, a study carried out by Almon et al on 551 individuals of the general population of the Canary Islands (31), found that the prevalence of central obesity was higher in CT+TT (62%) individuals than in CC individuals (55.9%). However, those differences did not reach statistical significance possibly due to a smaller sample size. Nevertheless, what they did observe was a greater risk of metabolic syndrome in CT+TT individuals than in CC individuals (OR: 1.56; 95%CI: 1.06–2.31). In our study, carried out on a Mediterranean population, and despite the tendency of the association with metabolic syndrome being similar to that observed in the Canary islands, our results did not reach statistical significance, given that we observed the effects of the SNP mainly on anthropometric measurements and not on lipid concentrations or blood pressure. The differences observed between these two studies may be due to the different ages of the population and associated risk factors. Thus, although in the study carried out in the Canary Islands CC individuals were found to consume less milk than CT+TT individuals (246 vs. 300 g/d; $P<0.05$), our study found no significant differences in the amount of total milk consumed by carriers of these genotypes. Possibly, on dealing with an elderly population in which calcium requirements are greater to minimize osteoporosis, medical advice recommending higher milk consumption may have a greater influence, that recommendation offsetting the genetic influence.

Other studies that have analysed the influence of the *LCT* SNP on milk and dairy product consumption have also found differing associations depending on the age and gender of the population analysed (16-20). In general, it seems that the influence of the *LCT* genotype on dairy consumption is higher in women than in men (32, 33). So, for example, in the study carried out by Laaksonen et al in Finland (17), it was observed that until the age of 12 years, the consumption of milk and milk products in males with the C/C and C/T213910 genotypes was very similar and lower than for the T/T 213910 genotype. In females with the C/C213910 genotype, milk and milk product consumption was lowest from the age of 6 years. In our study,

CC women presented lower skimmed and semi-skimmed milk consumption than T-carriers, as well as a lower total dairy product consumption, which was not observed in men. One possible explanation could be that females and males differ in their sensitivity to gastrointestinal symptoms caused by mal-digested lactose. Previous studies have shown that women with lactose mal-digestion may experience stronger symptoms than men with lactose mal-digestion (34).

Another possible explanation could be that as men consume less milk than women, the amount of lactose does not pose a problem even for the lactase non-persistent, as it has been reported that gastrointestinal symptoms of intolerance to lactose are not severe until consuming amounts greater than 12 g of lactose/day (approximately 1 cup of milk) (35-36). The symptoms of lactose intolerance result from bacterial fermentation of undigested lactose in the colon (10). Various studies have shown differences in the microbial composition of fecal samples of the LP and LNP groups (37, 38). For example, Szilagyi et al (38), observed that lactose mal-digesters had a mean change difference (0.72 log10 colony forming units/g stool; P=0.04) in bifidobacteria counts compared with lactose digesters. Lactobacilli counts were also increased. Bearing in mind that recent studies have shown differences between the gut microbiota in obese and non-obese individuals (39, 40), changes in the gut microbiota may be involved in the lower risk of obesity observed in CC individuals, due to differing lactase fermentation capacity. This hypothesis, as well as the identification of additional mechanisms to explain the association of the *LCT* SNP with anthropometric measurements, requires further studies in order to substantiate it. Supporting this hypothesis is our observation of a possible greater effect of the *LCT* SNP on anthropometric measurements when the amounts of lactose consumed are greater.

ACKNOWLEDGEMENTS

Funding/Support: This work was supported by grants from the Ministerio de Ciencia e Innovación, Spain (CIBER CB06/03/0035, RD07/0067/0006, PI6-1326, PI07-0954, PI08-90002 and SAF-09-12304), the Generalitat Valenciana, Spain (GVACOMP2010-181, BEST2010-211, BEST2010-032) and the National Heart, Lung, and Blood Institute grants HL-54776, National Institute of Diabetes and Digestive and Kidney Diseases, Grant Number DK075030 and by contracts 53-K06-5-10 and 58-1950-9-001 from the US Department of Agriculture Research.

6 REFERENCES

1. Pereira MA, Jacobs DR Jr, Van Horn L, Slattery ML, Kartashov AI, Ludwig DS. Dairy consumption, obesity, and the insulin resistance syndrome in young adults: the CARDIA Study. JAMA. 2002 Apr 24;287(16):2081-9.

2. Mirmiran P, Esmaillzadeh A, Azizi F. Dairy consumption and body mass index: an inverse relationship. Int J Obes (Lond). 2005 Jan;29(1):115-21.

3. Azadbakht L, Mirmiran P, Esmaillzadeh A, Azizi F. Dairy consumption is inversely associated with the prevalence of the metabolic syndrome in Tehranian adults. Am J Clin Nutr. 2005 Sep;82(3):523-30.

4. Elwood PC, Pickering JE, Fehily AM. Milk and dairy consumption, diabetes and the metabolic syndrome: the Caerphilly prospective study. J Epidemiol Community Health. 2007 Aug;61(8):695-8.

5. Snijder MB, van Dam RM, Stehouwer CD, Hiddink GJ, Heine RJ, Dekker JM. A prospective study of dairy consumption in relation to changes in metabolic risk factors: the Hoorn Study. Obesity (Silver Spring). 2008 Mar;16(3):706-9. Epub 2008 Jan 17.

6. Wennersberg MH, Smedman A, Turpeinen AM, Retterstøl K, Tengblad S, Lipre E, Aro A, Mutanen P, Seljeflot I, Basu S, Pedersen JI, Mutanen M, Vessby B. Dairy products and metabolic effects in overweight men and women: results from a 6-mo intervention study. Am J Clin Nutr. 2009 Oct;90(4):960-8. Epub 2009 Aug 26.

7. Lamarche B. Review of the effect of dairy products on non-lipid risk factors for cardiovascular disease. J Am Coll Nutr. 2008 Dec;27(6):741S-6S. Review.

8. German JB, Gibson RA, Krauss RM, Nestel P, Lamarche B, van Staveren WA, Steijns JM, de Groot LC, Lock AL, Destaillats F. A reappraisal of the impact of dairy foods and milk fat on cardiovascular disease risk. Eur J Nutr. 2009 Jun;48(4):191-203. Epub 2009 Mar 4. Review.

9. Warensjo E, Nolan D, Tapsell L. Dairy Food Consumption and Obesity-Related Chronic Disease. Adv Food Nutr Res. 2010;59C:1-41. Epub 2010 Jun 24.

10. Suchy FJ, Brannon PM, Carpenter TO, Fernandez JR, Gilsanz V, Gould JB, Hall K, Hui SL, Lupton J, Mennella J, Miller NJ, Osganian SK, Sellmeyer DE, Wolf MA. National Institutes of Health Consensus Development Conference: lactose intolerance and health. Ann Intern Med. 2010 Jun 15;152(12):792-6.

11. Enattah NS, Sahi T, Savilahti E, Terwilliger JD, Peltonen L, Järvelä I. Identification of a variant associated with adult-type hypolactasia. Nat Genet. 2002 Feb;30(2):233-7. Epub 2002 Jan 14.

12. Olds LC, Sibley E. Lactase persistence DNA variant enhances lactase promoter activity in vitro: functional role as a cis regulatory element. Hum Mol Genet. 2003 Sep 15;12(18):2333-40. Epub 2003 Jul 22.

13. Lewinsky RH, Jensen TG, Møller J, Stensballe A, Olsen J, Troelsen JT. 2.T-13910 DNA variant associated with lactase persistence interacts with Oct-1 and stimulates lactase promoter activity in vitro. Hum Mol Genet. 2005 Dec 15;14(24):3945-53. Epub 2005 Nov 21.

14. Pohl D, Savarino E, Hersberger M, Behlis Z, Stutz B, Goetze O, Eckardstein AV, Fried M, Tutuian R. Excellent agreement between genetic and hydrogen breath tests for lactase deficiency and the role of extended symptom assessment. Br J Nutr. 2010 Apr 19:1-8.

15. Swallow DM. Genetics of lactase persistence and lactose intolerance. Annu Rev Genet. 2003;37:197-219.

16. Sacerdote C, Guarrera S, Smith GD, Grioni S, Krogh V, Masala G, Mattiello A, Palli D, Panico S, Tumino R, Veglia F, Matullo G, Vineis P. Lactase persistence and bitter taste response: instrumental variables and mendelian randomization in epidemiologic studies of dietary factors and cancer risk. Am J Epidemiol. 2007 Sep 1;166(5):576-81. Epub 2007 Jun 27.

17. Laaksonen MM, Mikkilä V, Räsänen L, Rontu R, Lehtimäki TJ, Viikari JS, Raitakari OT; Cardiovascular Risk in Young Finns Study Group. Genetic lactase non-persistence, consumption of milk products and intakes of milk nutrients in Finns from childhood to young adulthood. Br J Nutr. 2009 Jul;102(1):8-17. Epub 2009 Jan 13.

18. Torniainen S, Hedelin M, Autio V, Rasinperä H, Bälter KA, Klint A, Bellocco R, Wiklund F, Stattin P, Ikonen T, Tammela TL, Schleutker J, Grönberg H, Järvelä I. Lactase persistence, dietary intake of milk, and the risk for prostate cancer in Sweden and Finland. Cancer Epidemiol Biomarkers Prev. 2007 May;16(5):956-61.

19. Gugatschka M, Hoeller A, Fahrleitner-Pammer A, Dobnig H, Pietschmann P, Kudlacek S, Obermayer-Pietsch B.Calcium supply, bone mineral density and genetically defined lactose maldigestion in a cohort of elderly men. J Endocrinol Invest. 2007 Jan;30(1):46-51.

20. Gugatschka M, Dobnig H, Fahrleitner-Pammer A, Pietschmann P, Kudlacek S, Strele A, Obermayer-Pietsch B. Molecularly-defined lactose malabsorption, milk consumption and anthropometric differences in adult males. QJM. 2005 Dec;98(12):857-63.

21. Smith GD, Lawlor DA, Timpson NJ, Baban J, Kiessling M, Day IN, Ebrahim S. Lactase persistence-related genetic variant: population substructure and health outcomes. Eur J Hum Genet. 2009 Mar;17(3):357-67. Epub 2008 Sep 17.

22. Heard-Costa NL, Zillikens MC, Monda KL, Johansson A, Harris TB, Fu M, Haritunians T, Feitosa MF, Aspelund T, Eiriksdottir G, Garcia M, Launer LJ, Smith AV, Mitchell BD, McArdle PF, Shuldiner AR, Bielinski SJ, Boerwinkle E, Brancati F, Demerath EW, Pankow JS, Arnold AM, Chen YD, Glazer NL, McKnight B, Psaty BM, Rotter JI, Amin N, Campbell H, Gyllensten U, Pattaro C, Pramstaller PP, Rudan I, Struchalin M, Vitart V, Gao X, Kraja A, Province MA, Zhang Q, Atwood LD, Dupuis J, Hirschhorn JN, Jaquish CE, O'Donnell CJ, Vasan RS, White CC, Aulchenko YS, Estrada K, Hofman A, Rivadeneira F, Uitterlinden AG, Witteman JC, Oostra BA, Kaplan RC, Gudnason V, O'Connell JR, Borecki IB, van Duijn CM, Cupples LA, Fox CS, North KE. NRXN3 is a novel locus for waist circumference: a genome-wide association study from the CHARGE Consortium. PLoS Genet. 2009 Jun;5(6):e1000539. Epub 2009 Jun 26.

23. Kettunen J, Silander K, Saarela O, Amin N, Müller M, Timpson N, Surakka I, Ripatti S, Laitinen J, Hartikainen AL, Pouta A, Lahermo P, Anttila V, Männistö S, Jula A, Virtamo J, Salomaa V, Lehtimäki T, Raitakari O, Gieger C, Wichmann EH, Van Duijn CM, Smith GD, McCarthy MI, Järvelin MR, Perola M, Peltonen L.European lactase persistence genotype shows

evidence of association with increase in body mass index. Hum Mol Genet. 2010 Mar 15;19(6):1129-36. Epub 2009 Dec 16.

24. Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Ruiz-Gutiérrez V, Covas MI, Fiol M, Gómez-Gracia E, López-Sabater MC, Vinyoles E, Arós F, Conde M, Lahoz C, Lapetra J, Sáez G, Ros E; PREDIMED Study Investigators. Effects of a Mediterranean-style diet on cardiovascular risk factors: a randomized trial. Ann Intern Med. 2006 Jul 4;145(1):1-11.

25. Grundy SM, Cleeman JI, Daniels SR; et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute Scientific Statement. Circulation. 2005;112(17):2735-2752.

26. Martin-Moreno JM, Boyle P, Gorgojo L, Maisonneuve P, Fernandez-Rodriguez JC, Salvini S, Willett WC. Development and validation of a food frequency questionnaire in Spain. Int J Epidemiol 2003; 22:512–519.

27. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, Perez-Bauer M, Martínez-González MA, Salas-Salvadó J, Martín-Moreno JM. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr. 2010 Jun;103(12):1808-16.

28. Mataix J. Tablas de composición de alimentos (Spanish food composition tables), 4th edn. Universidad de Granada, Granada (in Spanish). 2003.

29. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. PLoS Comput Biol. 2009 Aug;5(8):e1000491. Epub 2009 Aug 28.

30. Agueda L, Urreizti R, Bustamante M, Jurado S, Garcia-Giralt N, Díez-Pérez A, Nogués X, Mellibovsky L, Grinberg D, Balcells S. Analysis of three functional polymorphisms in relation to osteoporosis phenotypes: replication in a Spanish cohort. Calcif Tissue Int. 2010 Jul;87(1):14-24.

31. Almon R, Alvarez-Leon EE, Engfeldt P, Serra-Majem L, Magnuson A, Nilsson TK. Associations between lactase persistence and the metabolic syndrome in a cross-sectional study in the Canary Islands. Eur J Nutr. 2010 Apr;49(3):141-6. Epub 2009 Oct 22.

32. Lehtimäki T, Hemminki J, Rontu R, Mikkilä V, Räsänen L, Laaksonen M, Hutri-Kähönen N, Kähönen M, Viikari J, Raitakari O. The effects of adult-type hypolactasia on body height growth and dietary calcium intake from childhood into young adulthood: a 21-year follow-up study--the Cardiovascular Risk in Young Finns Study. Pediatrics. 2006 Oct;118(4):1553-9.

33. Almon R, Patterson E, Nilsson TK, Engfeldt P, Sjöström M. Body fat and dairy product intake in lactase persistent and non-persistent children and adolescents. Food Nutr Res. 2010 Jun 16;54.

34. Vesa TH, Seppo LM, Marteau PR, et al. Role of irritable bowel syndrome in subjective lactose intolerance. Am J Clin Nutr 1998; 67, 710–715.

35. Wilt TJ, Shaukat A, Shamliyan T, Taylor BC, MacDonald R, Tacklind J, Rutks I, Schwarzenberg SJ, Kane RL, Levitt M. Lactose intolerance and health. Evid Rep Technol Assess (Full Rep). 2010 Feb;(192):1-410.

36. Hertzler SR, Huynh BC, Savaiano DA. How much lactose is low lactose? J Am Diet Assoc. 1996 Mar;96(3):243-6.

37. Zhong Y, Priebe MG, Vonk RJ, Huang CY, Antoine JM, He T, Harmsen HJ, Welling GW. The role of colonic microbiota in lactose intolerance. Dig Dis Sci. 2004 Jan;49(1):78-83.

38. Szilagyi A, Shrier I, Heilpern D, Je J, Park S, Chong G, Lalonde C, Cote LF, Lee B. Differential impact of lactose/lactase phenotype on colonic microflora. Can J Gastroenterol. 2010 Jun;24(6):373-9.

39. Sanz Y, Santacruz A, Gauffin P. Gut microbiota in obesity and metabolic disorders. Proc Nutr Soc. 2010 Aug;69(3):434-41. Epub 2010 Jun 14.

40. Muccioli GG, Naslain D, Bäckhed F, Reigstad CS, Lambert DM, Delzenne NM, Cani PD.The endocannabinoid system links gut microbiota to adipogenesis. Mol Syst Biol. 2010 Jul 27;6:392.

Table 1. Demographic, anthropometric, dietary and genetic characteristics of the studied subjects

	Men (n=336)	Women (n=602)
	Mean (SD)	Mean (SD)
Age (years)*	66 (7)	67 (6)
Weight (Kg)*	81.2 (12.0)	73.4 (11.0)
Height (m)*	1.68 (0.06)	1.55 (0.06)
Waist circumference (cm)*	104 (12)	102 (12)
BMI (Kg/m^2)*	29.0 (3.6)	30.7 (4.4)
Physical activity (kcal/d)*	220 (217)	136 (128)
Total energy intake (kcal/d)*	2364 (669)	2117 (614)
Total fat intake (% energy)	38.7 (7.6)	39.1 (6.9)
Carbohydrates (% energy)	41.8 (7.9)	42.8 (6.9)
Proteins (% energy)*	16.3 (2.7)	17.4 (2.7)
Alcohol (g/d)*	11.7 (14.9)	2.6 (4.9)
Total dairy products (g/d)*	322.3 (193.9)	396.2 (229.6)
Whole-fat milk (g/d)*	43.1 (123.1)	35.3 (120.3)
Partially skimmed milk (g/d)	103.5 (153.1)	127.4 (158.7)
Skimmed milk (g/d)*	77.3 (146.0)	114.8 (191.0)
Whole-fat yoghurt (g/d)	21.9 (64.2)	2[illegible].0 (66.0)
Skimmed yoghurt (g/d)*	36.9 (84.4)	56.0 (79.9)
Total cheese (g/d)	30.5 (23.2)	34.3 (27.5)
Lactose (g/d)*	12.9 (9.6)	16.9 (10.4)
Nonconsumers of milk (%)	14.2	14.0
Current smokers (%)*	27.9	4.1
Obesity (%)*	42.4	63.8
Diabetes (%)*	54.4	41.9
Metabolic syndrome (%)*	53.5	66.8
LCT (-13910C>T) genotype (%)		
CC (LNP)	39.3	37.2
CT (LP)	43.5	47.0
TT (LP)	17.2	15.8

SD: Standard deviation
[1]In metabolic syndrome was defined according to updated ATP III criteria
**: Statistically significant differences between men and women (Student's t test for continuous variables with normal distribution or the non-parametric Mann-Whitney U test for dairy products, alcohol and physical activity. Chi square tests for categorical variables)*
LNP: Lactase non-persistence; LP: Lactase persistence

Table 2: Association of the LCT rs4988235 polymorphism with dairy product consumption in the elderly Mediterranean population

	Whole population			Men			Women		
	CC	CT+TT	P	CC	CT+TT	P	CC	CT+TT	P
	LNP (*n*=357)	LP (*n*=583)		LNP (n=133)	LP (*n*=205)		LNP (n=283)	LP (n=378)	
	Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)		Mean (SD)	Mean (SD)	
Age (years)	67.5 (6.1)	66.8 (6.1)	0.075	66.9 (6.4)	66.2 (7.0)	0.472	67.8 (6.0)	67.1 (5.5)	0.153
Total energy intake (kcal/d)	2228.2 (646.3)	2204.4 (648.1)	0.599	2404.8 (693.9)	2370.8 (661.3)	0.617	2121.9 (599.7)	2113.9 (623.4)	0.882
Total dairy products (g/d)	394.1 (198.6)	377.8 (231.8)	0.124	330.2 (187.0)	317.1 (198.9)	0.387	368.9 (204.3)	410.7 (242.0)	0.049
Whole-fat milk (g/d)	46.1 (134.7)	33.2 (112.3)	0.068	45.4 (113.7)	41.6 (129.0)	0.440	46.6 (146.1)	28.6 (101.9)	0.088
Partially skimmed milk (g/d)	112.2 (185.4)	122.8 (183.6)	0.842	107.8 (180.0)	100.7 (148.8)	0.855	114.8 (188.9)	134.8 (199.2)	0.491
Skimmed milk (g/d)	98.9 (186.0)	103.9 (182.9)	0.942	84.0 (148.7)	73.0 (144.1)	0.292	104.7 (176.5)	120.8 (199.0)	0.434
Condensed milk (g/d)	0.2 (2.2)	0.2 (2.0)	0.311	0.1 (1.1)	0.1 (0.8)	0.589	0.2 (2.7)	0.3 (2.5)	0.639
Total milk (g/d)	266.4 (181.9)	260.2 (199.0)	0.866	237.3 (182.0)	215.4 (174.4)	0.125	266.3 (182.5)	284.5 (207.4)	0.468
Whole-fat yoghurt (g/d)	21.2 (52.4)	21.4 (56.5)	0.541	26.8 (60.3)	18.7 (49.7)	0.204	17.8 (46.8)	22.8 (60.0)	0.818
Skimmed yoghurt (g/d)	39.4 (80.9)	55.0 (82.1)	0.004	29.2 (54.6)	41.9 (80.7)	0.150	45.8 (83.7)	62.1 (87.4)	0.016
Skimmed milk and yoghurt (g/d)	248.5 (200.6)	317.6 (239.2)	0.053	221.0 (185.6)	215.6 (184.9)	0.762	265.1 (207.9)	317.6 (239.2)	0.014
Whipped cream (g/d)	0.2 (1.2)	0.4 (4.6)	0.305	0.2 (1.0)	0.8 (7.4)	0.699	0.2 (1.3)	0.3 (1.6)	0.321
Milkshake (g/d)	0.9 (10.0)	1.4 (15.3)	0.478	1.3 (14.1)	0.8 (6.7)	0.259	0.6 (6.3)	1.8 (18.4)	0.975
Ricotta cheese or junket (g/d)	1.5 (8.4)	1.4 (8.0)	0.378	1.4 (8.7)	1.4 (8.6)	0.991	1.6 (9.2)	1.4 (7.8)	0.268
Petit Suisse cheese (g/d)	0.6 (4.8)	0.3 (3.0)	0.717	0.0 (0.0)	0.0 (0.0)	0.980	0.9 (6.1)	0.4 (3.5)	0.927
Spreadable cheese wedges (g/d)	0.9 (2.7)	1.4 (4.8)	0.646	1.1 (3.1)	1.2 (4.1)	0.602	0.8 (2.5)	1.5 (5.1)	0.323
Cottage cheese (g/d)	13.0 (14.4)	15.7 (20.2)	0.101	9.5 (11.1)	12.6 (17.9)	0.296	15.1 (15.7)	17.4 (21.2)	0.274
Other cheese (hard cheese) (g/d)	16.1 (18.1)	14.8 (14.5)	0.739	18.0 (18.2)	16.0 (13.6)	0.706	15.0 (18.0)	14.1 (14.9)	0.914
Custard (g/d)	3.1 (11.9)	2.5 (8.8)	0.515	3.7 (13.4)	3.5 (11.5)	0.324	2.8 (10.9)	2.0 (6.9)	0.805
Ice cream (g/d)	1.8 (4.7)	3.2 (14.8)	0.973	1.9 (3.7)	4.5 (21.2)	0.996	1.8 (5.3)	2.4 (9.7)	0.755
Lactose (g/d)	14.4 (9.2)	15.1 (10.4)	0.441	13.4 (8.3)	12.6 (9.1)	0.311	15.0 (9.6)	16.5 (10.9)	0.115

SD: Standard deviation

P values for the comparison of means between CC and CT+TT subjects for dairy products were carried out by the non-parametric Mann-Whitney U test

Table 3: Association of the LCT rs4988235 polymorphism with anthropometric variables in the elderly Mediterranean population

	CC LNP (n=357)		CT LP (n=430)		TT LP (n=163)	P1	CC LNP (n=357)		CT+TT LP (n=593)		P1	P2	P3
	Mean	(SD)	Mean	(SD)	Mean (SD)		Mean	(SD)	Mean	(SD)			
Age (years)	67.5	(6.1)	66.9	(6.0)	66.7 (6.4)	0.196	67.5	(6.1)	66.8	(6.1)	0.080		
Height (m)	1.59	(0.09)	1.58	(0.09)	1.59 (0.09)	0.849	1.59	(0.09)	1.58	(0.09)	0.282	0.314	0.333
Weight (Kg)	78.2	(12.3)	76.6	(12.1)	77.8 (10.8)	0.047	78.2	(12.3)	76.8	(11.7)	0.044	0.032	0.021
BMI (Kg/m2)	29.7	(4.2)	30.6	(4.2)	30.7 (4.0)	0.011	29.7	(4.2)	30.6	(4.2)	0.003	0.001	0.002
Waist circumference (cm)	101.1	(11.6)	103.1	(11.3)	104.6 (11.9)	0.000	101.1	(11.6)	103.6	(11.5)	0.005	0.002	0.002
Obesity prevalence (%)	45.4		52.4		57.4	0.027	45.4		53.7		0.014		
OR for obesity (OR and 95% CI)	1 (ref)		1.33	(0.99-1.78)	1.62 (1.11-2.36)	0.038	1 (ref)		1.36	(1.07-1.85)	0.014	0.017	0.021

SD: Standard deviation

P1: Unadjusted P value
P2: Model adjusted for sex and age
P3: Model adjusted for sex, age, diabetes, physical activity and total energy intake

Table 4: Association of the LCT rs4988235 polymphism with plasma lipids, glucose, and blood pressure. Adjusted means*

	CC		CT		
	LNP (n=357)		LP (n=430)		
	Mean	(SE)	Mean	(SE)	P
Total cholesterol (mg/dL)	206.6	(2.2)	204.4	(1.8)	0.418
LDL-C (mg/dL)	129.67	(2.09)	127.10	(1.64)	0.324
HDL-C (mg/dL)	51.3	(0.7)	52.1	(0.6)	0.349
Triglycerides (mg/dL))	136.7	(4.6)	129.3	(3.6)	0.198
Fasting glucose (mg/dL)**	121.3	(1.8)	121.7	(1.4)	0.868
Systolic blood pressure (mm Hg)	146.3	(1.2)	147.4	(0.9)	0.422
Diastolic blood pressure (mm Hg)	81.7	(0.7)	82.6	(0.5)	0.309

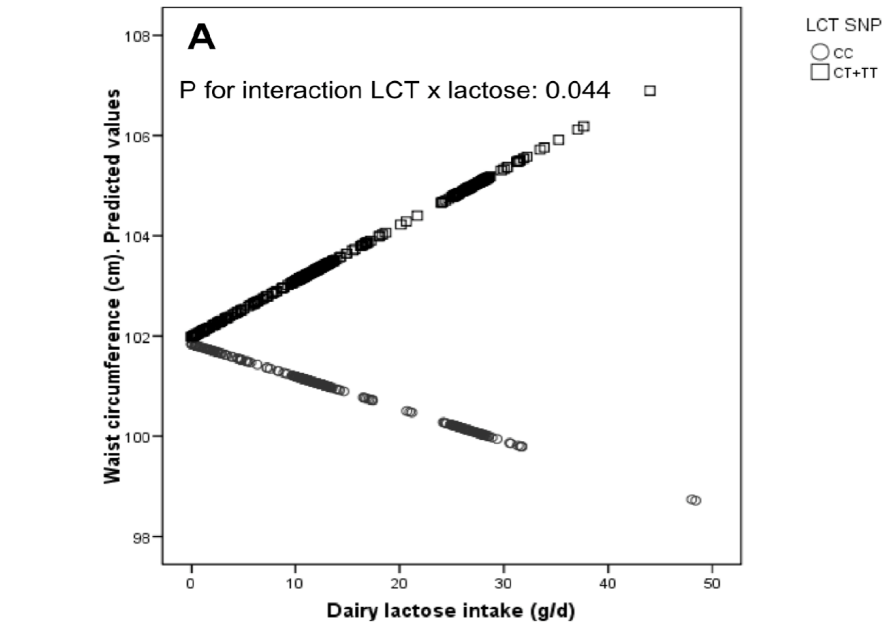
SE: Standard deviation
**Means were adjusted for sex, age and BMI*

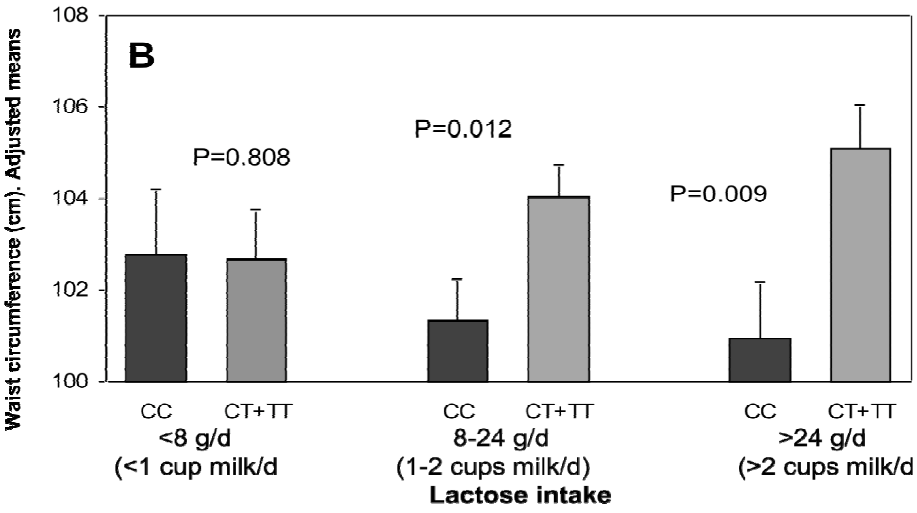
LEGENDS TO FIGURES

Figure 1 Modulation by dairy lactose intake of the association between the LCT −13910C>T polymorphism and waist circumference (cm) in the elderly Mediterranean population.

(a) Predicted values of waist circumference by the LCT −13910C>T

(b) Adjusted means of waist circumference (cm) in the study subjects (n = 940) depending on the LCT −13910C>T polymorphism according to three strata of lactose intake: low (≤8 g lactose/day; 20% of the population (n = 68 CC, 122 CT+TT)), intermediate (8–24 g lactose/day; 50% of the population (n = 188 CC, 284 CT+TT), and high (>24 g lactose/day; 30% of the population (n = 101 CC, 177 CT+TT).


A
P for interaction LCT x lactose: 0.044
LCT SNP
CC
CT+TT
Waist circumference (cm). Predicted values
108
106
104
102
100
98
0
10
20
30
40
50
Dairy lactose intake (g/d)


B
Waist circumference (cm). Adjusted means
108
106
104
102
100
P=0.808
P=0.012
P=0.009
CC
CT+TT
<8 g/d
(<1 cup milk/d
8-24 g/d
(1-2 cups milk/d)
>24 g/d
(>2 cups milk/d
Lactose intake

3

Significant associations of the rs2943634 (2q36.3) genetic polymorphism with adiponectin, high density lipoprotein cholesterol and ischemic stroke

Chapter 3

Significant associations of the rs2943634 (2q36.3) genetic polymorphism with adiponectin, high density lipoprotein cholesterol and ischemic stroke

Maria Arregui [a *], Eva Fisher [a, b], Sven Knüppel [a], Brian Buijsse [a], Romina di Giuseppe [a], Andreas Fritsche [c], Dolores Corella [d], Stefan N Willich [e], Heiner Boeing [a], Cornelia Weikert [a, e]

a Department of Epidemiology, German Institute of Human Nutrition (DIfE), Potsdam-Rehbrücke, Nuthetal, Germany.

b Administrative Office of the Commission on Genetic Testing, Robert Koch-Institute, Berlin.

c Department of Internal Medicine, Division of Endocrinology, Diabetology, Nephrology, Vascular Disease and Clinical Chemistry, University of Tübingen, Tübingen, Germany.

d Department of Preventive Medicine and CIBER Fisiopatología de la Obesidad y Nutrición. School of Medicine. University of Valencia, Valencia, Spain.

e Institute for Social Medicine, Epidemiology, and Health Economics, Charité University Medical Center, Berlin, Germany.

Key words: cardiovascular diseases, SNP, risk factor

Gene. 2012 Feb 25;494(2):190-5

Impact Factor 2012: 2.341

ABSTRACT

Background: rs2943634 C/A single nucleotide polymorphism (SNP), located in a non-coding region of chromosome 2q36.3, has been associated with coronary artery disease in two genome wide association studies. Our goal was to investigate its relation with myocardial infarction (MI) and ischemic stroke (IS), as well as with 12 intermediate risk phenotypes, in a population-based prospective cohort study.

Methods: rs2943634 was genotyped in a case-cohort study including a random sample of 1891 individuals (subcohort) and all incident MI (n = 211) and IS (n = 144) cases during a mean follow-up of 8.2 ± 2.2 years, nested within the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort comprising 27548 middle-aged men and women.

Results: rs2943634 minor allele (A) was associated in an additive fashion with lower risk of IS but not with MI [hazard ratio (HR) = 0.66; 95% confidence interval (CI): 0.50-0.87; P = 0.003; HR = 1.02; 95% CI: 0.82-1.28; P = 0.83 respectively, for the age and sex adjusted model]. Furthermore, it was related to slightly higher levels of plasma adiponectin [CC 6.94, CA 7.27, AA 7.86 µg/ml, P = 0.0002] and high density lipoprotein (HDL)-cholesterol (CC 52.08, CA 53.05 and AA 55.27 mg/dl, P = 0.002), based on additive models. Adjustment for adiponectin and HDL-cholesterol did not attenuate the association between the SNP and IS risk. In contrast, adjustment for adiponectin abolished the association between the SNP and HDL-cholesterol and adjustment for HDL-cholesterol attenuated the association between the SNP and adiponectin.

Conclusions: Our findings suggest that rs2943634 is associated with IS risk and with plasma levels of HDL-cholesterol and adiponectin in this German population. Further investigations are needed to confirm these results and to clarify the mechanisms underlying the association.3.1.

INTRODUCTION

rs2943634 C/A single nucleotide polymorphism (SNP) is located in a non-coding region on Chromosome 2q36.3, however the Welcome Trust Case Control Consortium (WTCCC) Study, which enrolled 1926 coronary artery disease (CAD) cases and 2938 controls of white European origin, imputed an association between its most frequent allele (C) and CAD (1). Authors could reproduce the association in the German myocardial (MI) family study which involved 875 MI cases and 1644 controls (1), however, further replication studies have not been able to confirm this association (2-7). Ischemic stroke (IS) shares common physiopathological mechanisms with MI due to atherosclerotic disease (8). Thus rs2943634 has been investigated on its relationship with IS (2) and carotid artery intima-media thickness, a subclinical marker of atherosclerosis associated with stroke (9), with no significant results. Association data between rs2943634 and intermediate risk phenotypes of cardiovascular diseases is scarce and inconsistent (1, 2, 9), although significant associations have been reported with body mass index (BMI) (1), blood pressure (2), low density lipoprotein (LDL) cholesterol (1) and high density lipoprotein (HDL) cholesterol (2). In the present study our goal was to investigate the association of rs2943634 with MI, IS, both cardiovascular diseases (CVD) combined, and with intermediate risk phenotypes that may implicate causal pathways, in a middle-aged population.

METHODS

Study population

The European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort comprises 27548 men and women from the general population of the Potsdam area in Germany, mainly aged 35–65 years at recruitment (1994–1998) (10). The associations of rs2943634 with risk of MI, IS and CVD were analysed using a case-cohort design. With this type of study design, the results are expected to be generalizable to the entire cohort (11, 12). The study population included a random sample of 2500 individuals (subcohort) and all newly occurred MI and IS cases from the EPIC-Potsdam cohort verified during a mean follow-up of 8.2 ± 2.2 years. Identification and verification of incident cases have been described in detail before (13). Of individuals who had both, MI and IS, we considered only the first event. After exclusion of individuals with a history of MI or IS at baseline, and those without fully obtained follow-up data

or with missing biomarkers, covariates or genotype data, the final study population comprised a subcohort of 1887 participants and a total of 211 MI cases,144 IS cases, (21 MI cases and 20 IS cases belonging to the subcohort). Participants were not required to be fasted at the baseline assessment, however, 654 individuals did not eat for at least 8 h before the blood drawing. The subcohort was used in a cross-sectional study to investigate the association between rs2943634 alleles and the following cardiovascular intermediate risk phenotypes: BMI, waist circumference (WC), systolic blood pressure (SBP), diastolic blood pressure (DBP), total cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides (TG), glucose, glycated haemoglobin (Hba1c), high-sensitivity C-Reactive Protein (hs-CRP), adiponectin and creatinine. Informed consent was obtained from all study participants, and approval was given by the Ethical Committee of the State of Brandenburg, Germany.

Biochemical and genetic analyses

At baseline 30 mL of venous blood was collected from all the study participants, it was fractionated into serum, plasma, buffy coat and erythrocytes and stored in liquid nitrogen until the time of analysis. All biomarkers were analysed between 2007 and 2008 at the Internal Medicine Department of the University of Tübingen, Germany. Plasma levels of total cholesterol, HDL-cholesterol, TG, hs-CRP, glucose and creatinine were determined with the automatic ADVIA 1650 analyser (Siemens Medical Solutions, Erlangen, Germany). LDL-cholesterol was determined using the Friedewald formula (14). Plasma total adiponectin was determined with an enzyme-linked immunosorbent assay (Linco Research, St Charles, Mo). Genotyping of whole genome amplified DNA samples was performed in 2008 with the TaqMan® System (Applied Biosystems, Foster City, CA, USA) at the Max Delbrück Center for Molecular Medicine (MDC), Berlin, Germany. The genotyping success rate was 99%.

Statistical analysis

Hardy–Weinberg equilibrium (HWE) of the SNP was tested using the χ^2 test. Baseline characteristics of the study populations were described as means ± standard deviation (SD) for continuous traits, medians and 25th and 75th percentiles for the skewed variables, and % for categorical variables. Association of the SNP with risk of MI, IS and CVD, was calculated as hazard ratios (HR) with 95% confidence interval (95% CI) using Cox proportional-hazard

regression, modified according to the Prentice method (11) to account for the case-cohort design, using a robust estimator for CI and considering the additive and co-dominant genetic models. Age was the underlying time variable in the counting processes, with entry defined as the individuals' age at the time of recruitment and exit defined as age at the diagnosis of CVD, or censoring. Analysis of covariance considering the additive genetic model was used to assess the association between rs2943634 genotypes (independent variables) and BMI, WC, SBP, DBP, Total Cholesterol, HDL-cholesterol, LDL-cholesterol, TG, blood glucose, Hba1c, hs-CRP, creatinine, and adiponectin (dependent variables). Data is reported as means and standard error (SEM), except for TG, hs-CRP, creatinine, and adiponectin which were used log-transformed in order to normalize their distributions, and data is reported as geometric means and 95% CI. Also to better reach normality of their distributions, glucose was used squared root transformed and Hba1c inversed transformed, data for them is respectively reported as squared and inverse means and 95% CI. P for trend was calculated with a linear regression model. Association analysis between rs2943634 and fasting depending variables were performed only in participants who were fasting at the time of blood drawing. Association between adiponectin and HDL-cholesterol in the subcohort was examined by means of Spearman partial correlation coefficient.

The statistical models were adjusted for sex and age. Further adjustment included known cardiovascular risk factors: smoking status (never smoker, former smoker, current smoker < 20 cigarettes per day, current smoker ≥ 20 cigarettes per day), sports activity (< 2 h/wk versus ≥ 2 h/wk), educational attainment (vocational school or less, technical school, university), BMI (continuous), WC (continuous), alcohol consumption (men: =0 g/d, >0 to 12 g/d, >12 to 24 g/d; >24 g/d; women: =0 g/d, >0 to 6 g/d, >6 to 12 g/d; >12 g/d), Hba1c, systolic blood pressure, antihypertensive medication, total cholesterol and hs-CRP. Effect modification by sex was evaluated by modeling the product term sex x genotype along with main effects (in the age-adjusted linear regression analysis). All statistical analyses were performed using the SAS Version 9.2 (SAS Institute, Cary, NC). The detectable odds ratio of the SNP with CVD, MI and IS was calculated with Quanto (http://hydra.usc.edu/GxE/) (15) concerning a log-additive model, a desirable power > 80%, $\alpha = 0.05$, and a baseline overall disease prevalence in the EPIC-Potsdam population of 3.1 % for CVD (n=847), 2.0% for MI (n=544) and 1.1% for IS (n=303). Our estimates suggested that the detectable odds ratio were 1.26 for CVD, 1.34 for MI and 1.41

for IS when considering a possible harmful effect of the SNP, and 0.79, 0.75 and 0.71 respectively when considering a possible protective effect.

RESULTS

Characteristics of participants

Demographic, biochemical, clinical and lifestyle characteristics of the subcohort and incident IS and MI cases are given in **Table 1**. The genotype frequency of rs2943634 in the subcohort followed HWE (P = 0.80). The A allele (ancestral allele) frequency across the subcohort was 36%, in accordance with other European populations (1, 2). No significant effect modifications by sex were found for the SNP on CVD, MI or IS and the studied intermediate risk phenotypes except for glucose (P = 0.03), for which we therefore present results further adjusted for the interaction term.

Association between rs2943634 and intermediate risk phenotypes of cardiovascular diseases

Table 2 shows results for age and sex adjusted association analysis between rs2943634 and intermediate risk phenotypes of cardiovascular diseases in the subcohort. No significant associations were found when considering BMI, WC, SBP, DBP, total cholesterol, LDL-cholesterol, glucose, Hba1c, hs-CRP or TG. However, the rs2943634 A allele, was associated with increased HDL-cholesterol and adiponectin, and very modestly decreased creatinine. Exclusion of participants taking lipid or antihypertensive medication did not modify these results (data not shown). Further (mutual) adjustment for lifestyle, biochemical and socio-demographical variables did not modify these results. However, further adjustment of the association analysis between rs2943634 and HDL-cholesterol for adiponectin, abolished the association, and further adjustment of the association analysis between the rs2943634 and adiponectin for HDL-cholesterol, slightly attenuated the association (**Table 2**). Adiponectin and HDL-cholesterol were positively correlated (Spearman correlation coefficient for the age and sex adjusted model: 0.44; P < 0.0001)

Association between rs2943634 and CVD, MI and IS

Disease risk was evaluated in 355 CVD, 211 MI and 144 IS incident cases, and 1846 persons who remained free of cardiovascular events during follow-up. **Table 3** shows HR and 95% CI, for the association analysis based in the co-dominant and additive models between rs2943634 and CVD, MI and IS, adjusted for age and sex (model 1) and further adjusted for known cardiovascular risk factors (model 2). No significant associations were found between the SNP and CVD or MI regardless of the adjusting model used. However rs2943634 minor allele A, presented lower risk of IS in an allele dosage fashion (HR = 0.66; 95% CI: 0.51-0.86; P = 0.002). Further adjustment for model 2, or for model 2 plus the intermediate risk phenotypes significantly associated with the SNP, HDL-Cholesterol and adiponectin (model 3) did not affect this association (HR = 0.66; 95% CI: 0.51-0.87; P = 0.003; and HR = 0.66; 95% CI: 0.50-0.86; P = 0.002, respectively).

DISCUSSION

In this study we found a novel association between rs2943634 minor allele (A) and decreased risk of IS. Interestingly, this allele was also associated with slightly higher HDL-cholesterol and adiponectin. HDL-cholesterol has been associated with lower IS risk (16), and although in the scarce prospective studies available (17-21) adiponectin has been reported not to be associated with stroke risk, it has been proposed to be a marker of conditions that prompt to cardiovascular diseases (19). However, adjustment for these biomarkers, did not attenuate the association between rs2943634 and IS, remaining still to be elucidated the biological mechanism underlying this association. Adiponectin is an adipocyte-derived hormone suggested to exert antiatherogenic, anti-inflammatory, insulin-sensitizing and cardioprotective effects (22). It is thought to regulate the synthesis of HDL-cholesterol by means of the up-regulation of Apolipoprotein A-I (apoA-I) synthesis and secretion (23), and by the enhance of the expression of ATP-binding cassette transporter (ABCA1) mRNA and protein levels (24). Adiponectin has also been suggested to have a role in the regulation of HDL-cholesterol catabolism and remodelling mediated by cholesteryl ester transfer protein (CETP) and lipoprotein lipase (LPL) (25, 26), and it has been shown to correlate negatively with the fractional catabolic rate of ApoA-I (27). Several epidemiological studies have shown adiponectin to be positively correlated with HDL-cholesterol

(27-30). In accordance to this, adiponectin and HDL-cholesterol were positively correlated in our study, and when we adjusted the association analysis of rs2943634 with HDL-cholesterol for adiponectin, the association was abolished, suggesting the SNP might be related to HDL-cholesterol by means of adiponectin. Besides the role of adiponectin on HDL-cholesterol regulation, it has recently been suggested that HDL-cholesterol exerts reciprocal effects on adiponectin expression and adipocyte metabolism (31). Thus, HDL-cholesterol has been seen to increase plasma adiponectin concentrations and enhance adiponectin expression in adipocytes in a phosphatidylinositol-3- kinase (PI3K) dependent manner (31). In agreement with this observation, when we adjusted rs2943634 association analysis with adiponectin for HDL-cholesterol, the association was slightly attenuated.

Since MI and IS share common risk factors, it is surprising that we found a strong association of the SNP with risk of IS, but not with risk of MI. However, several genetic polymorphisms (32-36) and also established risk factors for both endpoints, including hypertension or smoking (37, 38), also have shown to differ in their importance for the development of MI and IS. In line with our results for MI, no replication studies have been able to confirm the association between rs2943634 and CAD (2-7) after it was first suggested by the WTCCC study and replicated in the German MI family study (1). Thus, not the three case-control replication studies with European [3] Tunisian [5] and Caucasian from the Cleveland genebank (6) populations, nor a biracial, prospective cohort study of US persons (7) reported significant results. Further, a case-cohort study on the effect of rs2943634 on incident CHD and stroke, based on five cohorts within the MORGAM project (from Finland, Sweden, France and Northern Ireland), and including a subcohort of 2341 participants, 1436 incident CHD cases and 571 incident stroke cases, also found no significant associations with cardiovascular endpoints (2). This study is in addition the only previous to investigate the effect of the SNP on incident stroke, and its results are inconsistent with ours for this outcome (2). Unfortunately, authors do no provide separate risk estimates for the five European cohorts, thus although adjusted for the different cohorts, we cannot exclude that results might be driven by the two Finish cohorts, since they represent 80% of the study population. Also in contrast with our study, 88% of the participants were men, nearly half of them belonged to cohort of smokers (39), and the average age was 60.5 years. In our study, participants are of German origin, only 37% are men, the average age is 10 years younger,

and only 21% are current smokers. Perhaps these differences may have played a role in the inconsistent results.

rs2943634 was not associated with BMI, WC, SBP, DBP, total cholesterol, LDL-cholesterol, glucose, Hba1c and hs-CRP in our study population. Concerning the associations between rs2943634 and these intermediate risk phenotypes, results across studies are also inconsistent. Thus, rs2943634 has been reported to be associated with BMI, LDL-cholesterol (1) and blood pressure (1, 2). and not be associated with HDL-cholesterol (6, 9), however, and in line with our results, rs2943634 minor allele (A) has also been reported not to be associated with blood pressure, BMI, WC, cholesterol, hs-CRP (9), LDL-cholesterol and TG (6, 9), and to be associated with higher HDL-cholesterol (1.05 mg/dL increase per allele) (2). We are the first to report association analysis results between rs2943634 and adiponectin, glucose, Hba1c and creatinine.

Strengths and limitations of our study should be mentioned. The prospective design may include other cases than case-control studies and thereby come up with new insights into the relationship between genetic information and disease risk. The risk factors were assessed at baseline only, therefore, we assumed that these variables remained stable over time; fasting data was available only for about one third of the population, thus, association analysis for TG, LDL-cholesterol and glucose were performed in this reduced sample. Our study sample consisted of a modest number of incident MI cases (n = 211), however, given that the estimated hazard ratios for MI were close to 1.0 regardless of the adjusting model used or the heritage model considered, we believe the lack of association of the SNP with MI in our population was not derived from insufficient statistical power. Finally, all the reported significance levels were nominal p values and were not adjusted for multiple comparisons, however after stringent Bonferroni correction for the testing of multiple hypotheses ($P_{corrected}$ <0.003) results for IS, adiponectin and HDL-cholesterol remained significant.

CONCLUSIONS

rs2943634 minor allele A was associated with lower risk of IS in this European population but not by means of the classic risk factors investigated in this study. rs2943634 was also related to plasma levels of adiponectin and HDL-cholesterol. However, considering the inconsistent results across studies, these findings need to be replicated in further investigations.

ACKNOWLEDGEMENTS

The authors are grateful to Anna Bury, Sabina Herbert, Albrecht Pfäfflin and Bianca Krüger-Karstensen who were involved in the biochemical and genetic analyses. We are indepted to Wolfgang Fleischhauer for case ascertainment, to Ellen Kohlsdorf for data management and to Wolfgang Bernigau for statistical consultation. We thank Kathrin Saar and Norbert Hübner for support and discussion. This work was financially supported by the German Federal Ministry of Education. The recruitment phase of the EPIC-Potsdam study was funded by the German Federal Ministry of Science (01 EA 9401) and the European Union (SOC 95201408 05F02). The follow-up of the EPIC-Potsdam study was supported by the German Cancer Aid (70-2488-Ha I) and the European Union (SOC 98200769 05F02).

REFERENCES

1. Samani NJ, Erdmann J, Hall AS, Hengstenberg C, Mangino M, Mayer B, et al. Genomewide association analysis of coronary artery disease. N Engl J Med. 2007;357(5):443-53.

2. Karvanen J, Silander K, Kee F, Tiret L, Salomaa V, Kuulasmaa K, et al. The impact of newly identified loci on coronary heart disease, stroke and total mortality in the MORGAM prospective cohorts. Genet Epidemiol. 2009;33(3):237-46.

3. Samani NJ, Deloukas P, Erdmann J, Hengstenberg C, Kuulasmaa K, McGinnis R, et al. Large scale association analysis of novel genetic loci for coronary artery disease. Arterioscler Thromb Vasc Biol. 2009;29(5):774-80.

4. Muendlein A, Saely CH, Rhomberg S, Sonderegger G, Loacker S, Rein P, et al. Evaluation of the association of genetic variants on the chromosomal loci 9p21.3, 6q25.1, and 2q36.3 with angiographically characterized coronary artery disease. Atherosclerosis. 2009;205(1):174-80.

5. Ghazouani L, Khalifa SB, Abboud N, Perret C, Nicaud V, Ben Khalfallah A, et al. Association of three polymorphisms selected from a genome-wide association study with coronary heart disease in the Tunisian population. J Thromb Thrombolysis. 2010;29(1):114-8.

6. Wang AZ, Li L, Zhang B, Shen GQ, Wang QK. Association of SNP rs17465637 on Chromosome 1q41 and rs599839 on 1p13.3 with Myocardial Infarction in an American Caucasian Population. Ann Hum Genet. 2011.

7. Bressler J, Folsom AR, Couper DJ, Volcik KA, Boerwinkle E. Genetic variants identified in a European genome-wide association study that were found to predict incident coronary heart disease in the atherosclerosis risk in communities study. Am J Epidemiol. 2010;171(1):14-23.

8. Faxon DP, Fuster V, Libby P, Beckman JA, Hiatt WR, Thompson RW, et al. Atherosclerotic Vascular Disease Conference: Writing Group III: pathophysiology. Circulation. 2004;109(21):2617-25.

9. Cunnington MS, Mayosi BM, Hall DH, Avery PJ, Farrall M, Vickers MA, et al. Novel genetic variants linked to coronary artery disease by genome-wide association are not associated with carotid artery intima-media thickness or intermediate risk phenotypes. Atherosclerosis. 2009;203(1):41-4.

10. Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. Ann Nutr Metab. 1999;43(4):205-15.

11. Prentice RL, Self SG. Aspects of the use of relative risk models in the design and analysis of cohort studies and prevention trials. Stat Med. 1988;7(1-2):275-87.

12. Kulathinal S, Karvanen J, Saarela O, Kuulasmaa K. Case-cohort design in practice - experiences from the MORGAM Project. Epidemiol Perspect Innov. 2007;4:15.

13. Weikert C, Stefan N, Schulze MB, Pischon T, Berger K, Joost HG, et al. Plasma fetuin-a levels and the risk of myocardial infarction and ischemic stroke. Circulation. 2008;118(24):2555-62.

14. 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(6):499-502.

15. Gauderman WJ MJ. QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. 2006. http://hydauscedu/gxe Last accessed: August 2011.

16. Wannamethee SG, Shaper AG, Ebrahim S. HDL-Cholesterol, total cholesterol, and the risk of stroke in middle-aged British men. Stroke. 2000;31(8):1882-8.

17. Soderberg S, Stegmayr B, Stenlund H, Sjostrom LG, Agren A, Johansson L, et al. Leptin, but not adiponectin, predicts stroke in males. J Intern Med. 2004;256(2):128-36.

18. Stott DJ, Welsh P, Rumley A, Robertson M, Ford I, Sattar N, et al. Adipocytokines and risk of stroke in older people: a nested case-control study. Int J Epidemiol. 2009;38(1):253-61.

19. Matsumoto M, Ishikawa S, Kajii E. Association of adiponectin with cerebrovascular disease: a nested case-control study. Stroke. 2008;39(2):323-8.

20. Khalili P, Flyvbjerg A, Frystyk J, Lundin F, Jendle J, Engstrom G, et al. Total adiponectin does not predict cardiovascular events in middle-aged men in a prospective, long-term follow-up study. Diabetes Metab. 2010;36(2):137-43.

21. Rajpathak SN, Kaplan RC, Wassertheil-Smoller S, Cushman M, Rohan TE, McGinn AP, et al. Resistin, but Not Adiponectin and Leptin, Is Associated With the Risk of Ischemic Stroke Among Postmenopausal Women: Results From the Women's Health Initiative. Stroke. 2011.

22. Brochu-Gaudreau K, Rehfeldt C, Blouin R, Bordignon V, Murphy BD, Palin MF. Adiponectin action from head to toe. Endocrine. 2010;37(1):11-32.

23. Matsuura F, Oku H, Koseki M, Sandoval JC, Yuasa-Kawase M, Tsubakio-Yamamoto K, et al. Adiponectin accelerates reverse cholesterol transport by increasing high density lipoprotein assembly in the liver. Biochem Biophys Res Commun. 2007;358(4):1091-5.

24. Oku H, Matsuura F, Koseki M, Sandoval JC, Yuasa-Kawase M, Tsubakio-Yamamoto K, et al. Adiponectin deficiency suppresses ABCA1 expression and ApoA-I synthesis in the liver. FEBS Lett. 2007;581(26):5029-33.

25. van Hoek M, van Tol A, van Vark-van der Zee LC, Jansen H, Kastelein JJ, Sijbrands EJ, et al. Role of plasma adiponectin on the HDL-cholesterol raising effect of atorvastatin in patients with type 2 diabetes. Curr Med Res Opin2009. p. 93-101.

26. Cnop M, Havel PJ, Utzschneider KM, Carr DB, Sinha MK, Boyko EJ, et al. Relationship of adiponectin to body fat distribution, insulin sensitivity and plasma lipoproteins: evidence for independent roles of age and sex. Diabetologia. 2003;46(4):459-69.

27. Verges B, Petit JM, Duvillard L, Dautin G, Florentin E, Galland F, et al. Adiponectin is an important determinant of apoA-I catabolism. Arterioscler Thromb Vasc Biol. 2006;26(6):1364-9.

28. Altinova AE, Toruner F, Bukan N, Yasar DG, Akturk M, Cakir N, et al. Decreased plasma adiponectin is associated with insulin resistance and HDL cholesterol in overweight subjects. Endocr J. 2007;54(2):221-6.

29. Schulze MB, Rimm EB, Shai I, Rifai N, Hu FB. Relationship between adiponectin and glycemic control, blood lipids, and inflammatory markers in men with type 2 diabetes. Diabetes Care. 2004;27(7):1680-7.

30. Zietz B, Herfarth H, Paul G, Ehling A, Muller-Ladner U, Scholmerich J, et al. Adiponectin represents an independent cardiovascular risk factor predicting serum HDL-cholesterol levels in type 2 diabetes. FEBS Lett. 2003;545(2-3):103-4.

31. Van Linthout S, Foryst-Ludwig A, Spillmann F, Peng J, Feng Y, Meloni M, et al. Impact of HDL on adipose tissue metabolism and adiponectin expression. Atherosclerosis. 2010;210(2):438-44.

32. Sie MP, Uitterlinden AG, Bos MJ, Arp PP, Breteler MM, Koudstaal PJ, et al. TGF-beta 1 polymorphisms and risk of myocardial infarction and stroke: the Rotterdam Study. Stroke. 2006;37(11):2667-71.

33. Lin HF, Tsai PC, Liao YC, Lin TH, Tai CT, Juo SH, et al. Chromosome 9p21 Genetic Variants Are Associated With Myocardial Infarction But Not With Ischemic Stroke in a Taiwanese Population. J Investig Med. 2011.

34. Siegerink B, Rosendaal FR, Algra A. Genetic variation in fibrinogen; its relationship to fibrinogen levels and the risk of myocardial infarction and ischemic stroke. J Thromb Haemost. 2009;7(3):385-90.

35. M DEG, Quacquaruccio G, Pezzini A, Latella MC, A DIC, Del Zotto E, et al. Tissue factor gene polymorphisms and haplotypes and the risk of ischemic vascular events: four studies and a meta-analysis. J Thromb Haemost. 2009;7(9):1465-71.

36. Lemaitre RN, Rice K, Marciante K, Bis JC, Lumley TS, Wiggins KL, et al. Variation in eicosanoid genes, non-fatal myocardial infarction and ischemic stroke. Atherosclerosis. 2009;204(2):e58-63.

37. Heidemann C, Hoffmann K, Klipstein-Grobusch K, Weikert C, Pischon T, Hense HW, et al. Potentially modifiable classic risk factors and their impact on incident myocardial infarction: results from the EPIC-Potsdam study. Eur J Cardiovasc Prev Rehabil. 2007;14(1):65-71.

38. Weikert C, Berger K, Heidemann C, Bergmann MM, Hoffmann K, Klipstein-Grobusch K, et al. Joint effects of risk factors for stroke and transient ischemic attack in a German population: the EPIC Potsdam Study. J Neurol. 2007;254(3):315-21.

39. Group TACPS. The alpha-tocopherol, beta-carotene lung cancer prevention study: design, methods, participant characteristics, and compliance. The ATBC Cancer Prevention Study Group. Ann Epidemiol. 1994;4(1):1-10.

Table 1. Baseline characteristics in EPIC-Potsdam subcohort, incident MI and IS cases.

	Subcohort (N = 1891)	MI cases (N = 211)	IS cases (N = 144)
Men (%)	36.62	72.51	49.31
Age (year)	50.13 ± 9.03	55.79 ± 7.10	56.64 ± 8.05
Smoking status (%):			
Current ≥20 cigarettes/d	6.09	20.38	8.33
Current <20 cigarettes/d	14.85	23.22	15.28
Former ≤5 years	7.15	6.16	3.47
Former >5 years	24.27	22.75	29.86
Never	47.64	27.49	43.06
Use of medication (%)			
Antihypertensive	17.22	30.81	35.42
Lipid lowering medication	4.03	7.11	8.33
Educational attainment (%):			
Vocational school or less	37.41	39.81	42.36
Technical school	24.54	20.38	30.56
University degree	38.05	39.81	27.08
Physical activity (h/w):			
< 2	76.21	80.57	83.33
≥ 2	23.79	19.43	16.67
Alcohol intake (%):			
Men: = 0 g/d; women: = 0 g/d	0.00	0.00	0.00
Men: > 0-12 g/d; women: >0-6 g/d	47.32	57.35	47.92
Men: >12-24 g/d; women: >6-12 g/d	25.28	18.48	23.61
Men: > 24 g/d; women: > 12 g/d	27.40	24.17	28.47
Body mass index (kg/m^2)	26.02 ± 4.31	27.66 ± 4.02	26.70 ± 4.07
Waist circumference (cm)			
Men	94.06 ± 10.16	97.37 ± 10.05	95.09 ± 11.55
Women	80.59 ± 11.62	86.31 ± 11.96	83.74 ± 11.16
Systolic blood pressure (mm/Hg)	129.09 ± 17.66	139.69± 18.46	140.50 ± 19.99
Diastolic blood pressure (mm/Hg)	83.55 ± 10.78	87.91 ± 11.12	88.49 ± 10.47
Total cholesterol (mg/dl)	173.85 ± 36.50	190.18 ± 39.39	175.03± 40.06
HDL-cholesterol (mg/dl)	52.96 ± 14.25	45.42 ± 13.92	50.13 ± 15.82
LDL-cholesterol (mg/dl)[a]	108.56 ± 29.76	125.73 ± 30.21	108.74 ± 31.22
Triglycerides (mg/dl)[b]	101.20 (71.50; 148.50)	139.70 (90.20; 215.60)	113.30 (77.00; 183.70)
Blood glucose (mg/dl)	95.43 (88.25; 104.70)	99.86 (91.21; 110.20)	97.53 (86.76; 114.28)
HbA1c (%)	6.37 (6.09; 6.73)	6.67 (6.39; 7.07)	6.58 (6.29; 7.11)
Hs-C-reactive protein (mg/l)	6.60 (2.20; 20.90)	15.40 (5.50; 35.20)	9.90 (3.30; 34.10)
Adiponectin (ng/ml)	7.44 (5.30; 10.27)	6.04 (4.60; 8.90)	7.37 (5.44; 10.88)
Creatinine (mg/dl)	0.71 (0.62; 0.81)	0.79 (0.70; 0.90)	0.71 (0.64; 0.84)

Mean ± SD or %. a, Due to missing fasting data, LDL-cholesterol, TG and glucose are based on 543 subcohort individuals, 68 MI cases, and 55 IS cases. b, Median (25th percentile; 75th percentile), all such values. MI indicates myocardial infarction, IS indicates ischemic stroke.

Table 2. Associations between rs2943634 polymorphism and cardiovascular diseases intermediate risk phenotypes in the EPIC-Potsdam subcohort.

Characteristics	CC (N = 767)	CA (N = 876)	AA (N = 244)	P_{add}
Body mass index (kg/m^2)[a]	26.18 ± 0.15	25.88 ± 0.14	26.02 ± 0.27	0.29
Waist circumference (cm)				
Men	94.45 ± 0.61	93.57 ± 0.53	94.92 ± 1.13	0.82
Women	80.72 ± 0.49	80.69 ± 0.48	79.83 ± 0.85	0.46
SBP (mm/Hg)	129.18 ± 0.59	129.07 ± 0.55	128.92 ± 1.04	0.82
DBP (mm/Hg)	83.66 ± 0.38	83.54 ± 0.35	83.26 ± 0.67	0.62
Total Cholesterol (mg/dl)	f174.80 ± 1.27	173.07 ± 1.19	173.67 ± 2.26	0.40
HDL-cholesterol (mg/dl)	52.08 ± 0.49	53.05 ± 0.46	55.27 ± 0.87	0.002
HDL-cholesterol (mg/dl)[b]	52.47 ± 0.46	52.99 ± 0.43	54.39 ± 0.81	0.06
LDL-cholesterol (mg/dl)[c]	109.55 ± 2.04	108.64 ± 1.76	105.46 ± 3.43	0.35
Triglycerides (mg/dl)[d]	94.81 (88.14-101.98)	83.11 (78.02-88.52)	94.29 (83.45-106.55)	0.28
Blood glucose (mg/dl)[e, f]	98.97 (91.03-107.24)	96.10 (92.73-99.52)	92.22 (79.48-105.91)	0.56
HbA1c (%) [g]	6.47 (6.42-6.51)	6.43 (6.39-6.47)	6.41 (6.33-6,49)	0.15
Hs-C-reactive protein (mg/l)	7.59 (6.89-8.36)	6.94 (6.34-7.60)	7.93 (6.68-9.42)	0.84
Adiponectin (µg/ml)	6.94 (6.72-7.17)	7.27 (7.05-7.49)	7.86 (7.43-8.33)	0.0002
Adiponectin (µg/ml)[h]	7.01 (6.80-7.23)	7.26 (7.06-7.47)	7.66 (7.25-8.08)	0.005
Creatinine (µg g/dl)	0.72 (0.71-0.73)	0.70 (0.69-0.71)	0.70 (0.69-0.72)	0.004

All analyses were adjusted for age and sex. a, Means and standard error (SEM), all such values. b, Further adjusted for adiponectin; c, LDL-cholesterol, TG and glucose are based on 543 individuals due to missing fasting data. d, Geometric means and 95% CI, all such values. e, Squared means and 95% CI; f, Further adjusted by the sex x rs2943634 interaction term; g, inverse and 95% (CI); h, further adjusted for HDL-cholesterol.

Table 3. Hazard ratios (HR) and 95% confidence intervals (95% CI) for the associations between rs2943634 polymorphism and cardiovascular endpoints.

	CC	CA	AA	Per A allele	P add
CVD cases	N = 158	N =158	N =39		
Model 1[a]	1.0	0.84 (0.66-1.08)	0.74 (0.49-1.10)	0.85 (0.71-1.02)	0.08
Model 2[b]	1.0	0.83 (0.64-1.07)	0.73 (0.49-1.10)	0.85 (0.71-1.02)	0.08
Model 3[c]	1.0	0.87 (0.66-1.13)	0.71 (0.46-1.10)	0.85 (0.70-1.03)	0.09
MI cases	N = 86	N = 96	N = 29		
Model 1[a]	1.0	1.02 (0.74-1.41)	1.06 (0.64-1.73)	1.02 (0.82-1.28)	0.83
Model 2[b]	1.0	1.02 (0.73-1.42)	1.06 (0.64-1.74)	1.02 (0.82-1.29)	0.83
Model 3[c]	1.0	1.03 (0.72-1.47)	1.04 (0.62-1.74)	1.02 (0.81-1.30)	0.86
IS cases	N = 72	N = 62	N = 10		
Model 1[a]	1.0	0.72 (0.50-1.04)	0.38 (0.18-0.81)	0.66 (0.50-0.87)	0.003
Model 2[b]	1.0	0.70 (0.48-1.02)	0.36 (0.17-0.77)	0.64 (0.49-0.85)	0.002
Model 3[c]	1.0	0.75 (0.51-1.10)	0.34 (0.15-0.78)	0.65 (0.49-0.86)	0.003

CVD, cardiovascular diseases; MI, myocardial infarction; IS, ischemic stroke

a Model 1: Derived from Cox proportional-hazards regression, with age as underlying time variable, stratified by age at baseline, and adjusted for sex.

b Model 2: Adjusted for model 1, and further for HDL-cholesterol (continuous), and adiponectin (continuous).

c Model 3: Adjusted for model 2 and further for smoking status (never smoker, former smoker, current smoker < 20 cigarettes per day, current smoker ≥ 20 cigarettes per day), sports activity (< 2 h/wk versus ≥ 2 h/wk), educational attainment (vocational school or less, technical school, university), BMI (continuous), WC (continuous), alcohol consumption (men: =0 g/d, >0 to 12 g/d, >12 to 24 g/d; >24 g/d; women: =0 g/d, >0 to 6 g/d, >6 to 12 g/d; >12 g/d), Hba1c (continuous), SBP (continuous), antihypertensive medication, total cholesterol (continuous) and hs-CRP (continuous).

Supplementary table 1. Associations between rs2943634 polymorphism and cardiovascular diseases intermediate risk phenotypes in the EPIC-Potsdam subcohort assuming a dominant heritage model.

Characteristics	CC (N = 769)	CA + AA (N = 1122)	P_{dom}
Body mass index (kg/m^2) [a]	26.18 ± 0.15	25.91 ± 0.12	0.17
Waist circumference (cm)			
Men	93.95 ± 0.40	94.92 ± 1.13	0.42
Women	80.72 ± 0.49	80.49 ± 0.42	0.71
SBP (mm/Hg)	129.18 ± 0.58	129.04 ± 0.48	0.85
DBP (mm/Hg)	83.66 ± 0.38	83.48 ± 0.31	0.71
Total Cholesterol (mg/dl)	174.8 ± 1.27	171.138 ± 1.05	0.33
HDL-cholesterol (mg/dl)	52.09 ± 0.49	53.56 ± 0.40	0.02
LDL-cholesterol (mg/dl)[b]	109.55 ± 2.04	107.97 ± 2.04	0.54
Triglycerides (mg/dl)[c]	94.80 (88.12-101.98)	85.36 (80.70-90.28)	0.03
Blood glucose (mg/dl)[d, e]	96.80 (92.17-101.54	96.54 (93.43-99.69)	0.94
HbA1c (%) [f]	6.47 (6.42-6.51)	6.42 (6.39-6.46)	0.15
Hs-C-reactive protein (mg/l)	7.59 (6.89-8.36)	7.15 (6.60-7.74)	0.35
Adiponectin (µg/ml)	6.94 (6.72-7.17)	7.40 (7.20-7.60)	0.003
Creatinine (µg g/dl)	0.72 (0.71-0.73)	0.70 (0.69-0.71)	0.001

All analyses were adjusted for age and sex. a, Means and standard error (SEM), all such values. b, LDL-cholesterol, TG and glucose are based on 543 individuals due to missing fasting data. c, Geometric means and 95% CI, all such values. d, Squared means and 95% CI; e, Further adjusted by the sex x rs2943634 interaction term; f, inverse and 95% (CI).

Supplementary table 2. Associations between rs2943634 polymorphism and intermediate risk phenotypes of cardiovascular diseases in the EPIC-Potsdam subcohort assuming a recessive heritage model.

Characteristics	CC+ CA (N = 1647)	AA (N = 244)	P_{rec}
Body mass index (kg/m^2) [a]	26.02 ± 0.10	26.02 ± 0.27	1.00
Waist circumference (cm)			
Men	94.45 ± 0.61	93.81 ± 0.48	0.41
Women	80.71 ± 0.34	79.83 ± 0.85	0.33
SBP (mm/Hg)	129.12 ± 0.40	128.92 ± 0.86	0.86
DBP (mm/Hg)	83.59 ± 0.257	83.26 ± 0.67	0.64
Total Cholesterol (mg/dl)	173.88 ±	173.66 ± 1.04	0.93
HDL-cholesterol (mg/dl)	52.62 ± 0.33	55.28 ± 0.87	0.004
LDL-cholesterol (mg/dl)[b]	109.03 ± 1.33	105.46 ± 3.43	0.33
Triglycerides (mg/dl)[c]	87.93 (83.81-92.25)	94.23 (83.33-106.55)	0.30
Blood glucose (mg/dl)[d, e]	96.83 (94.73-98.95)	95.36 (88.54-102.43)	0.71
HbA1c (%) [f]	6.45 (6.41-6.48)	6.41 (6.33-6.49)	0.40
Hs-C-reactive protein (mg/l)	7.24 (6.77-7.73)	7.93 (6.68-9.42)	0.33
Adiponectin (µg/ml)	7.12 (6.96-7.27)	7.87 (7.43-8.33)	0.001
Creatinine (µg g/dl)	0.71 (0.70-0.72)	0.70 (0.69-0.72)	0.32

All analyses were adjusted for age and sex. a, Means and standard error (SEM), all such values. b, LDL-cholesterol, TG and glucose are based on 543 individuals due to missing fasting data. c, Geometric means and 95% CI, all such values. d, Squared means and 95% CI; e, Further adjusted by the sex x rs2943634 interaction term; f, inverse and 95% (CI).

Supplementary table 3. Hazard ratios (HR) and 95% confidence intervals (95% CI) for the associations between rs2943634 polymorphism and cardiovascular endpoints according to the dominant and recessive heritage models.

	CA vs CC+AA	P dom	AA vs CC+CA	P rec
CVD cases				
Model 1 [a]	0.90 (0.71-1.14)	0.39	0.80 (0.54-1.18)	0.26
Model 2 [b]	0.89 (0.70-1.14)	0.35	0.80 (0.54-1.19)	0.27
Model 3 [c]	0.93 (0.72-1.21)	0.60	0.76 (0.50-1.16)	0.20
MI cases				
Model 1 [a]	1.01 (0.74-1.36)	0.97	1.04 (0.66-1.66)	0.85
Model 2 [b]	1.00 (0.73-1.37)	0.99	1.05 (0.65-1.68)	0.84
Model 3 [c]	1.02 (0.73-1.42)	0.92	1.03 (0.63-1.68)	0.92
IS cases				
Model 1 [a]	0.85 (0.59-1.23)	0.39	0.45 (0.21-0.93)	0.03
Model 2 [b]	0.84 (0.58-1.21)	0.34	0.43 (0.21-0.91)	0.03
Model 3 [c]	0.90 (0.61-1.33)	0.59	0.39 (0.17-0.89)	0.02

CVD, cardiovascular diseases; MI, myocardial infarction; IS, ischemic stroke

a Model 1: Derived from Cox proportional-hazards regression, with age as underlying time variable, stratified by age at baseline, and adjusted for sex. b Model 2: Adjusted for model 1, and further for HDL-cholesterol (continuous), and adiponectin (continuous). c Model 3: Adjusted for model 2 and further for smoking status (never smoker, former smoker, current smoker < 20 cigarettes per day, current smoker ≥ 20 cigarettes per day), sports activity (< 2 h/wk versus ≥ 2 h/wk), educational attainment (vocational school or less, technical school, university), BMI (continuous), WC (continuous), alcohol consumption (men: =0 g/d, >0 to 12 g/d, >12 to 24 g/d; >24 g/d; women: =0 g/d, >0 to 6 g/d, >6 to 12 g/d; >12 g/d), Hba1c (continuous), SBP (continuous), antihypertensive medication, total cholesterol (continuous) and hs-CRP (continuous).

4

Heterogeneity of the Stearoyl-CoA desaturase-1 (SCD1) gene and metabolic risk factors in the EPIC-Potsdam Study

Chapter 4

Heterogeneity of the Stearoyl-CoA desaturase-1 (*SCD1*) gene and metabolic risk factors in the EPIC-Potsdam Study

Maria Arregui [a *], Brian Buijsse [a], Norbert Stefan [b], Dolores Corella [c], Eva Fisher [d], Romina di Giuseppe [a], Oscar Coltell [e], Sven Knüppel [a], Krasimira Aleksandrova [a], Hans-Georg Joost [f], Heiner Boeing [a], Cornelia Weikert [a, g]

a Department of Epidemiology, German Institute of Human Nutrition (DIfE) Potsdam-Rehbruecke, 14558 Nuthetal, Germany.

b Department of Internal Medicine, Division of Endocrinology, Diabetology, Nephrology, Vascular Disease and Clinical Chemistry. Member of the Deutsches Zentrum für Diabetesforschung (DZD), University of Tübingen, Tübingen, Germany.

c Department of Preventive Medicine and CIBER Fisiopatología de la Obesidad y Nutrición. School of Medicine. University of Valencia, Valencia, Spain.

d Administrative Office of the Commission on Genetic Testing, Robert Koch-Institute, Berlin, Germany.

e Department of Computing Languages and Systems, Universitat Jaume I, Castellón, Spain.

f Department of Pharmacology, German Institute of Human Nutrition (DIfE) Potsdam-Rehbruecke, 14558 Nuthetal, Germany.

g Institute for Social Medicine, Epidemiology, and Health Economics, Charité University Medical Center, Berlin, Germany.

PLoS One. 2012;7(11):e48338

Impact factor 2012: 4.092

ABSTRACT

Background: Stearoyl-CoA desaturase-1 (SCD1) is an enzyme involved in lipid metabolism. In mice and humans its activity has been associated with traits of the metabolic syndrome, but also with the prevention of saturated fatty acids accumulation and subsequent inflammation, whereas for liver fat content inconsistent results have been reported. Thus, variants of the gene encoding SCD1 (*SCD1*) could potentially modify metabolic risk factors, but few human studies have addressed this question.

Methods: In a sample of 2157 middle-aged men and women randomly drawn from the Potsdam cohort of the European Prospective Investigation into Cancer and Nutrition, we investigated the impact of 7 *SCD1* tagging-single nucleotide polymorphisms (rs1502593, rs522951, rs11190480, rs3071, rs3793767, rs10883463 and rs508384) and 5 inferred haplotypes with frequency >5% describing 90.9% of the genotype combinations in our population, on triglycerides, body mass index (BMI), waist circumference (WC), glycated haemoglobin (HbA1c), high-sensitivity C-reactive protein (hs-CRP), gamma-glutamyltransferase (GGT), alanine aminotransferase (ALT) and fetuin-A.

Results: No significant associations between any of the SNPs or haplotypes and BMI, WC, fetuin-A and hs-CRP were observed. Associations of rs10883463 with triglycerides, GGT and HbA1c as well as of rs11190480 with ALT activity, were weak and became non-significant after multiple-testing correction. Also associations of the haplotype harbouring the minor allele of rs1502593 with HbA1c levels, the haplotype harbouring the minor alleles of rs11190480 and rs508384 with activity of ALT, and the haplotype harbouring the minor alleles of rs522951, rs10883463 and rs508384 with triglyceride and HbA1C levels and GGT activities did not withstand multiple-testing correction.

Conclusion: These findings suggest that there are no associations between common variants of *SCD1* or its inferred haplotypes and the investigated metabolic risk factors. However, given the results from animal models, heterogeneity of human *SCD1* warrants further investigation, in particular with regard to rare variants.

INTRODUCTION

The human stearoyl-CoA desaturase-1 (*SCD1*) gene maps to chromosome 10q24.31, has 6 exons and is remarkably expressed in adipose tissue and the liver. It encodes the endoplasmatic reticulum enzyme SCD1, which catalyses the conversion of the saturated fatty acids (SFAs) palmitic and stearic, into the monounsaturated fatty acids (MUFAs) palmitoleic and oleic respectively [1]. These MUFAs are the major components of triglycerides, likely due to its production within the environs of the enzyme diacylglycerol acyltransferase (DGAT) [1]. Accordingly, it has been suggested that the increased activity of SCD1 in the liver, could result in an excess assembly and accumulation of triglycerides and subsequent development of hepatic steatosis [2]. Further, the overflow of triglycerides could also be incorporated into very low density lipoprotein (VLDL) particles and transported to adipose tissue and other sites, contributing to the development of obesity [3]. Both these conditions have been associated with insulin resistance [4]. Additionally, SCD1 deficiency has been associated with the reduced expression of fatty acid synthesis genes [5] and the up-regulation of genes involved in fatty acid β-oxidation [5,6]. In fact mice with a natural or a targeted deletion of the *SCD1* gene, have shown to be protected against hypertriglyceridemia [7], hepatic steatosis [8-10], obesity [3,5,8,11,12] and insulin resistance [11,13,14]. Conversely, there is also evidence that by channelling SFA, into triglyceride pools, increased SCD1 activity may prevent from lipoapoptosis [15], steatohepatitis [2] and inflammation [6,15-17]. Therefore, SCD1 seams to convey both, positive and negative roles in the development of metabolic risk factors of cardiovascular diseases.

Despite the strong scientific interest in SCD1, most of the existing knowledge on its function comes from mice models [18]. Some human studies have provided indirect evidence of the role of its activity (approximated as fatty acids product-to-precursor ratios measured in serum, plasma, erythrocytes or adipose tissue), and have proposed that its elevation might be associated with harmful effects such as elevated plasma triglycerides levels [19,20], liver fat [21-23], obesity [19,24], diabetes [25], high-sensitivity C-reactive protein (hs-CRP) levels [26] and even with cardiovascular mortality [27]. Conversely the scarce human studies investigating the tissue-specific activity and expression of SCD1, suggest that elevated SCD1 activity may protect from liver fat accumulation [28,29]. Further, the impact of *SCD1* heterogeneity on metabolic risk factors, has so far only been investigated in four human studies with focus on diabetes and

obesity [30,31], metabolic syndrome (MetS) [32] or inflammation [33]. Thus, a case control study in men and women from the United Kingdom, found no associations between 6 *SCD1* single nucleotide polymorphisms (SNPs) or its inferred haplotypes and diabetes, body mass index (BMI) or waist-to-hip ratio [30]; a cross-sectional study in Swedish elderly men reported that 4 out of 8 *SCD1* tagging SNPs (tag-SNPs) related to decreased BMI and waist circumference (WC), and increased insulin sensitivity [31]; the haplotype consisting of the rare alleles of these SNPs was also associated with decreased WC; a cross-sectional study in Costa Rican middle-aged men and women reported that 1 out of 7 *SCD1* tag-SNPs was associated with an increased prevalence of MetS, and among women, also with elevated systolic blood pressure and fasting blood glucose levels [32]. Also 2 haplotypes carrying the minor allele of this SNP were associated with elevated prevalence of MetS; finally a cross-sectional study in European and Asian young adults found 1 out of 10 tag-SNPs to be associated with CRP levels [33].

In the Potsdam cohort of the European Prospective Investigation into Cancer and Nutrition (EPIC-Potsdam), we investigated the impact of common genetic variation in *SCD1*, captured by means of 7 tag-SNPs and their inferred haplotypes, on the modulation of 8 metabolic risk factors related to the activity of *SCD1*. The investigated traits were plasma triglyceride levels, traits related to obesity (BMI and waist circumference (WC)), glucose metabolism (glycated haemoglobin (HbA1c)) and chronic inflammation (hs-CRP), and, for the first time, crude estimates of the presence of liver fat (the liver enzymes gamma-glutamyltransferase (GGT) and alanine amino transferase (ALT) [34,35]), and fetuin-A, a biomarker which has been associated with fat accumulation in the liver [36-40] as well as with insulin resistance, type 2 diabetes and cardiovascular events [40-42].

METHODS

Ethics Statement

Written informed consent was obtained from all study participants, and approval was given by the Ethics Committee of the Medical Association of the State of Brandenburg, Germany.

Study population

EPIC-Potsdam comprises 27548 individuals (10904 men and 16644 women) from the general population of the Potsdam area in Germany. Men were mainly aged 40-65 and women 35-65 years old at recruitment, which took place between 1994 and 1998 [43]. The baseline examination included a personal interview and a questionnaire on sociodemographic and lifestyle characteristics and prevalent diseases as well as anthropometric measurements[44]. The associations of 7 *SCD1* tag-SNPs and their inferred haplotypes with anthropometric and metabolic markers were investigated in a random sample of 2500 individuals (subcohort) drawn from the participants in the total cohort who had provided blood samples at baseline, following a cross-sectional design. After exclusion of individuals with missing covariates or genotype data, the final study population comprised 2157 participants. Fasting was not required at the time of blood draw, however, 615 participants were in fasting state for at least 8 h.

Laboratory analyses

From all the study participants a 30 mL sample of venous blood was collected, fractionated into serum, plasma, buffy coat and erythrocytes, and stored in liquid nitrogen until the time of analysis. Plasma levels of triglycerides, HbA1c, GGT, ALT, fetuin-A, hs-CRP ,total cholesterol and high density lipoprotein (HDL)-cholesterol were determined with the automatic ADVIA 1650 analyser (Siemens Medical Solutions, Erlangen, Germany) at the Department of Internal Medicine of the University of Tübingen, Germany, in 2007.

SNP selection and genotyping

Seven *SCD1* tag-SNPs (ordered according to chromosomic location: rs1502593, rs522951, rs11190480, rs3071, rs3793767, rs10883463 and rs508384) were identified in the HapMap 22/phaseII CEU population data (Utah residents with ancestry from northern and western Europe) [45] using stringent criteria (minor allele frequency (MAF) >0.05 and pairwise r2 ≥0.8) by means of the Tagger software [46] implemented in the version 4.2 of Haploview [47]. The tagged region comprised the coding region of *SCD1* as well as a 4.1 Kb upstream (promoter) and 4.3 Kb downstream (3' untranslated) region of the gene. Six SNPs were located in intronic sites and one in the 3′ untranslated region (rs508384). Genotyping of whole genome amplified DNA

samples was performed with a 7900HT Sequence Detection System with TaqMan assays (Applied Biosystems, Foster City, CA, USA) at the Max Delbrück Centre for Molecular Medicine, Berlin, Germany, in 2009. The average genotyping success rate in the 7 SNPs was >98%.

Statistical analyses

Normality of variables was tested by estimating their skewness and kurtosis, by comparing their means and median values and by plotting their distributions in histograms. To better reach normality of their distributions, triglycerides, GGT, ALT and hs-CRP were natural log-transformed and HbA1c inverse-transformed, and were used like that in all analyses. Hardy–Weinberg equilibrium (HWE) of the SNPs was tested using the $\chi 2$ test. Linkage disequilibrium between SNPs was assessed with the r2 measure using Haploview 4.2 [47]. Each SNP was coded as 0, 1 and 2 according to the number of minor alleles a participant carried. Analysis of covariance considering the additive, dominant and recessive genetic models was used to assess the associations between the SNPs (independent variables) and triglycerides, BMI, WC, HbA1c, GGT, ALT, fetuin-A and hs-CRP (dependent variables). Haplotypes were constructed to test whether multiple genetic variants of *SCD1* or a possible unobserved risk variant captured by the haplotypes, modulated the investigated traits. Haplotype frequencies were estimated based on the observed unphased genotypes by the expectation-maximization algorithm [48]. The effects of a particular haplotype load (0, 1 or 2 copies) were tested also by means of a regression-based analysis (ANCOVA) as suggested by Zaykin et al. [49]. Analyses were restricted to participants with a probability of 49-50% to carry one copy of the haplotype, or 100% probability to carry either none or two copies. Only haplotypes with frequencies >5% were considered. The additive, dominant and recessive models were examined. Analyses for triglycerides were performed only in participants who were fasting at the time of blood draw. Data are reported as means and standard errors, geometric means and 95% confidence intervals (CI) or inverse and 95% CI as appropriate. Regression coefficients and standard errors (SE) were also estimated. All analyses were adjusted for age and sex. Further (mutual) adjustment for known cardiovascular risk factors including smoking status (never smoker, former smoker, current smoker <20 cigarettes per day, current smoker ≥20 cigarettes per day), sports activity (<2 h/wk versus ≥2 h/wk), educational attainment (vocational school or less, technical school, university), BMI (continuous), WC

(continuous), alcohol consumption (men: =0 g/d, >0 to 12 g/d, >12 to 24 g/d; >24 g/d; women: =0 g/d, >0 to 6 g/d, >6 to 12 g/d; >12 g/d), prevalent diabetes, prevalent hypertension, total cholesterol, HDL cholesterol and hs-CRP) were also explored for the tag-SNPs. Effect modification by sex was evaluated by modeling the cross product term sex times genotype or haplotype, along with main effects (in the age-adjusted general linear model). All statistical analyses were performed using SAS Enterprise Guide 4.3 (SAS Institute Inc., Cary, NC, USA).

Power calculations were performed with Quanto [50] considering an additive model, a desirable power of 80%, a two-sided α of 0.05, the means and standard deviations of the traits in the subcohort (non-normal variables were transformed to normality) and the genotype frequencies of the least (rs10883463, MAF = 8%) and most (rs522951, MAF = 46%) common tag-SNPs. The detectable differences in our study ranged between 0.1 and 0.2 SD for BMI, WC, GGT, ALT, fetuin-A, hs-CRP and HbA1c and between 0.2 and 0.3 SD for triglycerides. Conservative Bonferroni correction for multiple comparisons was performed (P Bonferroni = α/(n individual hypothesis tested) = 0.05/((7 SNPs + 5 haplotypes) x 3 genetic models per SNP x 8 traits investigated). The corrected significance threshold was P Bonferroni = 0.0002.

RESULTS

Characteristics of the study population

Demographic, lifestyle, clinical, biochemical and genetic characteristics of the study population are given in Table 1, both for the subcohort and separately for men and women. Men (38%) were older than women due to sampling strategy. After adjusting for age, they also showed to smoke and drink more and to take more often antidiabetic medication. Further their BMI, WC and their activities of GGT and ALT were higher. They were more likely to be higher educated and also had lower hs-CRP levels than women. The genotype frequency of all *SCD1* tag-SNPs followed HWE (P >0.05), their allele frequencies were comparable to those observed in HapMap 22/phaseII CEU population data [45] and did not differ among sexes. Table S1 presents further information regarding genotype and allelic frequencies of the tag-SNPs across the EPIC-Potsdam subcohort and separately for men and women.

Linkage disequilibrium values (r2) between each pair of tag-SNPs ranged from 0 to 0.66. Five haplotypes inferred from the genotyped SNPs, with frequency >5% described 90.9% of the genotype combinations in our population: A-B-A-A-B-A-A (34.1%); B-A-A-B-A-A-A (29.8%), B-A-A-A-A-A-A (11.0%), A-A-B-A-A-A-B (8.6%) and A-B-A-A-A-B-B (7.4%) (Haplotypes are composed of variants rs1502593 (C>T), rs522951 (G>C), rs11190480 (A>G), rs3071 (T>G), rs3793767 (T>C), rs10883463 (T>C), rs508384 (C>A) in that order; A indicates common allele, B indicates rare allele).

Associations between *SCD1* tag-SNPs and inferred haplotypes and the investigated traits

No significant effect modifications by sex were found for any of the tag-SNPs or inferred haplotypes on the investigated traits, thus results are presented combined for men and women. Table 2 shows results for age- and sex-adjusted mean values of the 8 investigated metabolic traits for each of the 7 *SCD1* tag-SNPs. No significant associations were found between rs1502593, rs522951, rs3071, rs3793767 or rs508384 and any of the investigated traits. Also no significant associations were found between any of the investigated SNPs and fetuin-A, BMI, WC or hs-CRP. However, carriers of the rs11190480 rare allele, presented slightly lower activities of ALT in a dominant fashion (19.91 vs. 21.07 U/L, P = 0.03). Carriers of the rs10883463 rare allele, showed higher triglycerides (160.97 vs. 94.02 mg/dL, P = 0.03) and lower HbA1c levels (6.10 vs. 6.49 %, P = 0.03) in a recessive fashion, and slightly higher activities of GGT (22.22 vs. 19.88 U/L, P = 0.02) in a dominant fashion. Results were weak in precision, as shown by the wide confidence intervals, and after applying the Bonferroni correction for the multiple hypothesis tested (P Bonferroni = 0.0002), none of them remained significant. After further (mutual) adjustment of the statistical models for known cardiovascular risk factors, results remained essentially similar (table S2). Also they were not modified by adjustment for fasting status or exclusion of participants taking lipid-lowering or antidiabetic medication.

Table 3 shows results for age- and sex-adjusted mean values of the investigated metabolic traits for each of the 5 *SCD1* haplotypes. No associations between the two most common haplotypes and the investigated variables became apparent. Homozygotes for haplotype B-A-A-A-A-A-A harbouring the minor allele of the SNP rs1502593 presented lower HbA1c levels (6.03 vs. 6.49 %, P = 0.003). Carriers of haplotype A-A-B-A-A-A-B harbouring the minor alleles of the SNPs

rs11190480 and rs508384 exhibited slightly lower activities of ALT in a dominant fashion (19.85 vs. 21.08 U/L, P = 0.02). Carriers of haplotype A-B-A-A-A-B-B harbouring the minor alleles of the SNPs rs522951, rs10883463 and rs508384 showed higher triglyceride values 179.77 vs. 94.08 mg/dL and lower HbA1C levels (6.09 vs. 6.49 %, P = 0.03) in a recessive fashion, and higher GGT activities (22.24 vs. 19.93 U/L, P = 0.03) in a dominant fashion. However, after correction for multiple testing, also none of these results remained significant. Table S3 summarizes all of these results in the form of age- and sex-adjusted regression coefficients.

DISCUSSION

In the present study of a middle-aged sample of German men and women, we evaluated the impact of 7 *SCD1* tag-SNPs and 5 inferred haplotypes on MetS related traits on suggested crude estimates of the presence of liver fat and on inflammation. Our study is, so far, the largest performed in a European population, and also the first to report association results between *SCD1* genetic variants and liver parameters. We hypothesized that any functional variant affecting the activity of *SCD1* would possibly result in the modulation of one or more of the traits. At most, we found some associations weak in magnitude, precision, and statistical significance, which after conservative Bonferroni-correction for the number of traits, SNPs and haplotypes tested, did not remain significant, thus being suggestive of chance findings.

Four previous studies in humans have investigated the association of *SCD1* polymorphisms with different metabolic traits [30-33]. In a UK case-control study of 608 cases and 600 controls, Liew et al. [30] reported upon the association of 6 *SCD1* SNPs with type 2 diabetes, BMI and waist-to-hip ratio. Three of these SNPs were in common or highly linked to SNPs of our study: rs670213 (a good proxy for rs522951, r2 = 0.87 according to HapMap data for Caucasians of European origin [45], rs3071 and rs11598233 (a perfect proxy for rs3793767, r2 = 1 [45]). Consistent with our results, they also reported no significant associations. Warensjö et al. [31] investigated associations of 8 *SCD1* tag-SNPs with obesity and insulin sensitivity in 1143 Swedish elderly men. Five of these SNPs, were in common or highly linked to SNPs of our study, rs3870747 (linked with rs1119040, r2= 0.94 [45]), rs3071, rs3793767, rs10883463 and rs508384. In line with our results, they found a tendency of rs10883463 carriers towards increased insulin sensitivity, and no significant associations for rs3870747 and rs3793767. In contrast, they

reported lower WC in carriers of rs10883463, lower WC and higher insulin sensitivity in homozygotes for rs508384 rare allele and lower insulin sensitivity in heterozygotes for rs3071. This last association was also inconsistent with the results from Liew et al. [30]. Compared with our study population, the study of Warensjö et al. [31] was smaller and included only elderly men. However, we found no significant sex interaction for any of the SNPs with any of the studied phenotypes, and further all our association analyses were adjusted for age, what makes the overall evidence for an association of *SCD1* genetic variability with WC and insulin sensitivity less consistent. Recently, Gong et al. [32] analysed the association of 7 *SCD1* tag-SNPs with MetS prevalence in 2152 Costa Rican adult men and women. Only one SNP, rs1502593, was found to be associated with MetS. This SNP was also analysed in our population. In line with our results, they did not observe significant associations between rs1502593 and triglyceride levels or WC. They did point out a borderline association of rs1502593 with elevated fasting blood glucose levels among women. In our population we found no significant associations for this SNP. Homozygotes for the haplotype harbouring its minor allele showed lower HbA1c levels, measure of the average plasma glucose levels over prolonged periods of time, but this association did not withhold after multiple-testing correction. We cannot discard the possibility that the different ethnic origin may explain the different findings. Finally, Stryjecki et al. [33] examined the relationships between 10 *SCD1* tag-SNPs and CRP levels in 279 European and 249 Asian young adults. Only one SNP located 9 Kb upstream *SCD1*, and thus not included in our study, was associated with CRP levels, and only among females of both groups.

Limitations of our study should be mentioned. It is possible that, due to sample size limitations we were not able to detect minor contributions of the alleles. This limitation was stronger in the case of the association analyses for triglycerides levels, as fasting data was available only for about one third of the population. Thus while some of our results could represent a replication for certain relationships inspected in the four association studies that precede ours [30-33], to evaluate the outcome of our work, further studies in independent cohorts analysing the same SNPs, or those in perfect linkage disequilibrium, are necessary [51]. Also, we cannot exclude the possibility that a rare causal variant exists within the typed region but was not picked up by the chosen markers.

In summary, our findings suggest that common variants of *SCD1* do not modulate the investigated metabolic factors in this European population. However, given its biological relevance, the still scarce number of studies available and the inconsistency of their results, genetic heterogeneity of human *SCD1* in relation to impaired metabolism rewards further investigation in independent study populations, in particular with regard to rare variants of *SCD1*.

SUPPORTING INFORMATION LEGENDS

Table S1 Genotype and allelic frequencies of the *SCD1* tag-SNPs across the EPIC-Potsdam subcohort and separately for men and women.

Table S2 Association analysis between the 7 *SCD1* tag-SNPs and the 8 investigated metabolic traits in the EPIC-Potsdam Study, (mutually) adjusted for known cardiovascular risk factors. Table S3 β regression coefficients for the age- and sex-adjusted association analysis between the *SCD1* tag-SNPs and inferred haplotypes and the 8 investigated metabolic traits in the EPIC-Potsdam study.

ACKNOWLEDGEMENTS

The authors are grateful to Anna Bury and Sabina Herbert who were involved in the biochemical analyses. We are indebted to Ellen Kohlsdorf for data management. We thank Kathrin Saar and Norbert Hübner for support and discussion.

REFERENCES

1. Paton CM, Ntambi JM (2009). Biochemical and physiological function of stearoyl-CoA desaturase. Am J Physiol Endocrinol Metab 297: E28-37.

2. Li ZZ, Berk M, McIntyre TM, Feldstein AE (2009). Hepatic lipid partitioning and liver damage in nonalcoholic fatty liver disease: role of stearoyl-CoA desaturase. J Biol Chem 284: 5637-5644.

3. Cohen P, Miyazaki M, Socci ND, Hagge-Greenberg A, Liedtke W, et al. (2002). Role for stearoyl-CoA desaturase-1 in leptin-mediated weight loss. Science 297: 240-243.

4. Petersen KF, Shulman GI (2006). Etiology of insulin resistance. Am J Med 119: S10-16.

5. Ntambi JM, Miyazaki M, Stoehr JP, Lan H, Kendziorski CM, et al. (2002). Loss of stearoyl-CoA desaturase-1 function protects mice against adiposity. Proc Natl Acad Sci U S A 99: 11482-11486.

6. MacDonald ML, van Eck M, Hildebrand RB, Wong BW, Bissada N, et al. (2009). Despite antiatherogenic metabolic characteristics, SCD1-deficient mice have increased inflammation and atherosclerosis. Arterioscler Thromb Vasc Biol 29: 341-347.

7. Miyazaki M, Kim YC, Gray-Keller MP, Attie AD, Ntambi JM (2000). The biosynthesis of hepatic cholesterol esters and triglycerides is impaired in mice with a disruption of the gene for stearoyl-CoA desaturase 1. J Biol Chem 275: 30132-30138.

8. Miyazaki M, Dobrzyn A, Sampath H, Lee SH, Man WC, et al. (2004). Reduced adiposity and liver steatosis by stearoyl-CoA desaturase deficiency are independent of peroxisome proliferator-activated receptor-alpha. J Biol Chem 279: 35017-35024.

9. Miyazaki M, Dobrzyn A, Man WC, Chu K, Sampath H, et al. (2004). Stearoyl-CoA desaturase 1 gene expression is necessary for fructose-mediated induction of lipogenic gene expression by sterol regulatory element-binding protein-1c-dependent and -independent mechanisms. J Biol Chem 279: 25164-25171.

10. Miyazaki M, Flowers MT, Sampath H, Chu K, Otzelberger C, et al. (2007). Hepatic stearoyl-CoA desaturase-1 deficiency protects mice from carbohydrate-induced adiposity and hepatic steatosis. Cell Metab 6: 484-496.

11. MacDonald ML, Singaraja RR, Bissada N, Ruddle P, Watts R, et al. (2008). Absence of stearoyl-CoA desaturase-1 ameliorates features of the metabolic syndrome in LDLR-deficient mice. J Lipid Res 49: 217-229.

12. Sampath H, Flowers MT, Liu X, Paton CM, Sullivan R, et al. (2009). Skin-specific deletion of stearoyl-CoA desaturase-1 alters skin lipid composition and protects mice from high fat diet-induced obesity. J Biol Chem 284: 19961-19973.

13. Gutierrez-Juarez R, Pocai A, Mulas C, Ono H, Bhanot S, et al. (2006). Critical role of stearoyl-CoA desaturase-1 (SCD1) in the onset of diet-induced hepatic insulin resistance. J Clin Invest 116: 1686-1695.

14. Flowers JB, Rabaglia ME, Schueler KL, Flowers MT, Lan H, et al. (2007). Loss of stearoyl-CoA desaturase-1 improves insulin sensitivity in lean mice but worsens diabetes in leptin-deficient obese mice. Diabetes 56: 1228-1239.

15. Listenberger LL, Han X, Lewis SE, Cases S, Farese RV Jr, et al. (2003). Triglyceride accumulation protects against fatty acid-induced lipotoxicity. Proc Natl Acad Sci U S A 100: 3077-3082.

16. Liu X, Strable MS, Ntambi JM (2011). Stearoyl CoA desaturase 1: role in cellular inflammation and stress. Adv Nutr 2: 15-22.

17. Brown JM, Chung S, Sawyer JK, Degirolamo C, Alger HM, et al. (2008). Inhibition of stearoyl-coenzyme A desaturase 1 dissociates insulin resistance and obesity from atherosclerosis. Circulation 118: 1467-1475.

18. Sampath H, Ntambi JM (2011). The role of stearoyl-CoA desaturase in obesity, insulin resistance, and inflammation. Ann N Y Acad Sci 1243: 47-53.

19. Zhou YE, Egeland GM, Meltzer SJ, Kubow S (2009). The association of desaturase 9 and plasma fatty acid composition with insulin resistance-associated factors in female adolescents. Metabolism 58: 158-166.

20. Petersson H, Basu S, Cederholm T, Riserus U (2008). Serum fatty acid composition and indices of stearoyl-CoA desaturase activity are associated with systemic inflammation: longitudinal analyses in middle-aged men. Br J Nutr 99: 1186-1189.

21. Petersson H, Arnlov J, Zethelius B, Riserus U (2010). Serum fatty acid composition and insulin resistance are independently associated with liver fat markers in elderly men. Diabetes Res Clin Pract 87: 379-384.

22. Kotronen A, Seppanen-Laakso T, Westerbacka J, Kiviluoto T, Arola J, et al. (2009). Hepatic stearoyl-CoA desaturase (SCD)-1 activity and diacylglycerol but not ceramide concentrations are increased in the nonalcoholic human fatty liver. Diabetes 58: 203-208.

23. Tomita K, Teratani T, Yokoyama H, Suzuki T, Irie R, et al. (2011). Plasma free myristic acid proportion is a predictor of nonalcoholic steatohepatitis. Dig Dis Sci 56: 3045-3052.

24. Gong J, Campos H, McGarvey S, Wu Z, Goldberg R, et al. (2011). Adipose tissue palmitoleic acid and obesity in humans: does it behave as a lipokine? Am J Clin Nutr 93: 186-191.

25. Kroger J, Zietemann V, Enzenbach C, Weikert C, Jansen EH, et al. (2011). Erythrocyte membrane phospholipid fatty acids, desaturase activity, and dietary fatty acids in relation to risk of type 2 diabetes in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Am J Clin Nutr 93: 127-142.

26. Petersson H, Lind L, Hulthe J, Elmgren A, Cederholm T, et al. (2009). Relationships between serum fatty acid composition and multiple markers of inflammation and endothelial function in an elderly population. Atherosclerosis 203: 298-303.

27. Warensjö E, Sundstrom J, Vessby B, Cederholm T, Riserus U (2008). Markers of dietary fat quality and fatty acid desaturation as predictors of total and cardiovascular mortality: a population-based prospective study. Am J Clin Nutr 88: 203-209.

28. Stefan N, Peter A, Cegan A, Staiger H, Machann J, et al. (2008). Low hepatic stearoyl-CoA desaturase 1 activity is associated with fatty liver and insulin resistance in obese humans. Diabetologia 51: 648-656.

29. Peter A, Cegan A, Wagner S, Elcnerova M, Konigsrainer A, et al. (2011). Relationships between hepatic stearoyl-CoA desaturase-1 activity and mRNA expression with liver fat content in humans. Am J Physiol Endocrinol Metab 300: E321-326.

30. Liew CF, Groves CJ, Wiltshire S, Zeggini E, Frayling TM, et al. (2004). Analysis of the contribution to type 2 diabetes susceptibility of sequence variation in the gene encoding stearoyl-CoA desaturase, a key regulator of lipid and carbohydrate metabolism. Diabetologia 47: 2168-2175.

31. Warensjö E, Ingelsson E, Lundmark P, Lannfelt L, Syvanen AC, et al. (2007). Polymorphisms in the SCD1 gene: associations with body fat distribution and insulin sensitivity. Obesity (Silver Spring) 15: 1732-1740.

32. Gong J, Campos H, McGarvey S, Wu Z, Goldberg R, et al. (2011). Genetic variation in stearoyl-CoA desaturase 1 is associated with metabolic syndrome prevalence in Costa Rican adults. J Nutr 141: 2211-2218.

33. Stryjecki C, Roke K, Clarke S, Nielsen D, Badawi A, et al. (2011). Enzymatic activity and genetic variation in SCD1 modulate the relationship between fatty acids and inflammation. Mol Genet Metab.

34. Stefan N, Kantartzis K, Haring HU (2008). Causes and metabolic consequences of Fatty liver. Endocr Rev 29: 939-960.

35. Schindhelm RK, Diamant M, Dekker JM, Tushuizen ME, Teerlink T, et al. (2006). Alanine aminotransferase as a marker of non-alcoholic fatty liver disease in relation to type 2 diabetes mellitus and cardiovascular disease. Diabetes Metab Res Rev 22: 437-443.

36. Yilmaz Y, Yonal O, Kurt R, Ari F, Oral AY, et al. (2010). Serum fetuin A/alpha2HS-glycoprotein levels in patients with non-alcoholic fatty liver disease: relation with liver fibrosis. Ann Clin Biochem 47: 549-553.

37. Haukeland JW, Dahl TB, Yndestad A, Gladhaug IP, Loberg EM, et al. (2012). Fetuin A in nonalcoholic fatty liver disease: in vivo and in vitro studies. Eur J Endocrinol 166: 503-510.

38. Mussig K, Staiger H, Machicao F, Machann J, Hennige AM, et al. (2009). AHSG gene variation is not associated with regional body fat distribution--a magnetic resonance study. Exp Clin Endocrinol Diabetes 117: 432-437.

39. Reinehr T, Roth CL (2008). Fetuin-A and its relation to metabolic syndrome and fatty liver disease in obese children before and after weight loss. J Clin Endocrinol Metab 93: 4479-4485.

40. Stefan N, Hennige AM, Staiger H, Machann J, Schick F, et al. (2006). Alpha2-Heremans-Schmid glycoprotein/fetuin-A is associated with insulin resistance and fat accumulation in the liver in humans. Diabetes Care 29: 853-857.

41. Stefan N, Fritsche A, Weikert C, Boeing H, Joost HG, et al. (2008). Plasma fetuin-A levels and the risk of type 2 diabetes. Diabetes 57: 2762-2767.

42. Weikert C, Stefan N, Schulze MB, Pischon T, Berger K, et al. (2008). Plasma fetuin-a levels and the risk of myocardial infarction and ischemic stroke. Circulation 118: 2555-2562.

43. Boeing H, Korfmann A, Bergmann MM (1999). Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. Ann Nutr Metab 43: 205-215.

44. Kroke A, Bergmann MM, Lotze G, Jeckel A, Klipstein-Grobusch K, et al. (1999). Measures of quality control in the German component of the EPIC study. European Prospective Investigation into Cancer and Nutrition. Ann Nutr Metab 43: 216-224.

45. Consortium IH (2005). A haplotype map of the human genome. Nature 437: 1299-1320.

46. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, et al. (2005). Efficiency and power in genetic association studies. Nat Genet 37: 1217-1223.

47. Barrett JC, Fry B, Maller J, Daly MJ (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263-265.

48. Excoffier L, Slatkin M (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. Mol Biol Evol 12: 921-927.

49. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, et al. (2002). Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 53: 79-91.

50. Gauderman WJ MJ (2006). QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. http://hydauscedu/gxe. Last accessed: August 2011.

51. Studies N-NWGoRiA, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, et al. (2007). Replicating genotype-phenotype associations. Nature 447: 655-660.

Table 1. Baseline characteristics of the EPIC-Potsdam subcohort and separately for men and women.

	Subcohort (N = 2157)	**Men** (N = 819; 38%)	**Women** (N = 1338; 62%)	**P value** [a]
Age (year)	50.3±9.0	52.4 (51.8-53.0)	49.0 (48.5-49.4)	<0.0001
Educational attainment (%):				
Vocational school or less	37.4	30.8	41.5	<0.0001
Technical school	24.9	15.9	30.4	<0.0001
University degree	37.7	53.4	28.1	<0.0001
Physical activity (hours/week):				
< 2	75.9	74.7	76.7	0.3
≥ 2	24.1	25.3	23.3	0.3
Smoking status (%):				
Current ≥20 cigarettes/d	6.3	11.3	3.2	<0.0001
Current <20 cigarettes/d	14.9	16.7	13.8	0.07
Former ≤5 years	7.5	9.5	6.3	0.006
Former >5 years	24.4	35.3	17.7	<0.0001
Never	46.9	27.2	59.0	<0.0001
Alcohol intake:				
Men: = 0 g/d; women: = 0 g/d	0.05	0.1	0.0	0.3
Men: > 0-12 g/d; women: >0-6 g/d	47.2	36.4	53.9	<0.0001
Men: >12-24 g/d; women: >6-12 g/d	24.7	26.1	23.9	0.3
Men: > 24 g/d; women: > 12 g/d	28.0	37.5	22.2	<0.0001
Use of medication (%):				
Antidiabetic	2.6	3.5	2.0	0.03
Antihypertensive	19.3	19.6	19.2	0.8
Lipid lowering medication	5.1	5.7	4.7	0.3
Body mass index (kg/m^2)	26.1±4.3	26.6 (26.3-26.8)	25.8 (25.6-26.1)	0.0002
Waist circumference (cm):	85.8±12.9	93.4 (92.6-94.1)	81.2 (80.6-81.7)	<0.0001
Triglycerides (mg/dL) [b]	90.2 (64.9-126.5)	109.3 (102.1-117.0)	82.1 (77.6-86.7)	<0.0001
Gamma-glutamyltransferase (U/L)	16.8 (11.0-30.8)	28.3 (26.8-29.9)	14.4 (13.8-15.0)	<0.0001
Glutamic-pyruvate transaminase (U/L)	18.7 (14.3-26.4)	25.9(25.1-26.7)	16.8 (16.4-17.3)	<0.0001
Fetuin-A (mg/dL) [c]	0.25±0.06	0.25 (0.24-0.25)	0.25 0.25-0.26)	0.03
Glycated haemoglobin (%)	6.4 (6.1-6.8)	6.5 (6.4-6.6)	6.4 (6.4-6.5)	0.002
hs-C-reactive protein (mg/L)	0.8 (0.2-2.1)	0.6 (0.6-0.7)	0.8 (0.6-0.7)	<0.0001
Minor allele frequency (%) [d]				
rs1502593 (C>T)	44	45	43	0.2
rs522951 (G>C)	46	46	47	0.6
rs11190480 (A>G)	9	9	9	0.6
rs3071 (T>G)	35	35	34	0.8
rs3793767 (T>C)	38	36	39	0.2
rs10883463 (T>C)	8	8	8	0.7
rs508384 (C>A)	17	17	17	1.0

Subcohort: Mean ± SD, %, or median (25th percentile; 75th percentile), all such values. **Men and women**: mean and 95% confidence interval (CI) or %. Results obtained using analysis of covariance, all variables other than age are adjusted for age. **a** P value for the difference between men and women. **b**, based on the 615 participants fasting at blood draw. **c**, based on 2077 participants due to missing biomarker data. **d** Alleles given in brackets (most >less frequent allele).

Table 2. Age- and sex-adjusted association analyses between the 7 *SCD1* tag-SNPs and the 8 investigated metabolic traits: the EPIC-Potsdam Study.

	Triglycerides [a, b] (mg/dL)	BMI [c] (kg/m^2)	WC [c] (cm)	HbA1c [d] (%)	GGT [b] (U/L)	ALT [b] (U/L)	Fetuin-A [c, e] (mg/dL)	hs-CRP [b] (mg/L)
rs1502593 (n)								
0 (679)	96.45 (89.23-104.26)	26.22±0.16	87.47±0.41	6.48 (6.43-6.53)	20.13 (18.97-21.35)	20.52 (19.81-21.25)	0.25±0.002	0.73 (0.66-0.81)
1 (1072)	92.49 (86.98-98.34)	26.25±0.13	87.36±0.33	6.50 (6.46-6.55)	20.24 (19.31-21.21)	20.85 (20.28-21.44)	0.25±0.002	0.72 (0.67-0.79)
2 (406)	96.46 (87.67-106.14)	26.00±0.21	86.63±0.53	6.45 (6.39-6.52)	20.27 (18.79-21.87)	21.52 (20.57-22.50)	0.26±0.003	0.73 (0.63-0.83)
P_{add}	0.89	0.47	0.25	0.63	0.87	0.11	0.19	0.88
P_{dom}	0.53	0.83	0.53	0.80	0.87	0.24	0.75	0.86
P_{rec}	0.63	0.30	0.18	0.25	0.93	0.14	0.06	0.96
rs522951 (n)								
0 (607)	94.06 (86.85-101.88)	26.03±0.17	86.79±0.44	6.47 (6.41-6.52)	19.93 (18.73-21.21)	20.96 (20.20-21.75)	0.25±0.002	0.72 (0.64-0.80)
1 (1095)	95.22 (89.54-101.26)	26.23±0.13	87.26±0.32	6.50 (6.46-6.54)	20.46 (19.53-21.43)	20.88 (20.31-21.46)	0.25±0.002	0.73 (0.67-0.79)
2 (455)	93.32 (85.14-102.30)	26.31±0.20	87.87±0.50	6.49 (6.43-6.55)	19.98 (18.59-21.47)	20.73 (19.87-21.64)	0.25±0.003	0.73 (0.64-0.83)
P_{add}	0.93	0.28	0.10	0.50	0.90	0.70	0.46	0.82
P_{dom}	0.90	0.27	0.20	0.35	0.60	0.78	0.20	0.80
P_{rec}	0.77	0.52	0.17	0.90	0.72	0.73	0.87	0.91
rs11190480 (n)								
0 (1787)	93.97 (89.60-98.56)	26.23±0.10	87.38±0.26	6.49 (6.46-5.52)	20.30 (19.57-21.06)	21.08 (20.62-21.54)	0.25±0.001	0.73 (0.68-0.78)
1 (357)	96.89 (87.11-107.77)	26.05±0.22	86.66±0.57	6.46 (6.39-6.53)	19.93 (18.39-21.61)	19.85 (18.92-20.82)	0.25±0.003	0.72 (0.63-0.83)
2 (13)	114.86 (39.83-331.24)	25.60±1.16	86.71±2.95	6.66 (6.29-7.06)	15.46 (10.14-23.57)	21.67 (16.88-27.81)	0.24±0.015	0.59 (0.28-1.25)
P_{add}	0.56	0.39	0.25	0.63	0.41	0.05	0.35	0.78
P_{dom}	0.58	0.42	0.24	0.49	0.54	0.03	0.36	0.85
P_{rec}	0.72	0.61	0.85	0.38	0.21	0.77	0.73	0.59
rs3071 (n)								
0 (944)	95.09 (89.07-101.53)	26.36±0.14	87.65±0.35	6.47 (6.43-6.52)	20.64 (19.63-21.70)	20.75 (20.15-21.38)	0.25±0.002	0.77(0.71-0.84)
1 (936)	94.84 (88.81-101.29)	25.96±0.14	86.79±0.35	6.50 (6.45-6.54)	19.76 (18.79-20.78)	20.86 (20.25-21.49)	0.25±0.002	0.68(0.63-0.75)
2 (277)	91.32 (81.28-102.61)	26.42±0.25	87.49±0.64	6.52 (6.44-6.60)	20.30 (18.53-22.25)	21.30 (20.17-22.48)	0.25±0.003	0.73(0.63-0.86)
P_{add}	0.62	0.47	0.36	0.28	0.46	0.46	0.92	0.22
P_{dom}	0.79	0.10	0.13	0.32	0.27	0.62	0.86	0.08
P_{rec}	0.54	0.34	0.69	0.46	0.92	0.43	0.64	0.90
rs3793767 (n)								
0 (845)	95.94 (89.64-102.68)	26.17±0.15	87.21±0.37	6.47 (6.42-6.51)	20.50 (19.45-21.61)	21.05 (20.41-21.72)	0.25±0.002	0.71 (0.65-0.78)
1 (998)	94.63 (88.81-100.83)	26.15±0.13	87.15±0.34	6.51 (6.47-6.55)	20.40 (19.43-21.41)	20.83 (20.23-21.44)	0.25±0.002	0.74 (0.68-0.81)
2 (314)	89.68 (79.88-100.67)	26.41±0.24	87.74±0.61	6.48 (6.41-6.56)	18.86 (17.30-20.57)	20.51 (19.49-21.59)	0.25±0.003	0.73 (0.62-0.84)

P_{add}	0.37	0.50	0.57	0.50	0.17	0.38	0.97	0.65
P_{dom}	0.56	0.81	0.86	0.26	0.49	0.48	0.50	0.50
P_{rec}	0.34	0.33	0.38	0.81	0.09	0.47	0.32	0.97
rs10883463 (n)								
0 (1840)	94.64 (90.34-99.15)	26.18±0.10	87.21±0.25	6.49 (6.46-6.52)	19.88 (19.17-20.61)	20.78 (20.34-21.23)	0.25±0.001	0.73 (0.68-0.78)
1 (304)	89.89 (79.74-101.33)	26.24±0.24	87.39±0.61	6.49 (6.41-6.56)	22.43 (20.56-24.48)	21.35 (20.27-22.48)	0.25±0.003	0.72 (0.62-0.85)
2 (13)	160.99 (100.50-257.90)	27.49±1.16	90.77±2.95	6.10 (5.80-6.44)	17.85 (11.72-27.19)	22.48(17.52-28.85)	0.23±0.017	0.61 (0.29-1.28)
P_{add}	0.76	0.50	0.48	0.36	0.04	0.27	0.85	0.81
P_{dom}	0.80	0.64	0.63	0.62	0.02	0.30	0.99	0.88
P_{rec}	0.03	0.26	0.23	0.03	0.56	0.56	0.32	0.63
rs508384 (n)								
0 (1489)	93.94 (89.21-98.93)	26.18±0.11	87.27±0.28	6.50 (6.46-6.53)	20.01 (19.22-20.83)	21.01 (20.51-21.51)	0.25±0.002	0.73 (0.68-0.78)
1 (610)	93.61 (86.21-101.65)	26.24±0.17	87.16±0.43	6.47 (6.42-6.53)	20.66 (19.42-21.98)	20.55 (19.81-21.31)	0.25±0.002	0.74 (0.66-0.82)
2 (58)	119.44 (92.43-154.35)	26.18±0.55	87.90±1.40	6.44 (6.27-6.62)	20.67 (16.92-25.25)	20.85 (18.52-23.48)	0.25±0.008	0.64 (0.45-0.91)
P_{add}	0.36	0.81	0.96	0.36	0.40	0.39	0.27	0.86
P_{dom}	0.69	0.78	0.92	0.39	0.38	0.33	0.31	0.95
P_{rec}	0.07	0.98	0.64	0.58	0.82	0.99	0.49	0.46

Each SNP is coded as 0, 1 and 2 according to the number of minor alleles a participant carries. a, based on the 615 participants fasting at blood draw. b, geometric means and (95% CI); c, means and standard error; d, inverse and (95% CI), e, based on 2077 participants due to missing biomarker data. All the reported significance levels are nominal P values and are not adjusted for multiple comparisons. P_{add}, P for trend or P for the additive model; P_{dom}, P value for the dominant model. P_{rec}, P value for the recessive model.

Table 3. Age- and sex-adjusted association analyses between the 5 *SCD1* inferred haplotypes with frequency >5% and the 8 investigated metabolic traits in the EPIC-Potsdam study.

Haplotype (%)	**N (%)**	Triglycerides [a, b] (mg/dL)	BMI [c] (kg/m^2)	WC [c] (cm)	HbA1c [d] (%)	GGT [b] (U/L)	ALT [b] (U/L)	**Fetuin-A [c, e] (mg/dL)**	hs-CRP [b] (mg/L)
A-B-A-A-B-A-A (34.1)									
0 copies	919 (43.13)	95.76 (89.70-102.24)	26.18±0.14	87.20±0.35	6.47 (6.43-6.52)	20.59 (19.58-21.66)	21.07 (20.45-21.71)	0.25±0.002	0.71 (0.65-0.77)
1 copy	962 (45.14)	93.61 (87.77-99.85)	26.18±0.14	87.23±0.35	6.50 (6.46-6.55)	20.17 (19.20-21.20)	20.83 (20.23-21.45)	0.25±0.002	0.74 (0.68-0.81)
2 copies	250 (11.73)	91.55 (80.82-103.70)	26.37±0.27	87.84±0.68	6.48 (6.40-6.56)	18.88 (17.14-20.81)	20.43 (19.29-21.64)	0.25±0.004	0.72 (0.61-0.85)
P_{add}		0.49	0.63	0.51	0.58	0.15	0.34	0.70	0.59
P_{dom}		0.54	0.83	0.74	0.37	0.31	0.44	0.54	0.44
P_{rec}		0.62	0.51	0.39	0.82	0.14	0.42	0.89	0.94
B-A-A-B-A-A-A (29.8)									
0 copies	1041 (48.94)	94.50 (88.75-100.64)	26.29±0.13	87.58±0.33	6.48 (6.44-6.52)	20.73 (19.76-21.74)	20.73 (20.15-21.33)	0.25±0.002	0.76 (0.70-0.82)
1 copy	885 (41.61)	95.12 (88.91-101.77)	26.07±0.14	86.98±0.36	6.51 (6.46-6.55)	19.64 (18.65-20.69)	20.91 (20.27-21.56)	0.25±0.002	0.69 (0.63-0.76)
2 copies	201 (9.45)	86.77 (75.63-99.56)	26.15±0.30	86.59±0.75	6.46 (6.37-6.55)	20.02 (17.98-22.30)	21.35 (20.03-22.75)	0.25±0.004	0.71 (0.59-0.86)
P_{add}		0.46	0.35	0.13	0.75	0.23	0.42	0.93	0.25
P_{dom}		0.80	0.25	0.14	0.44	0.14	0.54	0.82	0.16
P_{rec}		0.23	0.88	0.36	0.55	0.87	0.46	0.56	0.86
B-A-A-A-A-A-A (11.0)									
0 copies	1213 (73.34)	92.97 (87.36-98.93)	26.22±0.13	87.20±0.32	6.48 (6.45-6.52)	19.99 (19.11-20.91)	20.76 (20.22-21.32)	0.25±0.002	0.72 (0.66-0.77)
1 copy	425 (25.7)	97.72 (88.58-107.81)	26.23±0.21	87.69±0.52	6.49 (6.43-6.55)	21.45 (19.92-23.10)	21.03 (20.14-21.97)	0.25±0.003	0.77 (0.68-0.88)
2 copies	16 (0.97)	96.50 (63.43-146.81)	25.35±1.07	84.24±2.69	6.03 (5.76-6.33)	23.32 (15.94-34.12)	24.69 (19.74-30.88)	0.27±0.015	0.91 (0.47-1.79)
P_{add}		0.42	0.79	0.74	0.33	0.08	0.31	0.36	0.26
P_{dom}		0.40	0.92	0.55	0.69	0.09	0.46	0.53	0.30
P_{rec}		0.91	0.41	0.25	0.003	0.48	0.14	0.11	0.51
A-A-B-A-A-A-B (8.6)									
0 copies	1795 (83.29)	93.99 (89.62-98.57)	26.23±0.10	87.38±0.26	6.49 (6.46-6.52)	20.30 (19.56-21.06)	21.08 (20.63-21.55)	0.25±0.001	0.73 (0.68-0.78)
1 copy	348 (16.15)	96.83 (87.00-107.76)	26.07±0.23	86.71±0.57	6.46 (6.39-6.53)	20.01 (18.44-21.72)	19.79 (18.85-20.77)	0.25±0.003	0.73 (0.63-0.85)
2 copies	12 (0.56)	114.87 (39.83-331.27)	26.10±1.21	87.90±3.07	6.68 (6.30-7.11)	17.22 (11.10-26.72)	21.96 (16.94-28.48)	0.25±0.017	0.57 (0.26-1.23)
P_{add}		0.57	0.54	0.35	0.70	0.59	0.04	0.32	0.88
P_{dom}		0.60	0.52	0.31	0.55	0.67	0.02	0.29	0.97
P_{rec}		0.72	0.94	0.84	0.33	0.47	0.70	0.87	0.53
A-B-A-A-A-B-B									

(7.4)									
0 copies	1855 (86.6)	94.51 (90.23-99.00)	26.17±0.10	87.21±0.25	6.49 (6.46-6.52)	19.93 (19.22-20.66)	20.80 (20.36-21.25)	0.25±0.001	0.73 (0.68-0.77)
1 copy	275 (12.84)	90.86 (80.18-102.97)	26.30±0.25	87.54±0.64	6.51 (6.43-6.59)	22.43 (20.47-24.59)	21.40 (20.27-22.60)	0.25±0.004	0.71 (0.61-0.84)
2 copies	12 (0.56)	179.79 (106.09-304.69)	27.90±1.20	92.07±3.07	6.09 (5.78-6.45)	18.23 (11.76-28.26)	22.96 (17.70-29.78)	0.23±0.017	0.66 (0.30-1.42)
P_{add}		0.62	0.32	0.30	0.63	0.05	0.25	0.75	0.76
P_{dom}		0.96	0.45	0.44	0.97	0.03	0.28	0.91	0.79
P_{rec}		0.02	0.16	0.12	0.03	0.64	0.48	0.30	0.80

Haplotypes are composed of variants rs1502593 (C>T), rs522951 (G>C), rs11190480 (A>G), rs3071 (T>G), rs3793767 (T>C), rs10883463 (T>C), rs508384 (C>A) in that order. A indicates common allele, B indicates rare allele. a, based on the 615 participants fasting at blood draw. b, geometric means and (95% CI); c, means and standard error; d, inverse and (95% CI); e, based on 2077 participants due to missing biomarker data. All the reported significance levels are nominal P values and are not adjusted for multiple comparisons. P_{add}: P for trend or P for the additive model; P_{dom}: P value for the dominant model. P_{rec}: P value for the recessive model.

Table S1. Genotype and allelic frequencies of the *SCD1* tag-SNPs across the EPIC-Potsdam subcohort and separately for men and women.

SNP	Group	Genotype N (%)			Allele N (%)	
		AA	AB	BB	A	B
rs1502593	Subcohort	679 (31.48)	1072 (49.70)	406 (18.82)	2430 (56)	1884(44)
	Men	233 (28.45)	434 (52.99)	152 (18.56)	900 (55)	738 (45)
	Women	446 (33.33)	638 (47.68)	254 (18.98)	1530 (57)	1146 (43)
rs522951	Subcohort	607 (28.14)	1095 (50.76)	455 (21.09)	2309 (54)	2005 (46)
	Men	223 (27.23)	441 (53.85)	155 (18.93)	887 (54)	751 (46)
	Women	384 (28.70)	654 (48.88)	300 (22.42)	1422 (53)	1254 (47)
rs11190480	Subcohort	1787 (82.85)	357 (16.55)	13 (0.60)	3931 (91)	383 (9)
	Men	681 (83.15)	136 (16.61)	2 (0.24)	1498 (91)	140 (9)
	Women	1106 (82.66)	221 (16.52)	11 (0.82)	2433 (91)	243 (9)
rs3071	Subcohort	944 (43.76)	936 (43.39)	277 (12.84)	2824 (65)	1490 (35)
	Men	356 (43.47)	353 (43.10)	110 (13.43)	1065 (65)	573 (35)
	Women	588 (43.95)	583 (43.57)	167 (12.48)	1759 (66)	917 (34)
rs3793767	Subcohort	845 (39.17)	998 (46.27)	314 (14.56)	2688 (62)	1626 (38)
	Men	330 (40.29)	384 (46.89)	105 (12.82)	1044 (64)	594 (36)
	Women	515 (38.49)	614 (45.89)	209 (15.62)	1644 (61)	1032 (39)
rs10883463	Subcohort	1840 (85.30)	304 (14.09)	13 (0.60)	3984 (92)	330 (8)
	Men	695 (84.86)	120 (14.65)	4 (0.49)	1510 (92)	128 (8)
	Women	1145 (85.58)	184 (13.75)	9 (0.67)	2474 (92)	202 (8)
rs508384	Subcohort	1489 (69.03)	610 (28.28)	58 (2.69)	3588 (83)	726 (17)
	Men	561 (68.50)	244 (29.79)	14 (1.71)	1366 (83)	272 (17)
	Women	928 (69.36)	366 (27.35)	44 (3.29)	2222 (83)	454 (17)

A, most frequent allele in each SNP; B, least frequent allele in each SNP (A>B): rs1502593 (C>T), rs522951 (G>C), rs11190480 (A>G), rs3071 (T>G), rs3793767 (T>C), rs10883463 (T>C), rs508384 (C>A).

Table S2. Association analysis between the 7 *SCD1* tag-SNPs and the 8 investigated metabolic traits in the EPIC-Potsdam Study, (mutually) adjusted for known cardiovascular risk factors.

	Triglycerides [a, b] (mg/dL)	**BMI [c] (kg/m^2)**	**WC [c] (cm)**	**HbA1c [d] (%)**	**GGT [b] (U/L)**	**ALT [b] (U/L)**	**Fetuin-A [c] (mg/dL)**	**hs-CRP [b] (mg/L)**
rs1502593 (n)								
0 (679)	94.54 (88.67-100.80)	26.09±0.13	85.86±0.34	6.47 (6.42-6.52)	18.57 (17.63-19.55)	19.47 (18.87-20.10)	0.251±0.002	0.76 (0.69-0.83)
1 (1072)	88.95 (84.55-93.58)	26.21±0.11	86.03±0.27	6.49 (6.46-6.53)	18.65 (17.90-19.43)	19.76 (19.27-20.26)	0.249±0.002	0.75(0.69-0.80)
2 (406)	95.39 (88.10-103.29)	25.88±0.17	85.03±0.44	6.44 (6.38-6.51)	18.71 (17.50-20.00)	20.60 (19.77-21.45)	0.257±0.003	0.77 (0.68-0.87)
P_{add}	0.20	0.28	0.15	0.40	0.98	0.10	0.09	0.92
P_{dom}	0.30	0.86	0.79	0.71	0.87	0.18	0.90	0.90
P_{rec}	0.30	0.15	0.06	0.27	0.90	0.04	0.03	0.74
rs522951 (n)								
0 (607)	93.70 (87.68-100.14)	26.02±0.14	85.47±0.36	6.46 (6.41-6.51)	18.24 (17.27-19.26)	19.93 (19.28-20.61)	0.254±0.002	0.75 (0.68-0.83)
1 (1095)	91.26 (86.74-96.00)	26.17±0.11	85.86±0.27	6.49 (6.45-6.53)	19.02 (18.26-19.80)	19.89 (19.40-20.39)	0.249±0.002	0.76 (0.71-0.82)
2 (455)	91.16 (84.52-98.33)	26.08±0.16	86.04±0.42	6.47 (6.41-6.53)	18.26 (17.15-19.45)	19.51 (18.78-20.28)	0.252±0.003	0.75 (0.66-0.84)
P_{add}	0.80	0.92	0.92	0.62	0.38	0.67	0.35	0.94
P_{dom}	0.51	0.46	0.31	0.38	0.36	0.70	0.21	0.88
P_{rec}	0.81	0.87	0.50	0.90	0.48	0.37	0.78	0.81
rs11190480 (n)								
0 (1787)	91.10 (87.65-94.69)	26.10±0.08	85.82±0.21	6.48 (6.45-6.51)	18.71 (18.13-19.32)	19.99 (19.61-20.38)	0.25±0.001	0.76 (0.71-0.80)
1 (357)	96.28 (88.11-105.21)	26.14±0.19	85.61±0.47	6.45 (6.39-6.52)	18.42 (17.15-19.78)	18.96 (18.16-19.81)	0.25±0.003	0.75 (0.66-0.85)
2 (13)	103.59 (42.75-251.04)	25.76±0.98	85.86±2.47	6.66 (6.31-7.05)	14.40 (9.92-20.93)	20.90 (18.16-19.81)	0.25±0.016	0.65 (0.33-1.28)
P_{add}	0.52	0.45	0.45	0.46	0.37	0.09	0.57	0.91
P_{dom}	0.26	0.89	0.70	0.57	0.53	0.04	0.29	0.85
P_{rec}	0.79	0.72	0.98	0.32	0.18	0.18	0.82	0.67
rs3071 (n)								
0 (944)	91.93 (87.06-97.06)	26.21±0.11	86.00±0.29	6.46 (6.41-6.50)	18.96 (18.14-19.81)	19.64 (19.12-20.17)	0.251±0.002	0.79 (0.73-0.85)
1 (936)	92.54 (87.64-97.71)	25.97±0.12	85.59±0.29	6.49 (6.45-6.53)	18.19 (17.41-19.01)	19.83 (19.30-20.37)	0.252±0.002	0.72 (0.67-0.78)
2 (277)	90.07 (81.73-99.27)	26.23±0.21	85.75±0.53	6.50 (6.43-6.58)	19.07 (17.58-20.67	20.43 (19.45-21.46)	0.251±0.003	0.76 (0.65-0.88)
P_{add}	0.89	0.27	0.62	0.38	0.36	0.39	0.97	0.36
P_{dom}	0.99	0.23	0.34	0.17	0.31	0.37	0.82	0.19
P_{rec}	0.66	0.52	0.94	0.50	0.55	0.20	0.96	0.97
rs3793767 (n)								
0 (845)	93.56 (88.42-99.01)	26.17±0.12	85.96±0.31	6.46 (6.41-6.50)	18.83 (17.98-19.72)	19.97 (19.41-20.54)	0.252±0.002	0.73 (0.67-0.80)
1 (998)	90.97 (86.33-95.86)	26.02±0.11	85.59±0.28	6.50 (6.46-6.54)	18.85 (18.07-19.67)	19.82 (19.31-20.34)	0.249±0.002	0.78 (0.72-0.84)
2 (314)	90.55 (82.26-99.68)	26.20±0.20	85.96±0.50	6.47 (6.40-6.54)	17.47 (16.19-18.85)	19.43 (18.55-20.36)	0.254±0.003	0.74 (0.65-0.85)
P_{add}	0.74	0.59	0.96	0.42	0.20	0.62	0.31	0.56
P_{dom}	0.44	0.49	0.46	0.26	0.58	0.51	0.51	0.38

P_{rec}	0.74	0.63	0.72	0.79	0.07	0.37	0.28	0.77
rs10883463 (n)								
0 (1840)	92.66 (89.21-96.24)	26.07±0.08	85.69±0.21	6.48 (6.45-6.51)	18.38 (17.81-18.96)	19.77 (19.40-20.15)	0.251±0.001	0.76 (0.72-0.80)
1 (304)	85.11 (76.99-94.09)	26.28±0.20	86.22±0.51	6.47 (6.40-6.55)	20.44 (18.92-22.08)	20.12 (19.19-21.09)	0.251±0.003	0.74 (0.65-0.86)
2 (13)	132.82 (88.81-198.6)5	27.56±0.98	89.15±2.48	6.08 (5.79-6.41)	15.23 (10.48-22.14)	20.34 (16.18-25.56)	0.241±0.016	0.53 (0.27-1.04)
P_{add}	0.06	0.21	0.25	0.06	0.03	0.78	0.84	0.57
P_{dom}	0.28	0.22	0.23	0.56	0.02	0.48	0.96	0.67
P_{rec}	0.07	0.14	0.17	0.02	0.29	0.83	0.56	0.30
rs508384 (n)								
0 (1489)	91.70 (87.91-95.67)	26.03±0.9	85.67±0.23	6.49 (6.45-6.52)	18.54 (17.90-19.20)	19.99 (19.57-20.42)	0.252±0.001	0.76 (0.71-0.81)
1 (610)	91.14 (85.08-97.64)	26.25±0.14	85.91±0.36	6.46 (6.41-6.51)	18.87 (17.87-19.93)	19.44 (18.80-20.09)	0.249±0.002	0.75 (0.68-0.83)
2 (58)	106.69 (85.97-132.40)	26.58±0.56	87.49±1.17	6.44 (6.28-6.61)	18.55 (15.53-22.16)	19.64 (17.62-21.89)	0.249±0.008	0.65 (0.47-0.90)
P_{add}	0.39	0.26	0.29	0.62	0.86	0.38	0.47	0.65
P_{dom}	0.84	0.14	0.36	0.34	0.61	0.17	0.17	0.68
P_{rec}	0.17	0.31	0.14	0.66	0.96	0.87	0.74	0.36

Each SNP is coded as 0, 1 and 2 according to the number of minor alleles a participant carries. a, based on the 615 participants fasting at blood draw; b, geometric means and (95% CI); c, means and standard error; d, inverse and (95% CI); e based on 2077 participants due to missing biomarker data; P_{add}, P for trend or P for the additive model; P_{dom}, P value for the dominant model; P_{rec}, P value for the recessive model. All the reported significance levels are nominal P values and are not adjusted for multiple comparisons. All analyses are (mutually) adjusted for age, sex, smoking status (never smoker, former smoker, current smoker <20 cigarettes per day, current smoker ≥20 cigarettes per day), sports activity (<2 h/wk versus ≥2 h/wk), educational attainment (vocational school or less, technical school, university), BMI (continuous), WC (continuous), alcohol consumption (men: = 0 g/d, >0 to 12 g/d, >12 to 24 g/d; >24 g/d; women: = 0 g/d, >0 to 6 g/d, >6 to 12 g/d; >12 g/d), prevalent diabetes, prevalent hypertension, total cholesterol, HDL-cholesterol and hs-CRP.

Table S3. β regression coefficients for the age- and sex-adjusted association analysis between the *SCD1* tag-SNPs and inferred haplotypes and the 8 investigated metabolic traits in the EPIC-Potsdam study.

	N	Ln (Triglycerides) [a] (mg/dL)	BMI (kg/m^2)	WC (cm)	1/(HbA1c) (%)	Ln (GGT) (U/L)	Ln (ALT) (U/L)	Fetuin-A[b] (mg/dL)	Ln (hs-CRP) (mg/L)
rs1502593									
β (SE) $_{add}$	0 (679)	-0.004 (0.031)	-0.09 (0.13)	-0.38 (0.33)	0.0002 (0.0005)	0.004 (0.024)	0.02 (0.01)	0.002 (0.002)	-0.01 (0.04)
β (SE) $_{dom}$	1 (1072)	-0.03 (0.05)	-0.04 (0.19)	-0.31 (0.49)	-0.0002 (0.0007)	0.01 (0.04)	0.03 (0.02)	0.0009 (0.003)	-0.01 (0.06)
β (SE) $_{rec}$	2 (406)	0.03 (0.05)	-0.24 (0.23)	-0.78 (0.59)	0.0010 (0.0009)	0.004 (0.042)	0.04 (0.03)	0.006 (0.003)	-0.03 (0.08)
P $_{add}$		0.89	0.47	0.25	0.63	0.87	0.11	0.19	0.88
P $_{dom}$		0.53	0.83	0.53	0.80	0.87	0.24	0.75	0.86
P $_{rec}$		0.63	0.30	0.18	0.25	0.93	0.14	0.06	0.96
rs522951									
β (SE) $_{add}$	0 (607)	-0.003 (0.031)	0.14 (0.13)	0.53 (0.33)	-0.0003 (0.0005)	0.003 (0.02)	-0.01(0.01)	-0.001 (0.002)	0.01 (0.04)
β (SE) $_{dom}$	1 (1095)	0.01 (0.05)	0.22 (0.20)	0.65 (0.51)	-0.0007 (0.0008)	0.02 (0.04)	-0.01(0.02)	-0.004 (0.003)	0.02 (0.07)
β (SE) $_{rec}$	2 (455)	-0.02 (0.05)	0.14 (0.22)	0.77 (0.56)	-0.0001 (0.0008)	-0.01 (0.04)	-0.01(0.02)	0.0005 (0.0031)	0.01 (0.07)
P $_{add}$		0.93	0.28	0.10	0.50	0.90	0.70	0.46	0.82
P $_{dom}$		0.90	0.27	0.20	0.35	0.60	0.78	0.20	0.80
P $_{rec}$		0.77	0.52	0.17	0.90	0.72	0.73	0.87	0.91
rs11190480									
β (SE) $_{add}$	0 (1787)	0.03 (0.06)	-0.20 (0.23)	-0.66 (0.58)	0.0004 (0.0009)	-0.03 (0.04)	-0.05 (0.02)	-0.003 (0.003)	-0.02 (0.07)
β (SE) $_{dom}$	1 (357)	0.03 (0.06)	-0.19 (0.24)	-0.71 (0.61)	0.0006 (0.0009)	-0.03 (0.04)	-0.06 (0.03)	-0.003 (0.003)	-0.01 (0.08)
β (SE) $_{rec}$	2 (13)	0.20 (0.54)	-0.60 (1.16)	-0.55 (2.96)	0.0039 (0.0044)	-0.27 (0.22)	0.04 (0.13)	-0.006 (0.017)	-0.21 (0.38)
P $_{add}$		0.56	0.39	0.25	0.63	0.41	0.05	0.35	0.78
P $_{dom}$		0.58	0.42	0.24	0.49	0.54	0.03	0.36	0.85
P $_{rec}$		0.72	0.61	0.85	0.38	0.21	0.77	0.73	0.59
rs3071									
β (SE) $_{add}$	0 (944)	-0.02 (0.03)	-0.10 (0.13)	-0.30 (0.33)	-0.001 (0.001)	-0.02 (0.02)	0.01 (0.01)	-0.0002 (0.0018)	-0.05 (0.04)
β (SE) $_{dom}$	1 (936)	-0.01 (0.04)	-0.30 (18)	-0.70 (0.46)	-0.001 (0.001)	-0.04 (0.03)	0.01 (0.02)	0.0004 (0.0025)	-0.11 (0.06)
β (SE) $_{rec}$	2 (277)	-0.04 (0.06)	0.26 (0.27)	0.27 (0.68)	-0.001 (0.001)	0.01 (0.05)	0.02 (0.03)	-0.002 (0.004)	0.01 (0.09)
P $_{add}$		0.62	0.47	0.36	0.28	0.46	0.46	0.92	0.22
P $_{dom}$		0.79	0.10	0.13	0.32	0.27	0.62	0.86	0.08
P $_{rec}$		0.54	0.34	0.69	0.46	0.92	0.43	0.64	0.90
rs3793767									

β (SE) $_{add}$	0 (845)	-0.03 (0.03)	0.09 (0.13)	0.19 (0.33)	-0.0003 (0.0005)	-0.03 (0.02)	-0.01 (0.01)	0.0001 (0.0018)	0.02 (0.04)
β (SE) $_{dom}$	1 (998)	-0.03 (0.04)	0.04 (0.18)	0.08 (0.47)	-0.0008 (0.0007)	-0.02 (0.03)	-0.01 (0.02)	-0.002 (0.003)	0.04 (0.06)
β (SE) $_{rec}$	2 (314)	-0.06 (0.06)	0.25 (0.26)	0.57 (0.65)	0.0002 (0.0010)	-0.08 (0.05)	-0.02 (0.03)	0.004 (0.004)	-0.003 (0.008)
P $_{add}$		0.37	0.50	0.57	0.50	0.17	0.38	0.97	0.65
P $_{dom}$		0.56	0.81	0.86	0.26	0.49	0.48	0.50	0.50
P $_{rec}$		0.34	0.33	0.38	0.81	0.09	0.47	0.32	0.97
rs10883463									
β (SE) $_{add}$	0 (1840)	0.02 (0.06)	0.16 (.24)	0.43 (0.61)	0.0008 (0.0009)	0.09 (0.04)	0.03 (0.03)	-0.001 (0.003)	-0.02(0.08)
β (SE) $_{dom}$	1 (304)	-0.02 (0.06)	0.12 (0.25)	0.32 (0.65)	0.0005 (0.0010)	0.11 (0.05)	0.03 (0.03)	0.00003 (0.004)	-0.01(0.08)
β (SE) $_{rec}$	2 (13)	0.54 (0.24)	1.30 (1.16)	3.54 (2.96)	0.0098 (0.0044)	-0.12 (0.22)	0.07 (0.13)	-0.017 (0.017)	-0.18(0.38)
P $_{add}$		0.76	0.50	0.48	0.36	0.04	0.27	0.85	0.81
P $_{dom}$		0.80	0.64	0.63	0.62	0.02	0.30	0.99	0.88
P $_{rec}$		0.03	0.26	0.23	0.03	0.56	0.56	0.32	0.63
rs508384									
β (SE) $_{add}$	0 (1489)	0.04 (0.04)	0.04 (0.17)	0.02 (0.44)	0.0006 (0.0007)	0.03 (0.03)	-0.02 (0.02)	-0.002 (0.002)	-0.01 (0.06)
β (SE) $_{dom}$	1 (610)	0.02 (0.05)	0.05 (0.19)	-0.05 (0.50)	0.0006 (0.0007)	0.03 (0.04)	-0.02 (0.02)	-0.003 (0.003)	0.004 (0.064)
β (SE) $_{rec}$	2 (58)	0.24 (0.13)	-0.01 (0.56)	0.66 (1.42)	0.0012 (0.0021)	0.02 (0.10)	-0.001 (0.061)	-0.005 (0.008)	-0.13 (0.18)
P $_{add}$		0.36	0.81	0.96	0.36	0.40	0.39	0.27	0.86
P $_{dom}$		0.69	0.78	0.92	0.39	0.38	0.33	0.31	0.95
P $_{rec}$		0.07	0.98	0.64	0.58	0.82	0.99	0.49	0.46
A-B-A-A-B-A-A									
β (SE) $_{add}$	0 (919)	-0.02 (0.03)	0.06 (0.13)	0.23 (0.34)	-0.0003 (0.0005)	-0.04 (0.03)	-0.01 (0.01)	-0.0007 (0.0019)	0.02 (0.04)
β (SE) $_{dom}$	1 (962)	-0.03 (0.04)	0.04 (0.18)	0.16 (0.47)	-0.0006 (0.0007)	-0.03 (0.03)	-0.02 (0.02)	-0.002 (0.003)	0.05 (0.06)
β (SE) $_{rec}$	2 (250)	-0.03 (0.07)	0.19 (0.28)	0.62 (0.72)	0.0002 (0.0011)	-0.08 (0.05)	-0.02 (0.03)	0.0006 (0.0039)	-0.01 (0.09)
P $_{add}$		0.49	0.63	0.51	0.58	0.15	0.34	0.70	0.59
P $_{dom}$		0.54	0.83	0.74	0.37	0.31	0.44	0.54	0.44
P $_{rec}$		0.62	0.51	0.39	0.82	0.14	0.42	0.89	0.94
B-A-A-B-A-A-A									
β (SE) $_{add}$	0 (1041)	-0.02 (0.03)	-0.13 (0.14)	-0.54 (0.35)	-0.0002 (0.0005)	-0.03 (0.03)	0.01 (0.02)	0.0002 (0.0019)	-0.052 (0.045)
β (SE) $_{dom}$	1 (885)	-0.01 (0.04)	-0.21 (0.18)	-0.68 (0.46)	-0.0005 (0.0007)	-0.05 (0.03)	0.01 (0.02)	-0.0006 (0.0025)	-0.08 (0.06)
β (SE) $_{rec}$	2 (201)	-0.09 (0.07)	-0.05 (0.31)	-0.72 (0.79)	-0.0007 (0.0012)	-0.01 (0.06)	0.03 (0.03)	0.0025 (0.0043)	-0.02 (0.10)
P $_{add}$		0.49	0.63	0.51	0.58	0.15	0.34	0.70	0.59
P $_{dom}$		0.54	0.83	0.74	0.37	0.31	0.44	0.54	0.44
P $_{rec}$		0.62	0.51	0.39	0.82	0.14	0.42	0.89	0.94
B-A-A-A-A-A-A									
β (SE) $_{add}$	0 (1213)	0.04 (0.05)	-0.06 (0.22)	0.19 (0.56)	0.0008 (0.0008)	0.07 (0.04)	0.02 (0.02)	0.003 (0.003)	0.08 (0.07)

β (SE) $_{dom}$	1 (425)	0.05 (0.06)	-0.02 (0.24)	0.36 (0.60)	0.0003 (0.0009)	0.07 (0.04)	0.02 (0.03)	0.002 (0.003)	0.08 (0.08)
β (SE) $_{rec}$	2 (16)	0.02 (0.22)	-0.88 (1.07)	-3.09 (2.70)	0.012 (0.004)	0.14 (0.20)	0.17 (0.11)	0.02 (0.02)	0.23 (0.34)
P $_{add}$		0.46	0.35	0.13	0.75	0.23	0.42	0.93	0.25
P $_{dom}$		0.80	0.25	0.14	0.44	0.14	0.54	0.82	0.16
P $_{rec}$		0.23	0.88	0.36	0.55	0.87	0.46	0.56	0.86
A-A-B-A-A-A-B									
β (SE) $_{add}$	0 (1795)	0.03 (0.06)	-0.14 (0.23)	-0.55 (0.58)	0.0003 (0.0009)	-0.02 (0.04)	-0.05 (0.03)	-0.003 (0.003)	-0.01 (0.07)
β (SE) $_{dom}$	1 (348)	0.03 (0.06)	-0.15 (0.24)	-0.63 (0.61)	0.0006 (0.0009)	-0.02 (0.04)	-0.06 (0.03)	-0.004 (0.003)	-0.003 (0.078)
β (SE) $_{rec}$	2 (12)	0.20 (0.54)	-0.10 (1.2)	0.63 (3.08)	-0.005 (0.005)	-0.16 (0.22)	0.05 (0.13)	-0.0034 (0.017)	-0.25 (0.40)
P $_{add}$		0.57	0.54	0.35	0.70	0.59	0.04	0.32	0.88
P $_{dom}$		0.60	0.52	0.31	0.55	0.67	0.02	0.29	0.97
P $_{rec}$		0.72	0.94	0.84	0.33	0.47	0.70	0.87	0.53
A-B-A-A-A-B-B									
β (SE) $_{add}$	0 (1855)	0.03 (0.06)	0.25 (0.25)	0.66 (0.63)	0.0005 (0.0010)	0.09 (0.05)	0.03 (0.03)	-0.001 (0.003)	-0.02 (0.08)
β (SE) $_{dom}$	1 (275)	-0.003 (0.066)	0.20 (0.26)	0.52 (0.67)	0.00003 (0.0010)	0.11 (0.05)	0.03 (0.03)	-0.0004 (0.0037)	-0.02 (0.09)
β (SE) $_{rec}$	2 (12)	0.65 (0.27)	1.71 (1.21)	4.82 (3.07)	0.010 (0.005)	-0.10 (0.22)	0.10 (0.13)	-0.02 (0.02)	-0.10 (0.39)
P $_{add}$		0.62	0.32	0.30	0.63	0.05	0.25	0.75	0.76
P $_{dom}$		0.96	0.45	0.44	0.97	0.03	0.28	0.91	0.79
P $_{rec}$		0.02	0.16	0.12	0.03	0.64	0.48	0.30	0.80

Each SNP and haplotype is coded as 0, 1 and 2 according to the number of copies of the least common variant a participant carries. Haplotypes are composed of variants rs1502593 (C>T), rs522951 (G>C), rs11190480 (A>G), rs3071 (T>G), rs3793767 (T>C), rs10883463 (T>C), rs508384 (C>A) in that order. A indicates common allele, B indicates rare allele. β (SE), regression coefficient and standard error obtained in linear regression analysis. a, based on the 615 participants fasting at blood draw; b, based on 2077 participants due to missing biomarker data. All the reported significance levels are nominal P values and are not adjusted for multiple comparisons. P $_{add}$, P for trend or P for the additive model; P $_{dom}$, P value for the dominant model. P $_{rec}$, P value for the recessive model.

5

Microsomal triglyceride transfer protein -164 T>C gene polymorphism and risk of cardiovascular disease: results from the EPIC-Potsdam case-cohort study

Chapter 5

Microsomal triglyceride transfer protein -164 T>C gene polymorphism and risk of cardiovascular disease: results from the EPIC-Potsdam case-cohort study

Romina di Giuseppe*[1], Sonali Pechlivanis[3], Eva Fisher[4], **Maria Arregui**[1], Beate Weikert[5], Sven Knüppel[1], Brian Buijsse[1], Andreas Fritsche[6], Stefan N. Willich[7], Hans-Georg Joost[2], Heiner Boeing[1], Susanne Moebus[3], Cornelia Weikert[1,7]

[1]Department of Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany

[2]Department of Pharmacology, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany

[3]Institute for Medical Informatics, Biometry and Epidemiology, University Hospital of Essen, University Duisburg-Essen, Essen, Germany,

[4]Administrative Office of the Commission on Genetic Testing, Robert Koch-Institute, Berlin, Germany

[5]Agency for Quality in Medicine, Berlin, Germany

[6]Department of Internal Medicine, Division of Endocrinology, Diabetology, Nephrology, Vascular Disease and Clinical Chemistry, University of Tübingen, Tübingen, Germany

[7]Institute for Social Medicine, Epidemiology and Health Economics and Chairman, Charité Center 1 for Humanities and Health Sciences, Charité University Medical Center, Berlin, Germany

BMC Med Genet. 2013 Jan 29;14:19

Impact Factor 2012: 2.328

Abstract

Background: The microsomal triglyceride transfer protein (MTTP) is encoded by the *MTTP* gene that is regulated by cholesterol in humans. Previous studies investigating the effect of *MTTP* on ischemic heart disease have produced inconsistent results. Therefore, we have tested the hypothesis that the rare allele of the -164T>C polymorphism in *MTTP* alters the risk of cardiovascular disease (CVD), depending on the cholesterol levels.

Methods: The -164T>C polymorphism was genotyped in a case-cohort study (193 incident myocardial infarction (MI) and 131 incident ischemic stroke (IS) cases and 1 978 non-cases) nested within the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study, comprising 27 548 middle-aged subjects. The Heinz Nixdorf Recall Study (30 CVD cases and 1,188 controls) was used to replicate our findings.

Results: Genotype frequencies were not different between CVD and CVD free subjects (P=0.79). We observed an interaction between the -164T>C polymorphism and total cholesterol levels in relation to future CVD. Corresponding stratified analyses showed a significant increased risk of CVD ($HR_{additve}$= 1.38, 95% CI: 1.07 to 1.78) for individuals with cholesterol levels <200 mg/dL in the EPIC-Potsdam Study. $HR_{additive}$ was 1.06, 95% CI: 0.33 to 3.40 for individuals in the Heinz Nixdorf Recall Study. A borderline significant decrease in CVD risk was observed in subjects with cholesterol levels ≥200 mg/dL ($HR_{additve}$=0.77, 95% CI: 0.58 to 1.03) in the EPIC-Potsdam Study.A similar trend was observed in the independent cohort ($HR_{additve}$=0.60, 95% CI: 0.29 to 1.25).

Conclusions: Our study suggests an interaction between *MTTP* -164T>C functional polymorphism with total cholesterol levels. Thereby risk allele carriers with low cholesterol levels may be predisposed to an increased risk of developing CVD, which seems to be abolished among risk allele carriers with high cholesterol levels.

Keywords: epidemiology, genetics, myocardial infarction, ischemic stroke, cholesterol, additive interaction.

BACKGROUND

The microsomal triglyceride transfer protein (MTTP), a lipid transfer protein encoded by the *MTTP* gene, is located in the luminal side of the endoplasmic reticulum [1, 2]. It plays an important role in the assembly and secretion of apolipoproteins B (ApoB) containing lipoproteins as chylomicrons in the intestine, and of very low density lipoproteins (VLDLs) in the liver [3-5]. In turn, chylomicrons transport exogenous lipids to cells, while VLDLs carry endogenous triglycerides through the bloodstream. Thus, considering the important role played by *MTTP* in fat absorption and lipoprotein regulation, several studies have been conducted on *MTTP* promoter polymorphisms [6-24].

Three polymorphisms of the *MTTP* gene (4q24), Ile128Thr (rs3816873), -164T>C (rs1800804), and -493G>T (rs1800591), have been described which are in complete linkage disequilibrium [6]. The rare alleles of these polymorphisms have been reported to decrease plasma lipoprotein-lipid levels [6-14] and some features of metabolic syndrome [16-19]. However, conflicting or negative findings have been reported by others [20-26]. The same inconsistent results have been published with regard to coronary heart disease (CHD) [7, 24-26]. Two studies reported null findings [23, 24]. In contrast, in the INTERGENE [26], ULSAM and WOSCOPS [7] studies homozygosity for the -164C and the -493T alleles was associated with increased risk of ischemic heart disease (IHD) though, in the latter, concomitantly to a decrease in total plasma cholesterol [7]. Nevertheless, considered the reduced expression of the *MTTP* gene in carriers of the rare alleles [26], it is reasonable to expect lower cholesterol levels which have previously been shown to regulate MTTP gene expression [27, 28]. Also, it is likely to assume an interaction between the genotype of MTTP and cholesterol levels [19] in affecting the risk of CHD.

Therefore, in the framework of the European Prospective Investigation into Cancer and Nutrition (EPIC)–Potsdam Study we investigated the association of the -164T>C variant, used as proxy for the three loci, with cardiovascular disease (CVD), myocardial infarction (MI) and ischemic stroke (IS) and the presence of an interaction between -164T>C polymorphism and total cholesterol in relation to CVD (MI and IS). We hypothesized that in carriers of the *MTTP* -164 C-allele lower cholesterol levels are associated with increased susceptibility to MI and IS.

METHODS

EPIC-Potsdam

The Ethics Committee of the Medical Association of the State of Brandenburg approved the study procedures and a written informed consent was obtained from all participants.

Heinz Nixdorf Recall Study

The study was approved by the local ethics committees, was conducted in accordance with German *Good Epidemiologic Practice* (GEP) including extended quality management procedures and re-certifications according to DIN ISO 9001:2001. Informed consent was obtained from all participants. Information on genotype, sex, age, diabetes, anti-hyperlipidemic drugs and CVD was available for up to 1,513 of 4,814 participants.

Study population

EPIC-Potsdam

Between 1994 and 1998, as a part of the large-scale European prospective cohort study EPIC, the EPIC-Potsdam Study enrolled from the general population 16 644 women (aged 35 to 65 years) and 10 904 men (aged 40 to 65 years for), for a total of 27 548 participants [29].

At baseline, self-administered questionnaires on diet and lifestyle, PC-guided interviews with additional questions on prevalent diseases, blood pressure and anthropometric measurements were collected following standard procedures [30].

Every two years, information on incident diseases and changes in lifestyle are collected by self-administered questionnaires [31], with response rates during follow-up exceeding 90% at all occasions.

A case-cohort study consisting of all incident cases identified during a mean follow-up of 8 ± 2.2 years [32] and a sub-cohort of 2 500 individuals randomly selected from the EPIC-Potsdam cohort [33], was used to assess the association of the 164T>C variant with CVD risk (including both MI and IS).

With this type of study, the results are expected to be representative of the entire cohort [34,

35]. After the exclusion of participants with prevalent MI and stroke at the baseline, 269 individuals with incident MI and 246 individuals with incident stroke were identified (199 IS, 41 haemorrhagic strokes, and 6 strokes with undefined pathogenesis). Among the sub-cohort, non-ischemic strokes were not considered as cases, while for individuals who experienced both MI and IS only the first event was considered [36]. After exclusion of prevalent CVD and missing follow-up dates, 2 368 participants remained to be in the sub-cohort. After further exclusion of subjects with a history of diabetes at the baseline and/or treated with anti-hyperlipidemic drugs, and those with missing *MTTP* genotype, biochemical or covariates data the final case-cohort consisted of 2 302 individuals (324 CVD cases: 193 MI and 131 IS, and 1 978 non-cases). Among CVD cases, 37 were part of the subcohort.

Heinz Nixdorf Recall Study

For the replication we analyzed data from the Heinz Nixdorf Recall (Risk Factors, Evaluation of Coronary Calcium and Lifestyle) study comprising 4,814 participants aged 45–75 years. The participants were randomly selected from registration lists of the densely populated Ruhr metropolitan area in Germany between 2000 and 2003. The rationale and design of the study have been described in detail [37]. Between 2006 and 2008 incident cases were identified during the 5-year follow-up examination.

The genotyping of the *MTTP* SNP –I128T (rs3816873) was already available in a random selected sample of n=1,513 *Heinz Nixdorf Recall* participants out of 4,814. After exclusion of participants with a history of CVD and/or diabetes at baseline and/or treated with anti-hyperlipidemic drugs, the final replication cohorts consisted of n=1,218 individuals (30 CVD and 1,188 non-cases).

Ascertainment of CVD

EPIC-Potsdam Study

As described elsewhere [32], all possible cases of MI or stroke were identified by self-report or by death certificate in one of the four follow-up questionnaires and further verified by contacting the patients' attending physician or by review of death certificates according to WHO MONICA criteria.

Heinz Nixdorf Recall Study

Incident CVD (n=30) included fatal and nonfatal MI (n=24) and other CVD (n=6), were identified.

Primary endpoints for this study were based on unequivocally documented incident coronary events that met predefined study criteria. We considered a myocardial infarction event based on symptoms, signs of electrocardiography, and enzymes (levels of creatine kinase (CK-MB)) as well as troponin T or I, and necropsy as 1) non-fatal acute myocardial infarction and 2) coronary death, which occurred between the baseline examination and five years after study entry [38, 39]. For all primary study endpoints, hospital and nursing home records including electrocardiograms, laboratory values, and pathology reports were collected. For deceased subjects, death certificates were collected and interviews with general practitioners, relatives and eyewitnesses were undertaken if possible. Medical records were obtained in 100% of all reported endpoints. An external criteria and endpoint committee blinded for conventional risk factor status and CAC scores reviewed all documents and classified the endpoints thereafter. Due to the small number of cases only total CVD were considered for the replication analysis.

Other measurements

Prevalent diabetes mellitus was identified by a physician and based on self-reported medical diagnoses, medication records and dieting behavior. Prevalent hypertension was defined as follows: systolic blood pressure ≥140 mm Hg or diastolic blood pressure ≥90 mm Hg or self-reporting of a diagnosis or use of antihypertensive medication. Education, lifestyle characteristics (including alcohol consumption), regular sport activity and smoking history were assessed at baseline by trained interviewers during a PC-guided interview. Trained personnel took anthropometric and blood pressure measurements.

Biochemical analyses

EPIC-Potsdam Study

At baseline, 30 ml of venous blood was taken from the respected participant (non-fasting or fasting blood) and, after fractionation into serum, plasma, leukocytes, and erythrocytes immediately stored at -196°C in liquid nitrogen [31]. All biomarkers were determined in 2007

in the Department of Internal Medicine, University of Tübingen. Plasma glucose, high-density lipoprotein cholesterol (HDL), total cholesterol and triglycerides were measured with an automatic analyzer (ADVIA 1650, Siemens Medical Solutions, Erlangen, Germany). LDL-cholesterol was calculated using Friedewald's formula [40]. To account for citrate's dilution factor concentrations of total, HDL-, LDL-cholesterol and triglycerides were multiplied by 1.1.

Heinz Nixdorf Recall Study

At baseline, plasma cholesterol levels were measured with an automatic analyzer (ADVIA 1650, Siemens Medical Solutions, Erlangen, Germany).

Genetic analyses

EPIC-Potsdam Study

DNA extraction was performed using a commercial kit (Qiagen, Hilden, Germany). In 2009 at the Max Delbrück Center for Molecular Medicine, Berlin, Germany, the *MTTP* SNP -164T>C (rs1800804) was genotyped by TaqMan technology (Applied Biosystems, Foster City, CA, USA) using 5 ng of whole-genome amplified DNA per sample. The call rate for the SNP assay exceeded 98%.

Heinz Nixdorf Recall Study

Lymphocyte DNA was isolated from EDTA anti-coagulated venous blood by a Chemagic Magnetic Separation Module I (Chemagen, Baesweiler, Germany). The *MTTP* SNP –I128T (rs3816873) was genotyped using four different platforms: Illumina Hap300, Illumina Hap550, Illumina Human660W-Quad and Illumina HumanOmni1-Quad. The call rate for this SNP was 99.9%.

Statistical Analysis

For both EPIC-Potsdam and Heinz Nixdorf Recall Studies statistical analyses were performed with the use of SAS software package, release 9.2 (SAS Institute, Cary, NC).

EPIC-Potsdam Study

Data on triglyceride measurements were transformed into natural logarithms to reduce skewness and data were reported as geometric means and 95% confidence interval (CI).

The deviation from Hardy-Weinberg equilibrium (HWE) was measured using the χ^2 test.

The HWE was tested in the subcohort. Age and sex adjusted analysis of variance was used to describe general characteristics according to -164T>C genotype. Data were reported as means and standard error (SE). P for trend was calculated from age and sex adjusted linear regression model. To investigate the impact of total cholesterol on the associations between genotype and CVD we performed stratified analyses according to low (<200 mg/dL) and borderline-high/high (≥200 mg/dL) cholesterol levels as defined in the Adult Treatment Panel III (ATP III) report [41]. Multiplicative interaction between cholesterol levels and genotype as well as sex and genotype in relation to CVD was tested with cross-product term. As a measure of additive interaction between cholesterol levels and genotype we further calculated the Synergy Index (SI) and the relative excess risk due to interaction (RERI) and their 95% confidence interval (CI) [42], as suggested by Rothman [43]. SI>1 and RERI>0 suggest a positive interaction (superadditive effect); SI<1 and RERI<0 suggest a negative interaction (subadditive effect); SI=1 and RERI=0 suggest absence of interaction.

Both the multiplicative and additive interactions were also tested for triglycerides, HDL- and LDL-cholesterol. The choice of the cut-off values for these biomarkers was also based on the ATP III criteria (lower and higher than 130 mg/dL for LDL-cholesterol; lower and higher than 40 mg/dL and lower and higher than 50 mg/dL for HDL-cholesterol, respectively, in men and women; lower and higher than 150 mg/dl for triglycerides) [41].

Cox proportional-hazard regression modified according to the Prentice method [34] was used to compute the age and sex adjusted hazard ratio (HR) and 95% CI for the associations between *MTTP* -164T>C and risk of MI, IS and total CVD (combined endpoints). In the counting processes age was the underlying time variable with "entry time" defined as age at baseline and "exit time" as age at CVD event (MI or IS) or censoring. Associations between *MTTP* -164T>C and CVDs were tested in three models: additive, dominant and recessive. Furthermore, competing risk analyses were performed to test whether the associations of *MTTP* -164T>C with cardiovascular events differed between MI and IS, as described by Lunn and McNeil [44].

The power to detect an association between the *MTTP* -164T>C SNP and CVDs (MI, IS and combined endpoints) was computed with Quanto (http://hydra.usc.edu/GxE/) [45], in relation to a desirable power =80%, assuming $\alpha = 0.05$ and a disease prevalence of 3.1% for CVD (n=847), 2.0% for MI (n=544) and 1.1% for IS (n=303), reflecting the baseline prevalence data in the EPIC-Potsdam population. The detectable odds ratio per risk allele equals 1.30, 1.39 and 1.47, respectively for CVD, MI and IS.

Heinz Nixdorf Recall Study

Cox proportional-hazard regression was used to compute the age and sex adjusted hazard ratio (HR) and 95% CI for the associations between *MTTP* –I128T and risk of total CVD. Associations between *MTTP* –I128T and CVDs were tested using the additive model.

RESULTS

General characteristics

The genotype distribution of the -164T>C and –I128T SNPs followed the HWE (*P*=0.37 and *P*=0.71, respectively for the EPIC-Potsdam subcohort and for the replication Heinz Nixdorf Recall Study). There were 1 114, 758 and 143 subjects with genotypes TT, CT, CC respectively (738 men and 1 277 women) observed in the sub-cohort. The C allele frequency was 0.26 for both studies. According to -164T>C genotype, age- and sex-adjusted baseline characteristics of subjects who did or did not develop cardiovascular events during the follow-up period are shown in Table 1. In particular, 58.5% of subjects with incident CVD (n=324; 193 incident MI; 131 incident IS) were men, older, smokers and with a lower educational level than individuals who remained free of CVD (n=1 978) during a mean follow-up of 8.2 years. Furthermore, they had a higher prevalence of abdominal obesity and hypertension, slightly higher total and low density lipoprotein- (LDL-) and lower HDL-cholesterol levels (Table 1). According to genotype no significant differences in central obesity, obesity and hypertension and socio-demographic characteristics were observed in subjects with or without CVD (Table 1). In contrast, opposite trends were observed for total-, LDL-cholesterol and triglyceride levels. In the group free of CVD carriers of the C-allele showed slightly higher total- and LDL-cholesterol levels ($P_{trend} = 0.036$ and $P_{trend}=0.026$, respectively). In the group

of CVD carriers of the C-allele showed lower triglyceride (P_{trend}=0.016) along with slight lower total-cholesterol levels (P_{trend}=0.033) (Table 1).

Association between MTTP -164T>C polymorphism and incident CVD

The association between the *MTTP* -164T>C polymorphism and CVD events was tested also for MI and IS separately, taking into consideration the additive, dominant and recessive models (Table 2). Since there were no sex differences in the association between *MTTP* -164T>C and CVD (P for interaction = 0.86), we combined men and women in all analyses. After adjustment for age and sex, Cox regression analyses revealed no significant association between the -164T>C variant and CVD risk considering the additive ($HR_{additive}$: 1.04, 95% CI: 0.86 to 1.25; P=0.714), dominant ($HR_{CT+CC\ vs\ TT}$: 1.09, 95% CI: 0.85 to 1.39; P=0.505) and recessive ($HR_{CC\ vs\ CT+TT}$: 0.90, 95% CI: 0.55 to 1.46; P=0.662) models. Nevertheless, the multiplicative and additive interactions between the -164T>C polymorphism and total cholesterol (dichotomous) in relation to CVD risk were significant and in the same negative direction ($\beta_{multiplicative\ interaction}$ = -0.55 ± 0.19; P=0.004; $SI_{additive\ interaction}$ = 0.31, 95% CI: 0.16 to 0.62 and $RERI_{additive\ interaction}$ = -1.44, 95% CI: -2.37 to -0.51).

Stratified analyses according to low (<200 mg/dL) and borderline-high/high (≥200 mg/dL) cholesterol levels showed significant positive associations between *MTTP* -164T>C and CVD in subjects with cholesterol levels <200 mg/dL, considering both the additive ($HR_{additive}$= 1.38, 95% CI: 1.07 to 1.78; P=0.014) and the dominant models (HR_{CT+CC}= 1.76, 95% CI: 1.22-2,54; P=0.002) (Table 2). Analyzing MI and IS separately, the associations seemed to be stronger for stroke ($HR_{additive}$:=1.60, 95% CI: 1.16 to 2,20, P=0.004; $HR_{Dominant}$= 2.22, 95% CI: 1.35 to 3.64, P=0.002) than for MI ($HR_{additive}$= 1.19, 95% CI: 0.82 to 1.72, P=0.353; $HR_{Dominant}$= 1.41, 95% CI: 0.85 to 2.33, P=0.186). However, results from the competing risk analysis (IS versus MI) provided a Wald test P value equal to 0.28 and 0.20, respectively, for the additive and dominant model. Further adjustment for other CVD risk factors (i.e. body mass index, waist circumference, prevalent hypertension, sport activity and alcohol consumption) led to similar HRs depicted in Table 2 (data not shown).

In subjects with cholesterol levels ≥ 200 mg/dL we observed a borderline inverse association between *MTTP* -164T>C and CVD in the additive model ($HR_{additive}$= 0.77, 95% CI: 0.58 to 1.03, P=0.075) and significant relationships in the dominant model ($HR_{dominant}$= 0.67, 95% CI:

0.48 to 0.94, P=0.021; $HR_{dominant}$= 0.65, 95% CI: 0.43 to 0.97, P=0.036, respectively, for CVD and MI)(Table 2).

We performed additional analyses by testing both the multiplicative and additive interactions for triglycerides, HDL- and LDL-cholesterol. The interactions between *MTTP*/triglycerides and *MTTP*/HDL-cholesterol in relation to CVD were not significant (P=0.18 and P=0.11, respectively), whereas they were significant and in the same direction as those found for total cholesterol when LDL-cholesterol was analyzed (multiplicative interaction: P=0.023; $SI_{additive\ interaction}$ = 0.33 95% CI (0.15-0.73) and $RERI_{additive\ interaction}$ = -1.17 95% CI (-2.01—0.33). Stratified analysis according to the 2 LDL-cholesterol categories (<130 and ≥130 mg/dL) showed an increased CVD ($HR_{additive}$: 1.24; 95% CI: 0.98 to 1.56; $HR_{dominant}$: 1.51; 95% CI: 1.09 to 2.08) and IS risk ($HR_{additive}$: 1.30; 95% CI: 0.96 to 1.75; $HR_{dominant}$: 1.66; 95% CI: 1.07 to 2.57) in the low LDL group when *MTTP* was considered in a dominant fashion. A decreased CVD ($HR_{additive}$: 0.80; 95% CI: 1.58 to 1.09; $HR_{dominant}$: 0.69; 95% CI: 0.47 to 1.00) and MI ($HR_{additive}$: 0.74; 95% CI: 0.51 to 1.07; $HR_{dominant}$: 0.62; 95% CI: 0.39 to 0.96) risk was observed, instead, in the high LDL group, always in a dominant fashion (data not shown).

In the replication cohort we did observe a trend toward a decreased CVD risk in individuals with cholesterol levels higher than 200 mg/dL ($HR_{additive}$= 0.60, 95% CI: 0.29 to 1.25; P=0.17). No association was observed in the other strata (<200 mg/dL) ($HR_{additive}$= 1.06, 95% CI: 0.33 to 3.40; P=0.92).

DISCUSSION

In this study, we anticipated an interaction between total cholesterol levels and the *MTTP* -164T>C polymorphism with regard to the CVD risk. The presence of a statistically significant interaction confirmed our hypothesis and indicated carriers of the C allele of the *MTTP* -164T>C polymorphism with plasma total cholesterol levels lower than 200 mg/dL had an increased risk of CVD. The association seemed to be stronger for IS than for MI, but differences in the associations were not supported by competing risk analysis. Conversely, the *MTTP* -164 C-allele showed a lower CVD, and MI, risk in participants with cholesterol levels higher than 200 mg/dL. Similar relationships were observed considering LDL-cholesterol with levels lower and higher than 130 mg/dl suggesting that LDL is the driving cholesterol

component. However, the value of LDL levels seems to be limited as estimated based on the Friedewald formula [46]. In fact further studies are needed to replicate these findings.

The association between *MTTP* –I128T polymorphisms and CVD risk observed in the replication cohort showed a similar trend within the strata of cholesterol levels higher than 200 mg/dL. However, considering that the number of cases in the Heinz Nixdorf Recall Study is small further replication studies are needed. To our knowledge, this is the first prospective study showing such an effect of *MTTP* on risk of IS.

With regard to the association between the *MTTP* -164T>C polymorphism and cholesterol levels, previous studies observed inconsistent results. Few studies reported a slight cholesterol lowering effect of the rare alleles of the *MTTP* promoter polymorphisms [7, 10]. Ledmyr et al. investigated the association between the *MTTP* -493 G/T polymorphism and cholesterol in both healthy and hyper-cholesterolemic individuals, and observed decreased levels of total cholesterol in carriers of the -493 T variant [7, 10]. Furthermore, Phillips et al. in a small study including 82 patients with type 2 diabetes mellitus (T2DM) of a Caucasian population found that the subjects heterozygous for the -493 G/T had lower LDL-cholesterol and, in the postprandial phase, higher apoB48 levels in the VLDL fraction. The authors suggested that the -493 G/T polymorphism seemed to confer protection against atherosclerosis in T2DM patients [12]. In contrast, Jou et al. observed that total cholesterol, LDL-, and non HDL-cholesterol levels were higher according to the rare allele of the *MTTP* -493 G/T polymorphism when disease free young African Americans were investigated [19, 20]. Further, Lundahl et al. observed lower serum triglyceride levels in subjects affected by familial hypercholesterolemia and homozygous for the rare allele of the *MTTP* -493 G/T genotype [8].

Overall, these studies seem to suggest that *MTTP* regulates lipids differently in the presence or absence of disease, although the occurrence of a possible interaction between the LDL receptor and the *MTTP* genes is not excluded [7, 14]. Our results seem to be in line with these hypotheses. On one hand we observed slightly higher total and LDL-cholesterol levels in subjects free of CVD and homozygous for the rare -164 C allele, and on the other, lower total cholesterol and triglyceride levels according to the rare allele of the *MTTP* -164T>C polymorphism in the group of future CVD cases.

It has been shown that the C–allele of the *MTTP*-164 T>C polymorphism is homologous to a putative sterol response element (SRE) binding site and as such confers a reduced *MTTP* expression [26-28]. These findings come from an experimental study in which Hagan et al. demonstrated that human *MTTP* promoter activity is up-regulated by cholesterol [27]. The mechanism based on which cholesterol regulates *MTTP* gene expression is linked to the presence of a modified SRE in the *MTTP* promoter [27]. When cholesterol levels are low, the sterol regulatory element binding protein (SREBP) acts as transcription factor, binds to the SRE thereby inhibiting *MTTP* gene expression [28]. In contrast, in presence of cholesterol the modified SRE likely binds a new SREBP family member thus up-regulating *MTTP* expression [27, 28]. These observations suggest that *MTTP* gene expression is differently regulated by high and low cholesterol levels.

Despite the lack of significant associations between the -493G>T or -164T>C single nucleotide polymorphisms (SNPs), coronary heart disease and blood lipids observed in two previous studies [24, 25], recently Aminoff et al. put forward that carriers of the rare -164C allele are at increased risk of IHD [26]. They substantiated their findings by showing *in vivo* that the presence of the rare alleles of the -493G>T and -164T>C SNPs confer lower *MTTP* transcription in the heart, liver and macrophage. This mechanism, in turn, by causing the lipid accumulation in the heart would provoke an increased IHD risk. Indeed, our findings are in line with those of Aminoff et al. though they concluded that the increased IHD risk observed according to the -164C variant was independent of plasma lipids. As mentioned above, because human *MTTP* promoter activity is positively regulated by cholesterol [27], it is reasonable to assume that subjects with low cholesterol levels have, in general, a lower *MTTP* gene expression. Thus, in this low risk group carriers of the -164C variant, compared to carriers of the common allele, might be at increased CVD because of their lower *MTTP* gene expression. At the same time, if one would consider the observed associations as those mimicking *MTTP* inhibitors, then these findings could further highlight the concerns expressed by Aminoff et al. regarding the long term side effects *MTTP* inhibition may generate [26, 47].

In contrast, in subjects with higher cholesterol levels we observed a reduced, though borderline significant, CVD risk accordingly to the *MTTP* gene -164 C variant. Our findings

suggest that there could be an antagonistic (qualitative) interaction between cholesterol levels and *MTTP* -164 T>C polymorphism. These observations warrant further investigation.

The main limitation of this study is that the plasma lipoprotein and apolipoprotein levels, which are important in the effect of *MTTP* in cardiovascular disease, were not measured; our analyses on triglyceride levels were based on both fasting and non-fasting subjects; we estimated the LDL-cholesterol levels based on Friedewald equation. Strength of our study includes its prospective design. Furthermore, all cases of MI and IS were validated by medical records and were derived from a cohort population with a very high follow-up coverage.

CONCLUSIONS

The findings of this study suggest that in the subjects investigated an interaction between *MTTP* -164T>C functional polymorphism with total cholesterol levels predisposes risk allele carriers with low cholesterol levels to an increased risk of developing CVD, which seems to be abolished among risk allele carriers with high cholesterol levels. However, further studies are warranted in order to shed more light on these complex mechanisms.

ACKNOWLEDGMENTS

We are indebted to Wolfgang Fleischhauer for case ascertainment and to Ellen Kohlsdorf for data management. We thank Kathrin Saar and Norbert Hübner for support and discussion.

Part of the Heinz Nixdorf Recall genotyping was funded by Celgene, Munich Germany. The authors thank the participants of the Heinz Nixdorf Recall study and the entire community of Heinz Nixdorf Recall investigators and staff for their support and valuable contributions.

Sources of Funding

The recruitment phase of the EPIC-Potsdam Study was supported by the Federal Ministry of Science (01 EA 9401), Germany, and the European Union (SOC 95201408 05F02). The follow-up of the EPIC-Potsdam Study was supported by the German Cancer Aid (70-2488-Ha I) and the European Community (SOC 98200769 05F02). The Heinz Nixdorf Recall Study was supported by the Heinz Nixdorf Foundation (Germany), the Sarstedt AG & Co. (Nümbrecht, Germany) concerning laboratory equipment, the German Ministry of Education

and Science (BMBF) and the German Research Foundation (DFG) within the program HO 3314_2-1.

The funders had no role in study design, data collection and analysis, the decision to publish or preparation of the manuscript.

REFERENCES

1. Wetterau JR, Lin MC, Jamil H: Microsomal triglyceride transfer protein. *Biochim Biophys Acta* 1997, 1345:136-50.

2. Hussain MM, Iqbal J, Anwar K, Rava P, Dai K: Microsomal triglyceride transfer protein: a multifunctional protein. *Front Biosci* 2003, 8:s500-6.

3. White DA, Bennett AJ, Billett MA, Salter AM: The assembly of triacylglycerol-rich lipoproteins: an essential role for the microsomal triacylglycerol transfer protein. *Br J Nutr* 1998, 80:219-29.

4. Gordon DA, Jamil H: Progress towards understanding the role of microsomal triglyceride transfer protein in apolipoprotein-B lipoprotein assembly. *Biochim Biophys Acta* 2000, 1486:72-83.

5. Hussain MM, Shi J, Dreizen P: Microsomal triglyceride transfer protein and its role in apoB-lipoprotein assembly. *J Lipid Res* 2003, 44:22-32.

6. Berthier MT, Houde A, Paradis AM, Couture P, Gaudet D, Després JP, Vohl MC : Molecular screening of the microsomal triglyceride transfer protein: association between polymorphisms and both abdominal obesity and plasma apolipoprotein B concentration. *J Hum Genet* 2004, 49:684-90.

7. Ledmyr H, McMahon AD, Ehrenborg E, Nielsen LB, Neville M, Lithell H, MacFarlane PW, Packard CJ, Karpe F; WOSCOPS executive: The microsomal triglyceride transfer protein gene-493T variant lowers cholesterol but increases the risk of coronary heart disease. *Circulation* 2004, 109:2279-84.

8. Lundahl B, Leren TP, Ose L, Hamsten A, Karpe F. 2000: A functional polymorphism in the promoter region of the microsomal triglyceride transfer protein (MTTP -493G/T) influences lipoprotein phenotype in familial hypercholesterolemia. *Arterioscler Thromb Vasc Biol* 2000, 20:1784-8.

9. Lundahl B, Hamsten A, Karpe F. Postprandial plasma ApoB-48 levels are influenced by a polymorphism in the promoter of the microsomal triglyceride transfer protein gene. *Arterioscler Thromb Vasc Biol* 2002, 22: 289-93.

10. 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* 2002:43:51-8.

11. St-Pierre J, Lemieux I, Miller-Felix I, Prud'homme D, Bergeron J, Gaudet D, Nadeau A, Despres JP, Vohl MC: Visceral obesity and hyperinsulinemia modulate the impact of the microsomal triglyceride transfer protein -493G/T polymorphism on plasma lipoprotein levels in men. *Atherosclerosis* 2002, 160:317-24.

12. Phillips C, Mullan K, Owens D, Tomkin GH. Microsomal triglyceride transfer protein polymorphisms and lipoprotein levels in type 2 diabetes. *QJM* 2004, 97:211-8.

13. García-García AB, González C, Real JT, Martín de Llano JJ, González-Albert V, Civera M, Chaves FJ, Ascaso JF, Carmena R: Influence of microsomal triglyceride transfer protein promoter polymorphism -493 GT on fasting plasma triglyceride values and interaction with treatment response to atorvastatin in subjects with heterozygous familial hypercholesterolaemia. *Pharmacogenet Genomics* 2005, 15:211-8.

14. Lundahl B, Skoglund-Andersson C, Caslake M, Bedford D, Stewart P, Hamsten A, Packard CJ, Karpe F: Microsomal triglyceride transfer protein -493T variant reduces IDL plus LDL apoB production and the plasma concentration of large LDL particles. *Am J Physiol Endocrinol Metab* 2006, 290:E739-45.

15. Karpe F, Lundahl B, Ehrenborg E, Eriksson P, Hamsten A: A common functional polymorphism in the promoter region of the microsomal triglyceride transfer protein gene influences plasma LDL levels. *Arterioscler Thromb Vasc Biol* 1998, 18:756-61.

16. Rubin D, Helwig U, Pfeuffer M, Schreiber S, Boeing H, Fisher E, Pfeiffer A, Freitag-Wolf S, Foelsch UR, Doering F, Schrezenmeir J: A common functional exon polymorphism in the microsomal triglyceride transfer protein gene is associated with

type 2 diabetes, impaired glucose metabolism and insulin levels. *J Hum Genet* 2006, 51:567-74.

17. Austin MA, Talmud PJ, Luong LA, Haddad L, Day IN, Newman B, Edwards KL, Krauss RM, Humphries SE: Candidate gene studies of the atherogenic lipoprotein phenotype: a sib-pair linkage analysis of DZ women twins. *Am J Hum Genet* 1998, 62:406– 419.

18. Sposito AC, Gonbert S, Turpin G, Chapman MJ, Thillet J. Common polymorphism in the MTTP promoter attenuates the dyslipidemic and proatherogenic effects of excess body weight. *Arterioscler Thromb Vasc Biol* 2004, 24:e143.

19. Böhme M, Grallert H, Fischer A, Gieger C, Nitz I, Heid I, Kohl C, Wichmann HE, Illig T, Döring F; KORA Study Cohort: MTTP variants and body mass index, waist circumference and serum cholesterol level: Association analyses in 7582 participants of the KORA study cohort. *Mol Genet Metab* 2008, 95:229-32.

20. Juo SH, Han Z, Smith JD, Colangelo L, Liu K. Common polymorphism in promoter of microsomal triglyceride transfer protein gene influences cholesterol, ApoB, and triglyceride levels in young african american men: results from the coronary artery risk development in young adults (CARDIA) study. *Arterioscler Thromb Vasc Biol* 2000, 20:1316-22.

21. Juo SH, Colangelo L, Han Z, Smith JD, Liu K: Confirmation of the microsomal triglyceride transfer protein genetic effect on lipids in young African American men from the CARDIA study. *Arterioscler Thromb Vasc Biol* 2003, 23:912-3.

22. Schgoer W, Eller P, Mueller T, Tancevski I, Wehinger A, Ulmer H, Sandhofer A, Ritsch A, Haltmayer M, Patsch JR: The MTTP -493TT genotype is associated with peripheral arterial disease: results from the Linz Peripheral Arterial Disease (LIPAD) Study. *Clin Biochem* 2008, 41:712-6.

23. Okumura K, A. Imamura A, Murakami R, Takahashi R, Cheng XW, Numaguchi Y, Murohara T. Microsomal triglyceride transfer protein gene polymorphism strongly

influences circulating malondialdehyde-modified low-density lipoprotein. *Metabolism* 2009, 58:1306-11.

24. Herrmann SM, Poirier O, Nicaud V, Evans A, Ruidavets JB, Luc G, Arveiler D, Bao-Sheng C, Cambien F: Identification of two polymorphisms in the promoter of the microsomal triglyceride transfer protein (MTTP) gene: lack of association with lipoprotein profiles. *J Lipid Res* 1998, 39:2432-5.

25. 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, 148:337-43.

26. Aminoff A, H. Ledmyr H, P. Thulin P, K. Lundell K, L. Nunez L, E. Strandhagen E, C. Murphy C, U. Lidberg U, J. Westerbacka J, A. Franco-Cereceda A, J. Liska J, L.B. Nielsen LB, Gåfvels M, Mannila MN, Hamsten A, Yki-Järvinen H, Thelle D, Eriksson P, Borén J, Ehrenborg E: Allele-specific regulation of MTTP expression influences the risk of ischemic heart disease. *J Lipid Res* 2010, 51:103-11.

27. Hagan DL, Kienzle B, Jamil H, Hariharan N: Transcriptional regulation of human and hamster microsomal triglyceride transfer protein genes. Cell type-specific expression and response to metabolic regulators. *J Biol Chem* 1994, 269:28737-44.

28. Sato R, Miyamoto W, Inoue J, Terada T, Imanaka T, Maeda M: Sterol regulatory element-binding protein negatively regulates microsomal triglyceride transfer protein gene transcription. *J Biol Chem* 1999, 274:24714-20.

29. Boeing H, A. Korfmann A, Bergmann MM. Recruitment procedures of EPIC Germany: European Investigation into Cancer and Nutrition. *Ann Nutr Metab*. 1999, 43:205-215.

30. Boeing H, Wahrendorf J, Becker N: EPIC-Germany--A source for studies into diet and risk of chronic diseases. European Investigation into Cancer and Nutrition. *Ann Nutr Metab* 1999, 43:195-204.

31. Bergmann MM, Bussas U, Boeing H: Follow-up procedures in EPIC Germany— data quality aspects: European Prospective Investigation into Cancer and Nutrition. *Ann Nutr Metab* 1999, 43:225-234.

32. Rundle AG, Vineis P, Ahsan H. Design options for molecular epidemiology research within cohort studies: *Cancer Epidemiol Biomarkers Prev* 2005, 14:1899-1907.

33. Weikert C, Stefan N, Schulze MB, Pischon T, Berger K, Joost HG, Härung HU, H. Boeing H, Fritsche A: Plasma fetuin-A levels and risk of myocardial infarction and ischemic stroke. *Circulation* 2008, 118:2555-62.

34. Prentice R: A case-cohort design for epidemiologic cohort studies and disease prevention trials. *Biometrika* 1986, 73:1-11.

35. Miettinen O: Design options in epidemiologic research. An update. *Scand J Work Environ Health*. 1982, 8:7-14.

36. Weikert C, Berger K, Heidemann C, Bergmann MM, Hoffmann K, Klipstein-Grobusch K, Boeing H: Joint effects of risk factors for stroke and transient ischemic attack in a German population: the EPIC Potsdam Study. *J Neurol* 2007, 254:315-21.

37. Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H, Mann K, Siffert W, Lauterbach K, Siegrist J, Jöckel KH, Erbel R. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. *Am Heart J* 2002, 144:212-8.

38. Nomenclature and criteria for diagnosis of ischemic heart disease. Report of the Joint International Society and Federation of Cardiology/World Health Organization task force on standardization of clinical nomenclature. *Circulation* 1979, 59:607-9.

39. Alexander RW, Pratt CM, Roberts R: Diagnosis and management of patients with acute myocardial infarction. In Alexander RW, Schlant RC, Fuster V (eds.). Hurst´s The Heart, Arteries and Veins. McGraw-Hill, New-York; 1998:1345-1433.

40. Bohlscheid-Thomas S, Hoting I, Boeing H, Wahrendorf J: Reproducibility and relative validity of food group intake in a food frequency questionnaire developed for the German part of the EPIC project. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997, 26 (Suppl 1):S59-70.

41. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001, 285:2486-97.

42. Andersson T, Alfredsson L, Källberg H, Zdravkovic S, Ahlbom A: Calculating measures of biological interaction. *European Journal of Epidemiology* 2005, 20:575-79.

43. Rothman KJ: The estimation of synergy or antagonism. *Am J Epidemiol* 1976, 103:506-11.

44. Lunn M, McNeil D: Applying Cox regression to competing risks. *Biometrics* 1995, 51:524-532.

45. Gauderman WJ MJ (2006): QUANTO 1.1: A computer program for power and sample size calculations for genetic-epidemiology studies. 2006, http://hydauscedu/gxe Last accessed: August 2012.

46. Nordestgaard BG, Benn M. Fasting and nonfasting LDL cholesterol: to measure or calculate? *Clin Chem* 2009;55:845-7. doi: 10.1373/clinchem.2008.123083.

47. Joy TR, Hegele RA: Microsomal triglyceride transfer protein inhibition-friend or foe? *Nat Clin Pract Cardiovasc Med* 2008, 5:506-8.

Table 1. Baseline characteristics of subjects according to cardiovascular disease status and *MTTP* -164 T>C genotype in EPIC-Potsdam Study.

		CVD free by genotype					CVD by genotype			
	All CVD free	TT	CT	CC	P for trend[a]	**All CVD**	TT	CT	CC	P for trend[a]
n	1 978	1 096	742	140		324	178	126	20	
age, yrs	49.8 ± 0.2	49.5 ± 0.3	49.2 ± 0.3	49.4 ± 0.7	0.543	54.9 ± 0.5	56.1 ± 0.6	54.1 ± 0.7	56.6 ± 1.7	0.223
Men, %	36.8	36.7	35.3	35.0	0.687	58.5	64.3	60.9	69.4	0.654
Abdominal obesity, %[b]	21.4	20.9	20.3	18.8	0.567	26.9	32.2	30.9	30.0	0.837
Obesity, %[c]	14.9	14.5	15.5	9.7	0.134	18.4	19.9	19.4	19.6	0.975
Hypertension, %	47.6	45.6	44.7	43.7	0.592	60.3	70.8	68.3	49.0	0.114
Current smokers, %	21.4	21.2	22.5	21.4	0.957	39.7	35.7	31.6	47.9	0.264
High education, %	41.7	44.2	38.7	44.4	0.978	34.1	31.8	30.3	36.1	0.699
High sport activity, %	24.8	25.2	24.4	27.5	0.230	19.9	18.0	13.0	23.5	0.534
Cholesterol, mg/dL	191 ± 0.9	189 ± 1.2	192 ± 1.5	195 ± 3.4	0.036	198 ± 2.3	210 ± 3.3	199 ± 3.9	199 ± 9.7	0.033
HDL-cholesterol, mg/dL	52 ± 0.3	52 ± 0.4	52 ± 0.5	52 ± 1.2	0.840	50 ± 0.8	50 ± 1.1	51 ± 1.3	47 ± 3.2	0.575
LDL-cholesterol, mg/dL	114 ± 0.8	112 ± 1.0	115 ± 1.2	116 ± 2.8	0.026	120 ± 1.9	128 ± 2.7	121 ± 3.2	129 ± 7.9	0.305
Triglyceride, mg/dL[d]	107 (105-110)	105 (101-108)	108 (104-112)	114 (104-124)	0.087	118 (111-125)	131 (120-144)	116 (104-129)	99 (76-129)	0.016
Alcohol consumption, g/day	8.0 (7.5-8.5)	8.2 (7.6-8.9)	7.5 (6.8-8.2)	9.5 (7.6-11.7)	0.871	5.9 (5.1-6.9)	5.5 (4.3-7.0)	5.8 (4.4-7.6)	3.9 (2.0-7.9)	0.663

[a]Determined from linear regression model adjusted for age and sex (where appropriate) in reference to CVD free and CVD by genotype.Age was adjusted for sex. Sex was adjusted for age. [b]Abdominal obesity was defined according to the ATP III criteria [41] based on the following waist circumference cut-off points: men ≥ 102 cm and women ≥88 cm.[c]Obesity was defined as BMI ≥30 kg/m^2. [d]Geometric means and 95% (CI) all such values.

Table 2: Hazard rate ratios (HR) and 95% confidence intervals (95% CI) for the associations between MTTP -164 T/C polymorphism, CVD (combined endpoint), MI and IS.

	Cholesterol < 200 mg/dL			Cholesterol ≥ 200 mg/dL		
	HR (95% CI)*			HR (95% CI)*		
	CVD	MI	IS	CVD	MI	IS
Cases, n	139	69	70	185	77	40
C allele	1.38 (1.07-1.78)	1.19 (0.82-1.72)	1.60 (1.16-2.20)	0.77 (0.58-1.03)	0.76 (0.54-1.07)	0.77 (0.48-1.23)
$P_{additive}$	0.014	0.353	0.004	0.075	0.113	0.273
Dominant	1.76 (1.22-2.54)	1.41 (0.85-2.33)	2.22 (1.35-3.64)	0.67 (0.48-0.94)	0.65 (0.43-0.97)	0.69 (0.40-1.18)
$P_{dominant}$	0.002	0.186	0.002	0.021	0.036	0.172
Recessive	0.73 (0.31-1.69)	0.71 (0.21-2.39)	0.75 (0.25-2.27)	0.98 (0.53-1.82)	0.99 (0.47-2.06)	0.93 (0.33-2.66)
$P_{recessive}$	0.458	0.580	0.616	0.948	0.969	0.892

[a]Adjusted for age and sex.

6

Automation of Food Questionnaires in Medical Studies: A state-of-the-art review and future prospects

Chapter 6

Automation of Food Questionnaires in Medical Studies: A state-of-the-art review and future prospects

Zoe Falomir [a,n], **Maria Arregui** [a,b], Francisco Madueno [a], Dolores Corella [c,d], Oscar Coltell [a,d]

a Department of Languages and Computer Systems, Universitat Jaume I, Castelló, Spain

b Department of Epidemiology, German Institute of Human Nutrition (DIfE), Potsdam-Rehbruecke, Germany

c Department of Preventive Medicine and Public Health, Faculty of Medicine, Universitat de Valencia, Spain

d CIBER Physiopatology of Obesity and Nutrition, Spain

Comput Biol Med. 2012 Oct;42(10):964-74

Impact Factor 2012: 1.089

ABSTRACT

Applications for automating the most commonly used dietary surveys in nutritional research, Food Frequency Questionnaires (FFQs) and 24 h Dietary Recalls (24HDRs), are reviewed in this paper. A comprehensive search of electronic databases was carried out and findings were classified by a group of experts in nutrition and computer science into: (i) Computerized Questionnaires and Web-based Questionnaires; (ii) FFQs and 24HDRs and combinations of both; and (iii) interviewer-administered or self-administered questionnaires. A discussion on the classification made and the works reported is included. Finally, works that apply innovative technologies are outlined and the future trends for automating questionnaires in nutrition are identified.

1. INTRODUCTION

Dietary factors are linked to the leading non-communicable causes of death [1]: cardiovascular diseases, some cancers, type 2 diabetes, etc. The study of the interaction between diet and the genome is crucial to prevent and treat these diseases. The assessment of a person's diet is a painstaking task which consists of analysing their daily intake over one or more years. However, in epidemiological studies of diet-disease association, this assessment is not feasible and, in practice, only a portion of the intake information is evaluated and then the habitual participants' intake is extrapolated. In order to obtain enough statistical power to avoid measurement errors and changes in diet, it is necessary to obtain repeated measures of dietary information from a large number of participants over time. For extracting information on participants' diet, nutritionists use Food Frequency Questionnaires (FFQ), 24 h dietary recalls (24HDRs), dietary records or dietary histories. These surveys collect data on consumed foods or dishes, which can be transformed into energy and nutrient intake using food composition tables (FCTs).

FFQs and 24HDRs are the most widely used tools to extract information on diet in the field of epidemiological studies. Both techniques assume that participants have some regularity in their diet and are able to quantify it.

FFQs ask participants to report their usual frequency of consumption of each food from a list and for a specific period of time [2]. They assess habitual consumption of foods or nutrients. FFQs present general questions such as 'Do you eat olive oil?' and if participants respond 'yes', the FFQs ask the frequency of consumption (i.e. 'How often do you eat olive oil? Units per day/week/month/year, etc.'). The FFQs that include portion-sizes of foods are also referred to as semiquantitative FFQs. The use of FFQs is widespread because of their advantages such as easy administration and translation into nutrients, and also because they can cover seasonal intake variations and foods of occasional consumption. However, the automation of nutrient calculation is intensive and requires considerable computing and nutritional expertise [3]. Many FFQs were developed for different purposes, from capturing usual intake among large population-based samples [4] to tailoring the questionnaire to measure intake of a particular nutrient/food/food group in small specialized samples such as: iron [5], omega-3 fatty acids [6], calcium [7], phytosterols [8], etc.

24HDRs ask the respondent to recall all the foods and beverages consumed in the preceding 24 h or day [2]. They usually use open-answer questions, such as 'List all the beverages you drank and all the foods you ate yesterday between midnight and midnight' (unstructured recall) or 'List all the beverages you drank and all the foods you ate yesterday for breakfast/lunch/dinner/snacks' or 'What did you eat when you woke up?' (meal based recall). 24HDRs are used to collect high-quality dietary data because: (i) they are based on short-term memory, (ii) they do not consist of a closed list of foods and (iii) they provide quantitative information rather than consumption ranges. Therefore they do not require adaptation to specific populations such as FFQs. A single 24HDR is not considered representative of an individual's usual diet, so multiple 24HDRs are preferred for many nutritional studies. They also require highly trained interviewers, thus 24HDRs are not considered economical or practical in research settings with large samples and FFQs are frequently used. If 24HDRs could be self-administered using computer technologies to substitute the interviewer, they could be more feasible for large-scale studies [9].

Dietary records are more precise than 24HDRs because food intake is registered at the time of the eating. However, dietary records present high respondent burden, high investigator cost, and an extensive training and motivation of participants [2]. Therefore dietary records are usually substituted by 24HDRs in nutrition studies. The same disadvantages are present in diet histories in large population nutritional studies because they collect information not only about the food frequency of intake but also about the typical makeup of meals [2]. They also include more than one intake survey, such as the combination of a 24HDR, a FFQ and 3-day diet records [10]. Moreover they sometimes involve difficult cognitive tasks for the respondents, are not quantifiably precise and can have a high investigator burden [2].

There is no ideal method of reference or gold standard for estimating the validity of a food survey. In practice, relative validations of a questionnaire (i.e. FFQ) are performed with respect to another questionnaire (i.e. 24HDRs) taking into account that the sources of error between the reference questionnaire and that evaluated must be as independent as possible. By comparing both surveys, correlation coefficients are obtained which indicate the validity of the instrument and the calibration coefficients to be applied for correcting further executions. This is the reason why studies usually combine FFQ and 24HDRs for obtaining results and validate them [11,12].

Traditionally, FFQs and 24HDRs were administered in paper (i.e. Harvard paper FFQs1). However, as information and communication technologies have gained importance in recent years, great efforts have been made to automate the questionnaires involved in epidemiological and other nutritional studies in order to save costs.

First, computer programs were developed for helping researchers to administer questionnaires to participants and to accelerate the extraction and processing of the important data from FFQs [13–15,3,6,16] and from 24HDRs [17–19]. Then, other software applications automated self-administered FFQs [20,21,5,22,6,7], 24HDRs [23–25], and combinations of both [26] were developed. And all these tools needed a specific computer system in order to function correctly.

When the World Wide Web became widespread, dietary Web-based questionnaires substituted computerized questionnaires for improving accessibility and for obtaining a multi-platform functionality, such as on-line FFQs [8,27–39], on-line 24HDRs [9,40–43] and combinations of both [11,44,45].

In literature there are numerous works concerning the development, validation, repeatability of FFQs and 24HDRs in epidemiological studies. However, there are fewer approaches describing engineering developments for the automation, evolution and acceleration of results extraction in epidemiological studies. Our contribution here is: (i) to provide a broad state-of-the-art review from a technological point of view of studies that used an automated FFQ or 24HDR, (ii) to compare them and discuss their characteristics, (iii) to present recent developments that use innovative technologies, and (iv) to outline future prospects.

A few similar reviews are found in the literature [46,47,1,48]. A very brief review on technologies applied to FFQ was presented by García-Segovia et al. [46]. A classification of Web tools and other computer applications used in nutrigenomic research was undertaken by Stumbo et al. [47]. They presented the most commonly used tools in US and Europe providing a description from the point of view of a researcher in nutrition. Long et al. [1] presented a review of the evidence on the effectiveness of technology-based methods for dietary assessment by reporting six technology-based methods. Ngo et al. [48] reviewed automated nutrition questionnaires and outlined some innovative methods for automating questionnaires, such as smart cards, personal digital assistants (PDAs) and mobile phones. In this paper, a wider period

of time is taken into consideration, a broader description is presented regarding computerized and Web-based FFQs and 24HDRs, a comparative discussion is provided on the automation and administration of questionnaires and also innovative technologies are reported. An engineering perspective of all the works is given, when available, and useful recommendations for automating questionnaires in nutrition are outlined.

The rest of this paper is organized as follows. Section 2 explains the methodology used for carrying out the state-of-the-art review and presents a classification of all the works found in the literature. Section 3 describes the computerized systems that implemented FFQs and 24HDRs, whereas Section 4 explains the Web-based FFQs and 24HDRs. Section 5 reports on recent innovative technologies and, in Section 6 a discussion is presented. Section 7 presents conclusions and future prospects.

2. METHODOLOGY

Two researchers undertook an independent review of nutritional studies in which the design, development and/or use of automated (computerized or Web-based) FFQs and 24HDRs were described. Electronic databases including PubMed, EMBASE and Web of Science were searched from 1980 to December 2011. Hand searches of published conference proceedings, key nutrition journals and reference lists of retrieved papers were also undertaken. The search terms used were based on the following titles/ topics and keywords/abstracts searches:

- Computer* food* questionnaire*, computer* 24 h* recall*,

- Web* food* questionnaire*, Web* 24 h* questionnaire*, Web* 24 h* diet*

- Internet* food* questionnaire, Internet* 24 h* recall*, Internet* 24 h* questionnaire, Internet* 24 h* diet* recall*.

The papers obtained were first classified from a computational science point of view into two study types: (1) Computerized Questionnaires and (2) Web-based or Internet-based Questionnaires. Secondly, those works were classified from a nutritional point of view into: (1) FFQs, (2) 24HDRs and (3) combinations of FFQs and 24HDRs. And finally, a last classification

was made regarding the type of administration to participants: (a) interviewer-administered or used by the researcher for interviewing the participants, and (b) self-administered or completed by the participants themselves. Tables 1 and 2 show the results of this classification.

3. COMPUTERIZED QUESTIONNAIRES

The advantages that computerized questionnaires offered with respect to paper questionnaires caused their progressive introduction into nutritional studies, such as:

- The higher speed of data collection because answers are automatically stored on databases saving the time required to enter data manually;

- The higher quality of data extracted since there is an immediate and automatic control for incomplete and implausible data;

- The direct data transfer to the centre of study (avoiding costs of printing, postage and data typing);

- The easy and plausible introduction of help wizards on portion size estimation and recognition of food using real color illustrations;

- The small extra cost of adding a few thousand participants to the study once the system for handling the questionnaires is developed and established.

These advantages were provided by computerized questionnaires appearing in literature which have been classified in Table 1 and described below. Section 3.1 presents works that used computerized FFQs, whereas Section 3.2 describes computerized 24HDRs. Section 3.3 presents one computerized FFQ and 24HDR.

3.1. Computerized food frequency questionnaires (C-FFQs)

Several research studies describing the development and application of computer programs for automating data extraction from FFQs can be found in the literature [13,20,21,14,5,15,22,

3,6,16,7]. From the point of view of engineering technologies, perhaps some of the most recent and relevant studies would be the following:

1. CAFE [3]: Compositional Analyses from Frequency Estimates program, which automated the FFQ developed for the European Prospective Investigation into Cancer and Nutrition Study in the United Kingdom (EPIC-Norfolk and EPIC-Oxford). CAFE was written in SAS3 linked to an Oracle4 relational database which allowed predefined entries (numeric codes for frequency responses) and free text entries for registering new food items and nutrients. CAFE uses the EPIC nutrient database and the food composition table by McCance and Widdowson [49,50].

2. NUTRISOL [16]: a nutritional freeware program for analysing dietary food intakes and translating them into nutrients using the food composition tables (FCTs) developed by the Spanish National Research Council (CSIC) [51] and also using domestic commonly used measures for local dishes description. NUTRISOL was developed on MS Access using Visual Basic 6.0 and ran under Windows OS. It was PC-compatible and it needed a processor running at least at 200 MHz, 16 MB of RAM and 80 MB of hard disk free space. The results produced were exportable to statistical programs such as MS Excel and the Statistical Package for Social Sciences or SPSS5 v. 11. NUTRISOL was used for analysing the diet of some populations in Málaga and Cádiz (south Spain) through studies carried out by the Nutrition and Food Technology Institute6 at Granada University (Spain) and by the University Clinic Hospital in Málaga (Spain). The computerized FFQ developed also allowed the FCTs to be adapted/changed for other populations/countries.

Other studies also used computerized FFQ but technological descriptions of the design and development of their programs were not available:

i. RIBEPEIX [6]: a c-FFQ designed to quantify the intake of fish and shellfish omega-3 fatty acids and chemical contaminants in a general population. A MS Access database was created with the concentrations of fatty acids and pollutants of each sea food.

ii. Heath et al. [5] developed and validated a computerized FFQ in order to estimate the iron, vitamin C and calcium consumption of a population of young adult New Zealand women using the New Zealand Food Composition Database [52], and their phytate consumption using the Canadian food composition database [53]. Results were exportable to SPSS v. 6.1.1.

iii. Vandelanotte et al. [22] developed and validated a computerized FFQ to estimate fat intake in Belgium using the Belgian [54] and Dutch [55] food composition tables. Results were exportable to SPSS v. 11.

3.2. Computerized 24 h dietary recalls (C-24HDRs)

In the literature, few research studies on the development of computerized 24HDRs have been published [17,23,18,19,24,25]. Chronologically, these studies are:

1. EPIC-SOFT [17]: a computerized questionnaire developed to obtain standardized 24HDRs between the nine countries involved in the European Prospective Investigation into Cancer and Nutrition (EPIC). This program was written in Clipper, it ran on MSDOS v.3 and it was PC IBM-compatible needing 2MB of RAM and 5MB of hard disk free space. The nutrient databases implemented in the system and used to calculate energy and macronutrients were temporary FCTs derived from national values.

2. FIRSSt [23]: the Food Intake Recording Software System designed for use with fourth-grade children in USA. It used interactive multimedia to facilitate a child's self-report of diet by simulating a multiple pass 24HDR. FIRSSt included a computerized tutorial about how to use the program and it organized foods within a group hierarchy in which commonly consumed foods were also included. FIRSSt was programmed in Director 6.5, Foxpro was used to create a database, and data were downloaded into MS Access. The Continuing Survey of Food Intakes by Individuals food coding system was used to identify foods and identify amount of foods in food groups for mixed dishes.

3. The US Department of Agriculture (USDA) is the world leader in 24HDR methodology having used 24HDRs as the primary dietary assessment method in American surveys of food consumption since 1965. Their Dietary Intake Data System (summarized by Raper et al. [18]) is composed of three separated computer systems: (1) the Automated Multiple Pass Method (AMPM) for collecting food intakes using a 24HDR programmed using a MS Access database and Blaise programming language which was used by Moshfegh et al. [19] for assessing nutrient intake and appropriateness from the dietary interview component of the National Health and

Nutrition Examination Survey (NHANES): What We Eat In America (WWEIA); (2) the Post-Interview Processing System (PIPS) for reformatting data and assigning food codes; and (3) the Survey Net system for final coding and editing, quality review, and nutrient analysis. The Food and Nutrient Database for Dietary Studies (FNDDS) is used by the AMPM to calculate the food decomposition in nutrients. The FNDDS includes food descriptions, food portions and their weights, and their corresponding nutrients and it is reviewed periodically by dietary coders who make corrections if needed and assign codes to any unmatched responses.

4. YANA-C [24]: the "Young Adolescents Nutrition Assessment on Computer" is a self-administered 24HDR structured in six meal occasions including 18 food groups and a 19th group for the items not listed in the menu. It also included photographs for portion size estimation which changed every time a participant pushed the 'more' or 'less' button and other information such as 'Don't forget mayonnaise if you ate French fries...'. YANAC was developed using MS Visual Basic v. 6.0 and the total energy and nutrient intakes were computed using the Unilever Becel Nutrient Calculation Program v. 5.037 and the Belgian [54] and Dutch [55] FCTs. The YANA-C was also applied to 3 SAS SW v.8, SAS Institute Inc., Cary, NC, USA. Healthy Lifestyle in Europe by Nutrition in Adolescence or the HELENA study [56,57]. The feasibility of self-administration was analyzed [56] and compared to administration by a dietitian concluding that both methods were comparable.

5. NDSR [25]: Nutrition Data System for Research software by the Nutrition Coordinating Center (NCC) at University of Minnesota is a Windows-based dietary analysis program designed for the collection and analyses of 24HDRs and the analysis of food records, menus, and recipes. Calculation of nutrients occur immediately providing data per ingredient, food, meal, and day in report and analysis file formats. The software includes a dietary supplement assessment module, so that nutrient intake from both food and supplemental sources may be captured and quantified. NDSR may be installed on standalone computers or in a network environment and is supported on Microsoft Windows 7/Vista/XP. The minimum recommended hardware for NDSR 2011 on Windows 7 are: 2.2 GHz of processor speed, 2 GB of RAM, and 36 GB of hard disk free space. The NDSR uses the Food and Nutrient Database developed by the NCC [58], which includes over 18,000 foods, 7000 brand products, 162 nutrients, etc.

The NDSR has been used in several nutrition studies. For example, in order to increase portion size estimation accuracy, Toobert et al. [59] developed an interactive CD-ROM program to estimate fruit, vegetable and fat intake in USA populations from two dietary screeners: the National Cancer Institute's revised Fruit and Vegetable Scan [60] and the Block Fat Screener [61]. They validated that CD-ROM program comparing it to a computerized 24HDR carried out on participants by expert dietitians using the NDSR software v. V2007 and V2008.

3.3. Systems integrating both: C-FFQs and C-24HDRs

In the literature, CAPI [26] a computer-assisted personal interviewing seafood consumption survey tool is presented. The CAPI contains a 24-h recall and food frequency questionnaire, and assesses seasonal seafood consumption and temporary changes in consumption. It was developed from existing the Pacific NW Native American seafood consumption survey methodology. The CAPI is used with a booklet of harvest locationmaps and species and portion size images. Moreover, it is used by tribal interviewers reducing potential bias. CAPI survey software was developed using Microsoft Access 2000. Tribes are provided with a supervisor's software module to configure the CAPI for their tribe and they can select fish and shellfish consumed by a tribe from a prepopulated species library included in the CAPI. The supervisor can preview and test the customized CAPI and then deploy the survey software to one or more interviewer computers. The supervisor can add new species to the database that were identified during interviews with tribal members. Data can be exported for further analysis using statistical software.

4. WEB-BASED OR INTERNET-BASED QUESTIONNAIRES

In the last few years, a considerable number of research works in nutrition that use Web-based or Internet-based questionnaires have appeared in the literature. The significance of these questionnaires increased with the broadcasting of information and communication technologies. Web-based questionnaires provide all the advantages of computerized questionnaires and others such as:

- The higher compliance, provided by more flexibility of completion at any time and location, and the personalized feed-back and interactive help features provided (self-administered);

- The ability to communicate with geographically dispersed research groups (interviewer-administered) or populations (self-administered), potentially internationalizing research, and groups often difficult to sample.

The nutritional Web-based questionnaires most widely referenced in the literature can be classified as Table 2 shows into: (i) Web-FFQs, described in Section 4.1, and (ii) Web-24HDRs, described in Section 4.2.

4.1. Web-based food frequency questionnaires (Web-FFQs)

With the increased number of research studies using Web- FFQs in the last few years [8,28–30,32,35–39] their validity and reproducibility has also been evaluated [27,31,33,34]. From the point of view of engineering technologies, some of the most interesting developments that appeared chronologically in the literature are the following:

1. FITUVEROLES [8]: a Web service for calculating phytosterol intake in a Mediterranean population using a Web-FFQ developed by physicians, computer science engineers and other researchers in biomedicine for estimating the intake of phytosterols of the participants from previously validated FFQs that included 101 food items divided into 11 food groups. Results are calculated using three nutrient data sources for comparing and providing the robustness of the estimation: the Finnish Food Composition Database [62], the National Nutrient Database for Standard Reference by USDA [63] and the data provided by Jiménez-Escrig [64] for Spanish foods. FITUVEROLES was programmed using PHP, XHTML and JavaScript and the database engine was MySQL. The results produced were exportable to SPSS.

2. DHQ [30]: the Diet History Questionnaire funded by the American National Cancer Institute (NCI) with their DHQ Nutrient Database and data analysis used in cancer research. The DHQ is a food frequency questionnaire (FFQ) which consists of 124 food items and includes both portion size and dietary supplement questions. Beasley et al. [31] added pictures to this DHQ in order to

represent portion sizes and they developed the Web-based pictorial diet history questionnaire (PDHQ).

3. Galante and Colli's [28] on-line semi-quantitative FFQ to evaluate calcium and iron intake. The contents of the study were stored in a database with a copy on a CD-ROM. To ensure the confidentiality of the information and the privacy of the individuals, access to the communication tool was made available through an individual password, and data transmission was performed through personal e-mails. The statistical analysis of the results was carried out using the Virtual Nutri SW.8

4. Apovian et al. [33] developed and validated a Web-FFQ for capturing food servings in the Dietary Approaches to Stop Hypertension or DASH diet recommended by the US Department of Agriculture (USDA). It was implemented on HTML using check boxes and entry fields for self-administration and it was placed on a secure Web server at Boston University's Data Coordinating Center.

5. The Web-based assessment tools in Epidemiology9 by the German Institute of Human Nutrition Potsdam-Rehbrücke (DIfE) [36]: (1) EPIC Potsdam FFQ; (2) European FPQ [65] containing 116 foods organized into 10 families and graphics for displaying portion sizes; (3) German FPQ; (4) European- Hellenic FPQ. They are usually self-administered or telephone-assisted. The participants must log into the Web service with the user name (ID) and password they received from the study office so that they are connected to the right questionnaire. Then they receive further information about filling in the online forms and who to ask in case of needing assistance. These Web-FFQs are used in the study of chronic diseases such as cancer, cardiovascular diseases or diabetes.

6. The VioFFQ by Viocare enterprise [38]: a self-administered Web- FFQ that collects data on dietary behavior and food patterns, estimates nutrient intake, and delivers a dietary change report. Dietary analysis is undertaken using the food and nutrient information from the Nutrition Coordinating Center (NCC) Food and Nutrient Database by the Division of Epidemiology and Community Health at the University of Minnesota. The VioFFQ was used in several epidemiological studies, for example, the recent study by McDaniel et al. [66] on polyunsaturated fatty acid ratios and their association with numerous chronic inflammation-related diseases.

7. The Block FFQ [39] by NutritionQuest [67]: a turn-key system, which provides an on-line structure for integrated data collection, nutrient and physical activity analysis, and data management. Questionnaires can be interviewer- or self-administered. Block FFQ accessed on-line and self-administered was used by Anderson-Bill et al. [68] for estimating the intake of fat, fiber, fruits and vegetables of the participants of a study for examining the behavioral characteristics of Web-health users.

Other important studies that use Web-FFQs, but which do not provide technological descriptions of the design and development of their programs are also available:

i. Matthys et al. [27] implemented and validated a Web-FFQ for analysing adolescents' food habits and proved its validity and reproducibility. The security of the system was achieved by providing a login name and a session-specific password to the participants. This FFQ contained questions on the average consumption of 69 food items during the past month and the food items listed were based on the classification system described in the Flemish Food Guide, commonly referred as the Food Triangle [69]. The energy intake was calculated [70] using the Belgian [54] and the Dutch [55] FCTs;

ii. RIBEFOOD [29] developed by the same authors of the computerized program RIBEPEIX [6] determines the dietary intake of a number of chemical contaminants (i.e., metals, dioxins and furans, PCBs, polycyclic aromatic hydrocarbons, etc.) using a FFQ carried out on a Web service;

iiii. Probst et al. [32] developed a Web-FFQ which automated the sequence of questioning that a dietitian usually takes with a client when conducting a diet history interview and they also connected aWeb camera to the computer for video-recording and analysing the participants' behaviors related to the type of foods appearing in the questionnaires;

iv. Vereecken et al. [35] developed a Web-FFQ for the Healthy Lifestyle in Europe by Nutrition in Adolescence or HELENA Study [56]. The reproducibility of this FFQ was analyzed in comparison to the YANA-C computerized 24HDR [24] and it has been recently validated by Maes et al. [71].

4.2. Web-based 24 h dietary recalls (Web-24HDR)

Studies using automated 24HDRs and on-line technologies appearing in the literature [9,40–43] are:

1. ASA24 [9,72] a Web-based Automated Self-Administered 24HDR developed by the American National Cancer Institute (NCI) which was inspired by the USDA's Automated Multipass Method or AMPM [18,19] and that incorporated and extended the Food Nutrient Database for Dietary Studies or FNDDS [18] and the food portion pictures of Food Intake REcording SW System or FIRSSt [23].

2. SNAPTM or the Synchronized Nutrition and Activity Program [40]: a Web-based software designed and evaluated for obtaining self-reported 24HDRs and physical activity questionnaires in school children. SNAPTM was written using PHP v. 4.0.1, MySQL v. 3.22 and JavaScript v. 1.3 to ensure data transfer and storage and its interface is intuitive and user-friendly for children to use. The dietary intake was measured using counts for 21 food groups and the list of commonly consumed foods and drinks was developed from a combination of findings from the National Diet and Nutrition Survey [73]. A free-text option box was also included to capture any unlisted food or drink. The data analysis was carried out using STATA10 v.8. software and the accuracy of the computer tool was provided by a Passing-Bablok method comparison technique using the Analyze-it11 software.

3. DietDay [41,74]: a fully automated, self-administered, Web- 24HDR inspired on the Automated Multiple Pass Method (AMPM) by USDA which is consisting of 9349 foods and over 7000 food images in 61 modules. Portion sizes are quantified by household measures using images of different amounts of food on a standard plate, glass, or bowl. Food preparation methods are also assessed, as well as condiments and additions. In DietDay usual consumption is asked by time of day. Nutrient values in the program are based on USDA values and expanded to include mixed dishes and product labeling information.

4. Web-SPAN [42]: a Web-Survey of Physical Activity and Nutrition presented and evaluated the appropriateness of a Web-based method for assessing dietary intake and physical activity in young students. The Web-based 24HDR data was transferred to ESHA Food Processor [75] and

the Canadian Nutrient File database [76] was used for calculating the macronutrient and micronutrient of the intakes.

5. The Oxford WebQ [43]: a self-administered Web-24HDRs which estimates nutrients automatically, providing a low-cost method for measuring dietary intake in large-scale studies.

4.3. Systems integrating both: Web-FFQs and Web-24HDRs

Approaches that integrate Web-FFQs and Web-24HDRs [11,44,45] are:

1. Hanning et al. [11] developed a Web-based Food Behaviour Questionnaire (FBQ), which included a 24HDR and a FFQ, which analyzed nutrients using ESHA Food Processor [75] and the Canadian Nutrient File database [76]. This FBQ was used for assessing dietary energy as a function of gender and weight status among Ontario and Alberta adolescents [77]. Later this was validated with students of grades six to eight obtaining positively correlated energy and nutrient intakes between FFQ and dietitian interviews [12].

2. OBENUTIC [44]: This Web service was developed to assist in the research of the relation between food intake and obesity and related diseases in Spanish Mediterranean populations. It provides dynamic generation of FFQs and 24HDRs according to the study defined. OBENUTIC uses, as nutrient data sources, either one of the following four FCTs or combinations of them: the Spanish tables by Mataix [78] and CESNID [79], the Finnish Food Composition Database [62] and the USDA's National Nutrient Database for Standard Reference [63]. A nutrient and food ontology, NutriOntologia, was developed for the alignments between food names [80,81] in each database. Results were customizable and exportable in a format suitable for SPSS. OBENUTIC was programmed using PHP, XHTML and JavaScript languages. The database engine was MySQL and the framework XAJAX in PHP was used for developing interfaces.

3. my DIDeA [45]: a Web-based dietary intervention for type 2 diabetes patients which included a semi-FFQ for recording participants' dietary intake and two days 24HDR for analyzing their nutrient intake using the Axxya Systems Nutritionist ProTM Diet Analysis.12

5. INNOVATIVE TRENDS

In the symposium by the European Nutrigenomics Organization (NuGO) [82] the following innovative applications used in nutritional studies were presented:

1. The ACASI [83,84], an audio computer-assisted program for interviewing participants so that they fill in a diet history questionnaire while listening to the questions;

2. The IMM [85], an Interactive Multimedia computerized recall that incorporates a touch screen and audio functions to help data input;

3. The Wellnavi [86,87], a Pocket-PC with digital camera and mobile phone for direct electronic data transfer to a dietitian in the study centre;

4. The Technology Assisted Dietary Assessment (TADA) project [88–90] used a mobile device with an embedded camera to estimate daily food and nutrient intake from digital images taken by the participants. Food recognition in digital images is achieved using segmentation methods combined with Gabor filters for texture comparison and statistical pattern classification techniques such as Support Vector Machines (SVMs). Food portion size is estimated in cm3 and the Food and nutrient Database for Dietary Studies (FNDDS) is used for obtaining weight measures. Moreover, X-ray computerised microtomography (XMCT), 3D laser imaging, and other techniques are used to measure food density (g/cm3) to convert portion estimates to a weight measure.

Other innovative applications can be found in the literature [91,92]. Image-Diet Day [91] is an automated image-capture method to aid dietary recall consisting of a user-initiated camera-equipped mobile phone programmed to automatically capture and transmit images to a secure website in conjunction with computer-assisted and multipass 24HDRs. Participants used the device on three independent days and then they used the captured images using ImageViewer Software while completing the 24HDR on the following day. Image processing filters successfully eliminated underexposed, overexposed and blurry images. Participants concluded that the images were helpful but that wearing the device around their neck was not easy and they also had problems such as limited battery life, self-consciousness about wearing the device in public and concerns about the field of view of the camera. Therefore, there is still work to be

done to solve these problems and to meet the challenge of managing the thousands of images generated. Lambert et al. [92] tested how useful smart cards were for assessing food intake. They provided smart cards to school children and when they paid for their meal, the foods on the tray were recorded at the cash desk and sent to a central computer. Moreover, the date and time of the meal was stored and also the data on the computer can be linked to a nutrient database. However, they observed that this method was useful for collecting information about participants' food selection but not for their food intake because they sometimes exchanged trays or paid each other's meals and the real portion of the intake was not known and it had to be estimated.

6. DISCUSSION

After analysing all the works appearing in the literature related to the automation of questionnaires in nutrition, a discussion is presented comparing: (i) automated vs. paper questionnaires (Section 6.1); (ii) computerized vs. Web-based questionnaires (Section 6.2); (iii) interviewer- vs. self-administered questionnaires (Section 6.3); and (iv) Web-based questionnaires vs. Innovative technologies (Section 6.4).

6.1. Automated vs. paper questionnaires

In a comparison between the use of a Web-questionnaire and a similar printed questionnaire, Balter et al. [93] found that the willingness to answer the second part of the questionnaire was higher with a Web-questionnaire than with the printed questionnaire, which suggests that the participants responding to the Web-based questionnaire found the process more appealing than those who responded to the printed questionnaire. Therefore, they recommended the use of Web-questionnaires if Internet access is high for the population involved. Another study [94] for comparing Web vs. paper questionnaires in adolescents concluded that most of the participants required less time to fill in the Web-questionnaire and they also felt less observed and more independent while completing it. Finally, a comparison of Web and paper versions of self-administered questionnaires used in the NutriNet-Santé Study [95,96] concluded that the quality

of information provided by the Web-questionnaire was equal to, or better than, that of the paper version, with substantial logistic and cost advantages.

6.2. Computerized vs. web-based questionnaires

Administration of paper questionnaires started to be replaced in the early eighties by researcher-administered computerized questionnaires. These tools helped researchers to store dietary information and to calculate nutrients and energy intake in an easier way. At the beginning of 2000, the first Web-questionnaires started to be available. They offered the advantages that they were accessible around the world by means of Internet, and that information could be stored in a centralized Web-server. Due to the spread of the WWW, these new applications allowed the possibility of self-administration of questionnaires to a wide range of participants. Furthermore, since the Semantic Web gained importance, developments of ontologies related to food intake, recipes and nutrition have been developed. Pattern recognition algorithms were also improved allowing the introduction of free entrances in 24HDRs questionnaires because the contents of the answers were automatically interpreted. Also, Web-based questionnaire developments improved in their portability, allowing their administration on standard PCs, PDAs, tablet computers or mobile phones only using Internet connection.

6.3. Self-administered vs. interviewer-administered questionnaires

Interviewer-administered questionnaires avoid excluding from the studies segments of population without Internet access or ability to use computers. Moreover, as interviewer-administered questionnaires are made in front of the participants, an expert dietitian is able to recognize nonverbal behavior that potentially may aid in minimizing social desirability bias. As Probst et al. [32] concluded, recognizing that a client is shifting in the chair, touching their face, etc. when asked about dessert, for example, may assist in determining the appropriate types of questions to be asked and whether or not to ask additional probing questions in order to gain more detail.

On the other hand, as the WWW increasingly spread in recent years, self-administered questionnaires are becoming more popular. A comparison between interviewed-administered

24HDRs and interactive Web-based self-administered 24HDRs [97] concluded that the latter permit considerable logistic simplification and cost saving and that they may be highly advantageous for large population-based surveys. Moreover, Web-based self-administered questionnaires can also be completed at any time from any location with centralized monitoring of participant completion. However, from the point of view of information technologies, a self-administered Web-based questionnaire must be developed more accurately because it is not completed by an expert dietitian or computer scientist. Therefore, some considerations have to be taken into account:

(1) The participants must receive information regarding: (a) the Web address, and (b) a unique ID-number and a secure password for each participant in order to ensure the confidentiality of the data provided to the system. This information can be provided by e-mail to accelerate the process of data gathering.

(2) The interface must guide all the participants equally and minimize potential errors derived from their varying degrees of knowledge. For that, help screens, a tutorial or an animated agent who guides participants through the Webside can be used.

(3) Including colour photographs showing different portion sizes per food items or/and dishes to obtain trustworthy information regarding food portion-size.

Moreover, Web-based self-administered questionnaires are not usually designed to capture the social desirability of a particular food. However, they may incorporate video recordings of participants' body movement and facial expression at the time of food reporting. Later, this video may be processed by intelligent systems for identifying such desirability. An initial study described the differences in the observed behaviors according to the type of foods selected by participants using a prototype version of a Web-based video-recorded dietary assessment [32].

6.4. Web-based questionnaires vs. new innovative technologies

Some innovative technologies such as personal digital assistants (PDAs) with camera, dictaphones or mobile phones are interesting because they allow the participants to store electronic food records in the moment of eating.

However, as Illner et al. [98] mentioned: portable food records (i.e. the mobile phone with camera by Boushey et al. [88]) are innovative ideas suitable for reaching low-literacy groups, adolescents and elderly, but like all new techniques on food records, application in large-scale studies is hindered by costs and training burden. Moreover, new technologies such as mobile phones with camera or portable barcode readers with PDAs are more useful for capturing diet snapshots (detailed information on dietary practise at a certain point in time, such as single meals or purchases) than for capturing complete dietary sequences like diet throughout a day.

Finally, the identified future trends for administration of questionnaires from the point of view of computer engineering and nutrition are: (1) using mobile devices such as mobile telephones, PDA-like devices or tablets with a built-in camera, network connectivity, and a microprocessor for integrating image analysis, visualization tools, etc.; (2) programming nutritional assessment tools programed as mobile phone applications which the participants may download from public markets on their personal mobiles and use for self-administration; (3) introducing multimedia functions in questionnaires automated for mobile devices such as: selection of foods by touching the screen, naming the food appearing in photographs using a handwriting-note application; speech synthesizers for reading the questionnaire, etc.; (4) introducing computer vision technologies for recognition of foods (i.e. Gabor filters and Support Vector Machine algorithms, etc.), for calculating food portions and weight in the photos made by the participants' mobile devices; (5) using computer vision methods for identifying the barcode of some foods difficult to decompose in components (i.e. precooked foods) in order to obtain the specification of the exact ingredients used by the manufacturer; (6) using pattern recognition algorithms and language processing techniques for identifying the contents of free text entrances; (7) developing and using ontologies for assigning meaning (interpretable byWeb agents) to the food, recipes and other nutrition information extracted from the questionnaires; (8) cameras (in portables or mobile phones) for video-recording and recognizing participants' behaviors when completing questionnaires with the aim of minimizing social desirability bias.

7. CONCLUSIONS AND FUTURE PROSPECTS

FFQs and 24HDRs dietary recalls have been automated in a variety of ways in different studies. Depending on the aim of the study and on the population analyzed, a method for automating nutrition questionnaires would be more suitable than others. In this state-of-the-art review, nutritional studies conducted or published since 1980, in which the use of automated (computerized or Web-based) nutrition questionnaires, have been described. In order to identify relevant studies that describe the design, development and/or use of food questionnaires, a comprehensive search procedure was developed using the electronic databases PubMed, Embase and Web of Science and searches of published conference proceedings, key nutrition journals and reference lists of retrieved papers. The works obtained were classified from a nutritional point of view and regarding the type of administration to participants. Several studies that apply innovative technologies have been outlined too. And finally, a discussion has been provided on the classification made and the works reported.

Table 1 Classification of Computerized FFQs, 24HDRs, and combination of both. The works in each box are listed chronologically

Type of Administration	C-FFQs	C-24HDRs	C-FFQ + C-24HDRs
Interviewer-administered	Baghurst & Record [13]		
	Smith et al. [14]		
	Geekie et al. [15]	EPIC-SOFT [17]	CAPI [26]
	CAFE [3]	AMPM [18,19]	
	RIBEPEIX [6]		
	NUTRISOL [16]		
Self-administered	Engle et al. [20]		
	Suitor & Gardner [21]	FIRSSt [23]	–
	Heath et al. [5]	YANA-C [24]	
	Vandelanotte et al. [22]	NDSR [25]	
	Wong et al. [7]		

Table 2. Classification of Web-based FFQs, 24HDRs, and combination of both. The works in each box are listed chronologically.

Type of Administration	Web-FFQs	Web-24HDRs	Web-FFQ + Web 24HDRs
Interviewer-administered	FITUVEROLES [8]	–	OBENUTIC [44]
	RIBEFOOD [29]		
Self-administered	Matthys et al. [27]		
	Galante & Colli [28]		
	Probst et al. [32]		
	DHQ [30], PDHQ [31]	ASA24 [9,72]	
	Martı-Cid et al. [29]	SNAP [40]	FBQ [11]
	FFQ-DASH [33]	DietDay [41]	my DIDeA [45]
	Vereecken et al. [35]	Web-SPAN [42]	
	Labonté et al. [34]	Oxford WebQ [43]	
	EPIC-FFQs [36]		
	VioFFQ [38]		
	BlockFFQ [39]		

8 CONFLICT OF INTEREST STATEMENT

None declared.

9 ACKNOWLEDGMENTS

This work has been partially supported by: (1) Spanish Ministry of Science and Innovation (MICINN) under the project named GEWIMICS and referenced as SAF2009-12304/GEN; (2) the Cooperative Research Thematic Network (RETIC) on Computational Medicine (COMBIOMED) referenced as RD07/0067/0006 and funded by Carlos III Spanish Health Institute (ISCIII); (3) the Biomedical Research Centre Network (CIBER) named Fisiopatología de la Obesidad y Nutrición (CIBERobn) as an initiative of ISCIII; (4) Generalitat Valenciana under Grant numbers BEST/2010/211, BEST/2011/261, and ACOMP/2011/145; and (5) Spanish National Centre of Cardiovascular Research (CNIC) under project number CNIC06.

REFERENCES

[1] J. Long, L. Littlefield, G. Estep, H. Martin, T. Rogers, C. Boswell, B. Shriver, C. Roman-Shriver, Evidence review of technology and dietary assessment, Worldviews Evidence Based Nurs. 7 (4) (2010) 191–204.

[2] F.E. Thompson, A.M.Y.F. Subar, Dietary assessment methodology, ReCALL (2008) 12–15.

[3] A. Welch, R. Luben, K. Khaw, S. Bingham, The CAFE computer program for nutritional analysis of the EPIC-Norfolk food frequency questionnaire and identification of extreme nutrient values, J. Hum. Nutr. Diet. 18 (2005) 99–116.

[4] A. Subar, F. Thompson, V. Kipnis, D. Midthune, P. Hurwitz, S. McNutt, A. McIntosh, S. Rosenfeld, Comparative validation of the Block, Willett, and National Cancer Institute Food Frequency Questionnaires: the eating at America's table study, Am. J. Epidemiol. 154 (2001) 1089–1099.

[5] A.-L. Heath, C. Skeaff, R. Gibson, The relative validity of a computerized food frequency questionnaire for estimating intake of dietary iron and its absorption modifiers, Eur. J. Clin. Nutr. 54 (2000) 592–599.

[6] J. Domingo, A. Bocio, R. Martí-Cid, J. Llobet, Benefits and risks of fish consumption part II. RIBEPEIX, a computer program to optimize the balance between the intake of omega-3 fatty acids and chemical contaminants, Toxicology 230 (2007) 227–233.

[7] S. Wong, C. Boushey, R. Novotny, D. Gustafson, Evaluation of a computerized food frequency questionnaire to estimate calcium intake of Asian, hispanic, and non-hispanic white youth, J. Am. Diet. Assoc. 108 (3) (2008) 539–543.

[8] M. Arregui, O. Coltell, R. Vázquez, A. Fabregat, O. Portolés, D. Corella, FITUVEROLES: un portal Web piloto para la determinación de fitoesteroles ingeridos en la dieta mediante cuestionarios digitalizados, Public Health Nutr. 9 (2006) 255.

[9] A. Subar, F. Thomson, N. Potischman, B. Forsyth, R. Buday, D. Richards, S. McNutt, S. Hull, P. Guenther, A. Schatzkin, T. Baranowski, Formative research of a quick list for an automated self-administered 24-hour dietary recall, J. Am. Diet. Assoc. 107 (2007) 1002–1007.

[10] B.S. Burke, The dietary history as a tool in research, J. Am. Diet. Assoc. 23 (1947) 1041–1046.

[11] R.M. Hanning, L. Jessup, I. Lambraki, C. MacDonald, L. McCargar, AWeb-based approach to assessment of food intake and behaviour of school children and adolescents, Can. J. Diet. Pract. Res. 64 (2) (2003) 110.

[12] R.M. Hanning, D. Royall, J.E. Toews, L. Blashill, J. Wegener, P. Driezen, Webbased Food Behaviour Questionnaire: validation with grades six to eight students, Can. J. Diet. Pract. Res. 70 (4) (2009) 172–178.

[13] K. Baghurst, S. Record, A computerized dietary analysis system for use with diet diaries of food frequency questionnaires, Community Health Stud. 8 (1) (1984) 11–18.

[14] B. Smith, S. Morgan, et al., Reproducibility and comparability of a computerized, self-administered food frequency questionnaire, J. Am. Diet. Assoc. 99 (12) (1999) 1579–1581.

[15] M. Geekie, I. Kennedy, R. Holman, A computerized food frequency questionnaire to facilitate dietary modification in a UK population, DIABETES 51 (2) (2002) A471.

[16] M. Gutiérrez-Bedmar, J. Gómez-Arena, A. Mariscal, A. García-Rodríguez, E. Gómez-Gracia, M. Carnero-Varo, J. Villalobos, J. Navajas, NUTRISOL: a computer programme for communitary and hospital nutritional evaluation of free access, Nutr. Hosp. 23 (1) (2008) 20–26.

[17] N. Slimani, G. Deharveng, R. Charrondi_ere, A. van Kappel, M. Ocké , A. Welch, A. Lagiou, M. van Liere, A. Agudo, V. Pala, B. Brandstetter, C. Andren, C. Stripp, W. van Staveren, E. Riboli, Structure of the standardized computerized 24-h diet recall interview used as reference method in the 22 centers participating in the EPIC project, Comput. Methods Prog. Biomed. 58 (1999) 251–266.

[18] N. Raper, B. Perloff, L. Ingwersen, L. Steinfeldt, J. Anand, An overview of the USDA's dietary intake data system, J. Food Compos. Anal. 17 (2004) 545–555.

[19] A. Moshfegh, D. Rhodes, D. Baer, T. Murayi, J. Clemens, W. Rumpler, D. Paul, R. Sebastian, K. Kuczynski, L. Ingwersen, R. Staples, L. Cleveland, The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes, Am. J. Clin. Nutr. 88 (2008) 324–332.

[20] A. Engle, L. Lynn, K. Koury, A. Boyar, Reproducibility and comparability of a computerized, self-administered food frequency questionnaire, Nutr. Cancer 13 (4) (1990) 281–292.

[21] C. Suitor, J. Gardner, Development of an interactive, self-administered computerized food frequency questionnaire for use with low-income women, J. Nutr. Educ. 24 (2) (1992) 82–86.

[22] C. Vandelanotte, C. Matthys, I.D. Bourdeaudhuij, Reliability and validity of a computerized questionnaire to measure fat intake in Belgium, Nutr. Res. 24 (2004) 621–631.

[23] T. Baranowski, N. Islam, J. Baranowski, K. Cullen, D. Myres, T. Marsh, C. deMoor, The food intake recording software system is valid among fourth-grade children, J. Am. Diet. Assoc. 102 (2002) 380–385.

[24] C. Vereecken, M. Convents, C. Matthys, L. Maes, Young adolescents' nutrition assessment on computer (YANA-C), Eur. J. Clin. Nutr. 59 (2005) 658–667.

[25] NDSR, Nutrition Data System for Research software, Nutrition Coordinating Center, University of Minnesota, Minneapolis, /http://www.ncc.umn.edu/ products/ndsr.htmlS, 2011, accessed Jan 2012.

[26] L. Kissinger, R. Lorenzana, B. Mittl, M. Lasrado, S. Iwenofu, V. Olivo, C. Helba, P. Capoeman, A. Williams, Development of a computer-assisted personal interview software system for collection of tribal fish consumption data, Risk Anal. 30 (12) (2010) 1833–1840.

[27] C. Matthys, I. Pynaert, W.D. Keyzer, S.D. Henauw, Validity and reproducibility of an adolescent Web-based food frequency questionnaire, J. Am. Diet. Assoc. 107 (2007) 605–610.

[28] A.P. Galante, C. Colli, Development and use of an on-line semi-quantitative food-frequency questionnaire to evaluate calcium and iron intake, Rev. Bras. Epidemiol. 11 (3) (2008) 1–9.

[29] R. Martí-Cid, A. Bocio, J. Llobet, J. Domingo, Balancing health benefits and chemical risks associated to dietary habits: RIBEFOOD, a new Internet resource, Toxicology 244 (2008) 242–248.

[30] NCI, National Cancer Institute, Risk Factor Monitoring and Methods, Diet History Questionnaire (DHQ), /http://riskfactor.cancer.gov/DHQ/S, accessed Dec 2011.

[31] J. Beasley, A. Davis, W. Riley, Evaluation of aWeb-based pictorial, diet history questionnaire, Public Health Nutr. 12 (5) (2008) 651–659.

[32] Y.C. Probst, K.D. Agnoli, M. Batterham, L. Tapsell, Video-recorded participant behaviours: the association between food choices and observed behaviours from a Web-based diet history interview, J. Hum. Nutr. Diet. 22 (2009) 21–28.

[33] C. Apovian, M. Murphy, D. Cullum-Dugan, P.-H. Lin, K. Meyers-Gilbert, G. Coffman, M. Jenkins, P. Bakun, K. Tucker, T. Moore, Validation of a Webbased dietary questionnaire designed for the DASH (dietary approaches to stop hypertension) diet: the DASH Online Questionnaire, Public Health Nutr. 13 (5) (2010) 615–622.

[34] Labonté M., Cyr A., Baril-Gravel L., Royer M., Lamarche B., Validity and reproducibility of a Web-based, self-administered food frequency questionnaire, Eur. J. Clin. Nutr. 66 (2) (2012) 166–173.

[35] C. Vereecken, I.D. Bourdeaudhuij, L. Maes, The HELENA online food frequency questionnaire: reproducibility and comparison with four 24-h recalls in Belgian-Flemish adolescents, Eur. J. Clin. Nutr. 64 (5) (2010) 541–548.

[36] DIFE, German Institute of Human Nutrition Potsdam-Rehbr ¨ ucke, /http:// www.dife.de/S, accessed Dec 2011.

[37] A.-K. Illner, U. Harttig, M. Bergmann, E. Bower, P. Amiano, G. Tognon, D. Palli, S. Salvini, T. Kaasik, D. Engeset, H. Ward, H. Boeing, Feasibility of innovative dietary assessment

in epidemiological studies using an approach of combining instruments, Public Health Nutr. 14 (6) (2011) 1055–1063.

[38] Viocare, VioFFQ, /http://www.viocare.comS, accessed Dec 2011.

[39] G. Block, F. Thompson, A. Hartman, F. Larkin, K. Guire, Comparison of two dietary questionnaires validated against multiple dietary records collected during a 1-year period, J. Am. Diet. Assoc. 92 (6) (1992) 686–693.

[40] H. Moore, L. Ells, S. McLure, S. Crooks, D. Cumbor, C. Summerbell, A. Batterham, The development and evaluation of a novel computer program to assess previous-day dietary and physical activity behaviours in school children: the Synchronised Nutrition and Activity Program (SNAP), Br. J. Nutr. 99 (2008) 1266–1274.

[41] L. Arab, K. Wesseling-Perry, P. Jardack, J. Henry, A. Winter, Eight selfadministered 24-hour dietary recalls using the Internet are feasible in African, Americans and Caucasians: the energetics study, J. Am. Diet. Assoc. 110 (6) (2010) 857–864 http://www.24hrrecall.com, Accessed Dec 2011.

[42] K. Storey, L. McCargar, Reliability and validity of Web-SPAN, a Web-based method for assessing weight status, diet and physical activity in youth, J. Hum. Nutr. Diet. (2011) 1–10.

[43] B. Liu, H. Young, F. Crowe, V. Benson, E. Spencer, T. Key, P. Appleby, V. Beral, Development and evaluation of the Oxford WebQ, a low-cost, Web-based method for assessment of previous 24 h dietary intakes in large-scale prospective studies, Public Health Nutr. 14 (11) (2011) 1998–2005.

[44] O. Coltell, M. Arregui, A. Fabregat, E. Barrera, E.A. ibarro, D. Corella, OBENUTIC- 24 H: computer tools for data acquisition of diet and its transformation in nutrients from 24 h recalls in nutritional studies, Obesity Metab. 5 (1) (2009) 36 https://www.obenutic.uji.es/.

[45] A. Ramadas, K. Quek, C. Chan, B. Oldenburg, Z. Hussein, Randomised-controlled trial of a Web-based dietary intervention for patients with type 2 diabetes mellitus: study protocol of my DIDea, Obesity Metab. 11 (359) (2011) 1–8.

[46] P. García-Segovia, R. Gonzá lez-Carrascosa, J. Martínez-Monzó , J. Ngo, L. Serra- Majem, New technologies applied to food frequency questionnaires a current perspective, Nutr. Hosp. 26 (4) (2011) 803–806.

[47] P. Stumbo, R.Weiss, J.W. Newmann, J.A. Pennington, K.L. Tucker, P.L.Wiesenfeld, A.-K. Illner, D.M. Klurfeld, J. Kaput, Web-enabled and improved software tools and data are needed to measure nutrient intakes and physical activity for personalized health research, J. Nutr. (2010) 1–12.

[48] J. Ngo, A. Engelen, M. Molag, J. Roesle, P. García-Segovia, L. Serra-Majem, A review of the use of information and communication technologies for dietary assessment, Br. J. Nutr. 101 (Suppl. 2) (2009) S102–S112.

[49] B. Holland, J. Brown, D. Buss, Fish and Fish Products. The Third Supplement to McCance and Widdowson's the Composition of Foods, The Royal Society of Chemistry, Cambridge, UK, 1993.

[50] W. Chan, J. Brown, S.C.D. Buss, Meat Products and Dishes. Supplement to McCance and Widdowson's the Composition of Foods, The Royal Society of Chemistry, Cambridge, UK, 1996.

[51] O. Moreiras, A. Carvajal, L. Cabrera, Tablas de composición de alimentos, Ed. Pirámide, Madrid, 1996 1996.

[52] B. Burlingame, G. Milligan, R. Quigley, T. Springgs, FOODfiles Manual, New Zealand Institute for Crop and Food Research, Wellington, 1995.

[53] U. Donovan, R. Gibson, Iron and zinc status of young women aged 14 and 19 years consuming vegetarian and omnivorous diets, J. Am. Coll. Nutr. 14 (1995) 463–472.

[54] NUBEL, Table, Belgische voedingsmiddelentabel, 3rd ed., Ministry of Public, Brussels, Belgium, 1999, /http://www.nubel.com/S, accessed Dec 2011.

[55] NEVO, NEVO Tabel. Nederlands voedingsstoffenbestand, Velotekst, Den Haag, The Netherlands, 1996, /http://www.rivm.nl/nevo_en/S, accessed Dec 2011.

[56] L. Moreno, M. Gonzá lez-Gross, M. Kersting, D. Molná r, S.D. Henauw, L. Beghin, M. Sj ¨ostr¨om, M. Hagstr¨omer, Y. Manio, C. Gilbert, F. Ortega, J. Dallongeville, D. Arcella, J. W¨arnberg, M. Hallberg, H. Fredriksson, L. Maes, K. Widhalm, A. Kafatos, A. Marcos, Assessing, understanding and modifying nutritional status, eating habits and physical activity in European adolescents: the HELENA (HEalthy Lifestyle in Europe by Nutrition in Adolescence) study, Public Health Nutr. 11 (2008) 288–299 http://www.helenastudy.com/.

[57] C. Vereecken, M. Convents, W. Sichert-Hellert, J. Alvira, C.L. Conne, S.D. Henauw, T.D. Vriendt, M. Phillip, L. Beghin, Y. Manios, L. Hallstrom, E. Poortvliet, C. Matthys, M. Plada, E. Nagy, L. Moreno, Development and evaluation of a self-administered computerized 24-h dietary recall method for adolescents in Europe, Int. J. Obesity 32 (2008) 26–34.

[58] NCC, Nutrition Coordinating Center (NCC) at the University of Minnesota, Food and Nutrient Database, /http://www.ncc.umn.edu/S,2011, accessed Jan 2012.

[59] D. Toobert, L. Strycker, S.E. Hampson, E. Westling, S.M. Christiansen, T.G. Hurley, J.R. Hébert, Computerized portion-size estimation compared to multiple 24-hour dietary recalls for measurement of fat, fruit, and vegetable intake in overweight adults, J. Am. Diet. Assoc. 111 (2011) 1578–1583.

[60] G. Greene, K. Resnicow, F. Thompson, K. Peterson, T. Hurley, J. Hebert, D. Toobert, G. Williams, D. Elliot, T. Sher, A. Domas, L. Nebeling, D. Midthune, M. Stacewicz-Sapuntzakis, A. Yaroch, Correspondence of the NCI Fruit and Vegetable Screener to repeat 24-h recalls and serum carotenoids in behavioral intervention trials, J. Nutr. 138 (2008) 200S–204S.

[61] G. Block, C. Clifford, M. Naughton, M. Henderson, M. McAdams, A brief dietary screen for high fat intake, J. Nutr. Educ. 21 (1989) 199–207.

[62] Fineli, Finnish Food Composition Database, /http://www.fineli.fi/index. php?lang=enS, accessed Dec 2011.

[63] USDA, National Nutrient Database for Standard Reference, /http://www.ars. usda.gov/S, accessed Dec 2011.

[64] A. Jimenez-Escrig, A. Santos-Hidalgo, F. Saura-Calixto, Common sources and estimated intake of plant sterols in the Spanish diet, J. Agric. Food Chem. 54 (9) (2006) 3462–3471.

[65] H. Boeing, A.-K. Illner, U. Harttig, M. Bergmann, Web-Based European Food Propensity Questionnaire, Technical Report, Department of Epidemiology, German Institute of Human Nutrition Postdam-Rehbr ¨ ucke.

[66] J. McDaniel, K. Ahijevych, M. Belury, Effect of n-3 oral supplements on the n-6/n-3 ratio in young adults, West J. Nurs. Res. 32 (1) (2010) 64–80.

[67] NutritionQuest, Block questionnaires, /http://www.nutritionquest.com/S, accessed Dec 2011.

[68] S.E. Anderson-Bill, A.R. Winett, R.J. Wojcik, Social cognitive determinants of nutrition and physical activity among Web-health users enrolling in an online intervention: the influence of social support, self-efficacy, outcome expectations, and self-regulation, J. Med. Internet Res. 13 (1) (2011) e28.

[69] V.I. voor Gezondheidspromotie, The Food Triangle: A Practical Guide, Brussels, Vlaams Instituut voor Gezondheidspromotie, 2003.

[70] C. Matthys, S.D. Henauw, C. Devos, G.D. Backer, Estimated energy intake, macronutrient intake and meal pattern of Flemish adolescents, Eur. J. Clin. Nutr. 57 (2003) 366–375.

[71] L. Maes, T. Cook, C. Ottovaere, C. Matthijs, L. Moreno, M. Kersting, A. Papadaki, Y. Manios, S. Dietrich, L. Hallstr ¨om, L. Haerens, I.D. Bourdeaudhuij, C. Vereecken, Pilot evaluation of the HELENA (HEalthy lifestyle in Europe by nutrition in adolescence) food-o-meter, a computertailored nutrition advice for adolescents: a study in six European cities, Public Health Nutr. 31 (2011) 1–11.

[72] T. Zimmermann, S. Hull, S. McNutt, B. Mittl, N. Islam, P. Guenther, F. Thompson, N. Potischman, A. Subar, Challenges in converting an interviewer- administered food probe database to self-administration in the national cancer institute automated self-administered 24-hour recall (ASA24), J. Food Compos. Anal. 22S (2009) S48–S51 http://riskfactor.cancer. gov/tools/instruments/asa24/.

[73] J. Gregory, S. Lowe, C. Bates, A. Prentice, L. Jackson, G. Smithers, R. Wenlock, M. Farron, National Diet and Nutrition Survey: Young People Aged 4 to 18 Years, vol. 1, The Stationery Office, London, 2000.

[74] L. Arab, C.H. Tseng, A. Ang, P. Jardack, Validity of a multipass, Web-based, 24- hour self-administered recall for assessment of total energy intake in black and whites, Am. J. Epidemiol. (2011) 1–10.

[75] ESHA, Food processor, /http://www.esha.com/S, accessed Dec 2011.

[76] Canadian, Nutrient file database, /http://www.hc-sc.gc.ca/fn-an/nutrition/ fiche-nutri-data/index-eng.phpS, 2001, accessed Dec 2011.

[77] V. Vance, S. Woodruff, L. McCargar, J. Husted, R. Hanning, Self-reported dietary energy intake of normal weight, overweight and obese adolescents, Public Health Nutr. 12 (2) (2008) 222–227.

[78] J. Mataix, Tabla de Composición de Alimentos, Universidad de Granada, ISBN 9788433849809, 2009.

[79] CESNID, Tablas de Composición de Alimentos del Centro Ensen˜anza Superior de Nutrición y Dieté tica (CESNID), McGraw Hill, ISBN: 844860590X, 2003, /http://www.cesnid.ub.edu/es/que_es.htmS.

[80] A. Fabregat, M. Arregui, E. Barrera, O. Portolé s, D. Corella, O. Coltell, NutriGeneOntology: a biomedical ontology for nutrigenomics research, in: 2008 International Conference on BioMedical Engineering and Informatics, ISBN 978-0-7695-3118-2, 2008, pp. 915–919.

[81] A. Fabregat, E. Barrera, M. Arregui, O. Portolé s, D. Corella, O. Coltell, BOGENVI: a biomedical ontology for modelling gene*environment interactions on intermediate phenotypes in nutrigenomics research, in: 21st IEEE International Symposium on Computer-Based Medical Systems, 2008, pp. 302–307.

[82] L. Penn, H. Boeing, C. Boushey, L. Dragsted, J. Kaput, A. Scalbert, A. Welch, J. Mathers, Assessment of dietary intake: NuGO symposium report, Genes Nutr. 5 (2010) 205–213.

[83] M.L. Slattery, M.A. Murtaugh, M.C. Schumacher, J. Johnson, S. Edwards, R. Edwards, J. Benson, L. Tom-Orme, A. Lanier, Development, implementation, and evaluation of a computerized self-administered diet history questionnaire for use in studies of American Indian and Alaskan native people, J. Am. Diet. Assoc. 108 (1) (2008) 101–109.

[84] S. Edwards, M. Slattery, M. Murtaugh, R. Edwards, J. Bryner, M. Pearson, A. Rogers, A. Edwards, L. Tom-Orme, Development and use of touch-screen audio computer-assisted self-interviewing in a study of American Indians, Epidemiology 165 (2007) 1336–1342.

[85] J. Zoellner, J. Anderson, S. Martin-Gould, Comparative validation of a bilingual interactive multimedia dietary assessment tool, J. Am. Diet. Assoc. 105 (2005) 1206–1214.

[86] D. Wang, M. Kogashiwa, S. Ohta, S. Kira, Validity and reliability of a dietary assessment method: the application of a digital camera with a mobile phone card attachment, J. Nutr. Sci. Vitaminol. 48 (6) (2002) 498–504.

[87] S. Kikunaga, T. Tin, G. Ishibashi, The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey, J. Nutr. Sci. Vitaminol. 53 (2007) 109–116.

[88] C. Boushey, D. Kerr, J. Wright, K. Lutes, D. Ebert, E. Delp, Use of technology in children's dietary assessment, Eur. J. Clin. Nutr. 63 (Suppl 1) (2009) S50–S57.

[89] A. Mariappan, M. Ruiz, F. Zhu, C. Boushey, D. Kerr, D. Ebert, E. Delp, Personal dietary assessment using mobile devices, Proceedings SPIE-The International Society for Optical Engineering 7246 (72460Z) (2009) 1–12.

[90] F. Zhu, A. Mariappan, C. Boushey, D. Kerr, K. Lutes, D. Ebert, E. Delp, Technologyassisted dietary assessment, in: C.A. Bouman, E.L. Miller, I. Pollak (Eds.), Computational Imaging, SPIE Proceedings, vol. 6814, SPIE, 2008, p. 681411.

[91] L. Arab, D. Estrin, D. Kim, J. Burke, J. Goldman, Feasibility testing of an automated image-capture method to aid dietary recall, Eur. J. Clin. Nutr. (2011) 1–7.

[92] N. Lambert, J. Plumb, B. Looise, Using smart card technology to monitor the eating habits of children in a school cafeteria: 1. developing and validating the methodology, J. Hum. Nutr. Diet. 18 (2005) 243–254.

[93] K.A. B¨ alter, O. B¨ alter, E. Fondell, Y.T. Lagerros, Web-based and mailed questionnaires: a comparison of response rates and compliance, Epidemiology 16 (4) (2005) 577–579.

[94] C. Lygidakis, S. Rigon, S. Cambiaso, E. Bottoli, F. Cuozzo, S. Bonetti, C. Della- Bella, C. Marzo, A Web-based versus paper questionnaire on alcohol and tobacco in adolescents, Telemed. J. E Health (9) (2010) 925–930.

[95] M. Touvier, C. Mé jean, E. Kesse-Guyot, C. Pollet, A. Malon, K. Castetbon, S. Hercberg, Comparison between Web-based and paper versions of a selfadministered anthropometric questionnaire, Eur. J. Epidemiol. 25 (5) (2010) 287–296.

[96] A. Vergnaud, M. Touvier, C. Mé jean, E. Kesse-Guyot, C. Pollet, A. Malon, K. Castetbon, S. Hercberg, Agreement between Web-based and paper versions of a socio-demographic questionnaire in the NutriNet-Santé study, Int. J. Public Health 56 (4) (2011) 407–417.

[97] M. Touvier, E. Kesse-Guyot, C. Mé jean, C. Pollet, A. Malon, K. Castetbon, S. Hercberg, Comparison between an interactive Web-based self-administered 24 h dietary record and an interview by a dietitian for large-scale epidemiological studies, Br. J. Nutr. 17 (2010) 1–10.

[98] A.-K. Illner, U. N¨othlings, K.Wagner, H. Ward, The assessment of individual usual food intake in large-scale prospective studies, Ann. Nutr. Metab. 56 (2010) **99–105.**

7

DISCUSSION

Chapter 7: discussion

In this thesis, genetic determinants of intermediate and final phenotypes of cardiovascular risk were examined in Spanish and German studies. Also a systematic review on available tools for food questionnaires automation in medical studies was conducted. Although results from each study have already been discussed in the chapters where they were presented, in the present chapter the main findings are summarized and some methodological issues are discussed.

7.1 MAIN FINDINGS AND COMPARISON OF RESULTS OF RECENTLY PUBLISHED STUDIES

Chapter 2, entitled "*Association of the LCT-13910C>T polymorphism with obesity and its modulation by dairy products in a Mediterranean population*", investigated the impact of the SNP rs4988235 C/T from the *MCM6* gene, located upstream the *LCT* gene, on measures of anthropometry in the PREDIMED-Valencia cohort. Being the first study of its kind in a Mediterranean population, it was found that homozygotes for the C allele (genotype consistent with a lactase non-persistent phenotype), had a lower mean BMI and waist circumference and a lower odds of obesity compared to the rest of participants. This was independent of sex, age, diabetes, physical activity and energy intake. This finding agrees with the results from the only study that was subsequently published (1). That study had a cross-sectional design and included 551 randomly selected adults belonging to the Canary Islands Nutrition Survey in Spain (aged 18-75 years) (1). Further, in PREDIMED-Valencia, it was observed that differences in waist circumference and obesity between genotypes were greater when lactose intake was high.

In **Chapter 3** (*Significant associations of the rs2943634 (2q36.3) genetic polymorphism with adiponectin, high density lipoprotein cholesterol and ischemic stroke*) the role of the SNP rs2943634 C/A on the risk of MI and IS was investigated. This SNP is located in a non-coding region of chromosome 2q36.3 and had previously been associated with coronary artery disease in two GWA studies. While no significant associations were observed for MI, carriers of the minor

allele of rs2943634 had a lower risk of IS. Further, the association of rs2943634 with 12 intermediate risk phenotypes was cross-sectionally explored. In line with the lower stroke risk, carriers of the minor allele showed slightly higher HDL-cholesterol and adiponectin levels. These intermediate phenotypes, however, could not explain the association with IS, which leaves open the need of further studies to clarify the underlying mechanisms. Following this study, other authors also investigated the role of this SNP in atherosclerotic diseases. In line with our findings suggesting a protective effect of the rare allele of rs2943634, a cross-sectional study in 1886 participants from the two EPIC centers of the Netherlands reported an inverse association with the MetS (odds ratio 0.88, 95% CI 0.79; 0.97) (2). Further, its major allele, compared to the minor, was associated with incident coronary heart disease (CHD) in two out of five US cohorts of different race/ethnicity in a study where follow-up ranged between 9.1 and 15.7 years. Significant associations were observed in a cohort of 12316 white participants (3963 cases, HR 1.07 95%, CI 1.04-1.11), and in a cohort of 883 Asian and Pacific Islanders (65 CHD cases, HR 2.00 95%, CI 1.20-3.33). In contrast, no associations with incident CHD were found in 3 cohort studies, including a cohort of 6502 African Americans (677 cases, HR 1.02, 95% CI 0.93-1.12), a cohort of 203 Indian-Americans (20 cases, HR 1.29, 95% CI 0.76-2.16), and a cohort of 1897 Hispanics (113 cases, HR 0.95, 95% CI 0.76-1.18) (3). Finally, in a collaborative consortium of 12 studies, involving 25,000 participants and 5794 incident CHD events, the major allele of rs2943634 was associated with a slightly higher CHD risk (odds ratio 1.08, 95% CI 1.03-1.14). However, no associations were found for any of the 13 investigated biological risk factors for CVD (4).

Chapter 4, entitled "*Heterogeneity of the Stearoyl-CoA desaturase-1 (SCD1) gene and metabolic risk factors in the EPIC-Potsdam Study*", is the largest study conducted so far to evaluate the impact of *SCD1* heterogeneity on MetS-related traits and on markers of liver fat and inflammation. Despite the strong interest in SCD1 due to its important role in the de novo lipogenic pathway (5), most of the existing knowledge on its function comes from mice models (5). After publication of the EPIC-Potsdam studies, no other studies have been published. Prior to our study, the impact of *SCD1* heterogeneity on metabolic risk factors had only been investigated in four human previous studies (6-9) which had provided mixed results. The findings of the study described in Chapter 4 do not suggest a modulation of the investigated traits by means of the SCD1 common genetic variants investigated. However, given its biological relevance, genetic

heterogeneity of human *SCD1* in relation to impaired metabolism rewards further investigation in independent study populations, in particular with regards to rare genetic variants.

Chapter 5 entitled "*Microsomal triglyceride transfer protein -164 T > C gene polymorphism and risk of cardiovascular disease: results from the EPIC-Potsdam case-cohort study*" tested the hypothesis that a SNP in the promoter region of *MTTP*, could modulate the risk to develop a cardiovascular event, depending on cholesterol levels. This SNP had inconsistently been associated with CVD in previous studies. In EPIC-Potsdam, participants with lower cholesterol levels (<200 mg/dL) showed indeed an increased risk, while participants with higher cholesterol levels showed a decreased risk. This last result could also be observed in a smaller replication cohort, suggesting that risk allele carriers with low cholesterol levels may be predisposed to an increased risk of developing CVD, which seemed to be abolished among risk allele carriers with high cholesterol levels. Other similar studies have not been published to date.

Chapter 6, entitled "*Automation of Food Questionnaires in Medical Studies: A state-of-the-art review and future prospects*", presents a review of different technological approaches used in nutritional epidemiological studies, for the automatization of FFQs and 24HDRs. For this purpose, a search in different electronic databases for articles on studies published until December 2011 was undertaken. The identified tools were first classified into computerized questionnaires vs. Web-based or Internet-based questionnaires. A further division was then made based on method of dietary assessment (that is, FFQs, 24HDRs, or combinations of both). A final classification was made by taking into account the type of administration to participants, that is, interviewer vs. self-administered methods. Additionally, some recent developments in the use of innovative technologies, such as audio, multimedia touch-screen, pocket-PC, digital camera and mobile phone, were described. This review shows that FFQs and 24HDRs dietary recalls are used/can be used in several automated forms in different studies. Depending on the aim of the study and on the study population, a method for automating nutrition questionnaires would be more suitable than others.

7.2 METHODOLOGICAL CONSIDERATIONS

Study populations

Chapter 2 presents a cross-sectional genetic association study, conducted on a high cardiovascular risk population, the PREDIMED-Valencia cohort. This cohort has its basis in the Faculty of Medicine and Odontology of the University of Valencia. Details on recruitment procedures have been explained in detail elsewhere (10). Briefly, potential participants were selected from records of a convenience sample of general practitioners of the Valencia region, Spain. Practitioners directly invited their patients to participate, and 70% of them agreed. Inclusion criteria consisted on being aged 55–80 years old for men and 60–80 years old for women; not having prevalent CVD, and having either type 2 diabetes or at least three of the following cardiovascular risk factors: current smoking, hypertension, dyslipidemia, overweight and family history of premature CVD. Exclusion criteria included: severe chronic illness, immunodeficiency, human immunodeficiency virus positive status, illegal drug or alcohol misuse, allergy to olive oil or nuts and low predicted likelihood of changing dietary habits.

The characteristics of this highly selected population may have compromised the extent to which results inferred from this genetic association study can be generalizable to other populations (that is, its external validity). Nevertheless, the findings were consistent with those reported in other populations (1, 11).

Studies on Chapters 3 to 5 are based on a subsample of the EPIC-Potsdam cohort. EPIC-Potsdam has its basis in the German Institute of Human Nutrition and includes 27,548 persons from the general population of the city of Potsdam and surroundings (Brandenburg, Germany). Participants were recruited between 1994 and 1998 and the procedures have been described in detail elsewhere (12). Briefly, it was based on addresses from general population registries. Participants were initially approached by mail. Those who did not respond within 2 weeks were re-contacted again either by mail or telephone. The participation rate, compared to the invited number of subjects, was 22.7% with a considerable variation by municipality and gender. Men (39.6%) were mainly aged 40-65 and women 35-65 years old at recruitment. After recruitment, participants have been re-contacted and asked to fill in follow-up questionnaires every 2-3 years. This is done to identify incident cases of chronic diseases and changes in diet and lifestyle. The

response rate of the follow-up has so far ranged between 92-96%, depending on the follow-up wave.

Being a random sample of EPIC-Potsdam, studies on chapters 3 to 5 shared the limitations of cohort studies, which include been sensitive to selection bias due to participation. Selection bias occurs when there is a difference in characteristics between people who participate in the study and those who do not. First, this may occur because the people who participate are different than those in the source population. A comparison of general characteristic of the EPIC-Potsdam cohort with data from the German National Health Survey 1991/1992 showed that the study population had higher socio-economic status and was healthier than the source population (12). This is in line with the observed fact that health-conscious people are often more likely to participate in health-related studies than others. While the selective participation in this study may have affected the external validity of the findings, conversely, it most likely played a role safeguarding the internal validity of the prospective study, by ensuring high participation rates across the follow-up rounds (12). Second, selection bias may occur if many participants drop out of the study during follow-up due to factors related to the exposure or outcome. However, as noted above, response rates during follow-up were high, which makes this form of selection bias unlikely. Furthermore, it has been suggested that when disease associations are estimated as relative risks, selective participation is unlikely to lead to wrong etiological conclusions (12).

Study design

Classical study designs for genetic association studies, include, but are not limited to, case–control, nested case-control, cohort, and cross-sectional designs. More recently also the case-cohort design has been used.

Chapters 2 to 5 present cross-sectional results for the association between SNPs and intermediate phenotypes of CVD. Unlike prospective designs, cross-sectional studies have the limitation that exposure is measured at the same time as the phenotype of interest, which opens the possibility of reverse causality. For genetic association studies using SNPs as exposure, however, this is less of a problem given that genotype does not change over the time.

Studies on Chapter 4 and 5 follow a case-cohort design (13), which combines the advantages of a prospective cohort study with the efficiency of a case-control design. This type of design is useful in follow-up studies when large cohorts are needed to acquire enough cases and it is not feasible to collect data on all covariates for the entire cohort. In case-cohort studies, the study set is constituted by a random subsample (subcohort) of the study population, and all members of the cohort developing the disease of interest during follow-up, in this case myocardial infarction and ischemic stroke. With a sufficiently large subcohort, it is expected that the characteristics of the subcohort are comparable to those of the total cohort. For the EPIC-Potsdam Study, the subcohort was indeed similar to the entire cohort with regard to socio-demographic, lifestyle, and anthropometric characteristics at baseline. Therefore, the results from the case-cohort study are expected to be generalizable to the entire cohort without the need to measure biomarkers and genotype SNPs in the all participants. Further, compared to nested case-control studies, the case-cohort design is more flexible, as it allows the use of the same subcohort to investigate the effect of the exposure on different outcomes. However, members of the subcohort can potentially develop the disease outcome of interest (given that they are representative of the total cohort) and would, therefore be excluded to safeguard the prospective design. Therefore, slightly more controls need to be selected in case-cohort studies than in nested case-control studies to achieve the same statistical power for the same number of cases (14).

The theoretical background for the case-cohort design was formulated in 1986 by Prentice (15), and it is suitable to measure risk ratios (14). The statistical analyses involve models for survival data (e.g. Cox regression) with some form of weighting to take into account the sampling from the original cohort (16). Three weighting methods have been proposed by Prentice, Self-Prentice, and Barlow. They differ in the way to handle subcohort members and cases outside the subcohort (external cases) at the time the event of interest occurs (16). Methods proposed by Prentice and Self-Prentice use weights 1 and 0. Subcohort members have weight 1 at all times. For the Prentice method, external cases have weight 0 at all times, until precisely the time of failure, in which the weight becomes 1. For the Self-Prentice method, external cases have weight 0 at all times (16). In the approach from Barlow, cases' weight is always 1 and the subcohort is weighted by the inverse of the sampling fraction (16). The three methods have shown to result in very similar effect estimates and standard errors. However, when the size of the subcohort is small, the

estimates given by the Prentice method have been suggested to better resemble the estimates from a full-cohort analysis (17).

Metabolic syndrome

Study on Chapter 2 contemplated the concept of MetS, while study on Chapter 3 looked at single components of this syndrome (among other biomarkers) but not at overall MetS. The concept of MetS focuses attention on the complex multifactorial health problems of abdominal obesity, high triglycerides and low HDL-cholesterol concentrations, elevated blood pressure, and glucose intolerance (18). Individuals with MetS are at increased risk for CVD and type 2 diabetes; the more individual MetS traits a person clusters, the greater the chances of developing an adverse outcome (19).

Although MetS is influenced by lifestyle and genetics (20), its exact pathogenesis remains unclear (21). The last report of a World Health Organization Expert Consultation described MetS as a useful educational concept, but with limited practical utility as a diagnostic or management tool (18). Moreover, Reaven, the first to propose the MetS (22), recently questioned the etiological role of abdominal obesity in the development of the other components of the MetS, and also concluded that the Framingham Risk Score performs better predicting CVD, and fasting plasma glucose concentrations might be better for predicting type 2 diabetes (19). While these arguments would not favour the use of MetS as a concept, there is general agreement that glucose intolerance, high triglycerides, low HDL-cholesterol, and elevated blood pressure tend to cluster together, and that obesity increases chances for this cluster to occur (19). In this context the MetS concept provides a framework to explore possible basis for the observed group of risk factors. Since the term was first classified, various definitions for MetS have been proposed by different organizations (23-26), the most recent come from the International Diabetes Federation (IDF) (25) and the American Heart Association /National Heart, Lung, and Blood Institute (AHA/NHLBI) (26). After harmonization, the difference among them concerns the measure for central obesity, which for the IDF is lower and ethnic specific (27). Due to the absence of a clear mechanism for the pathogenesis of MetS, these definitions can only be considered provisional (18).

Genetic factors as exposure

Unlike other exposures, single-nucleotide polymorphisms (SNPs) are stable throughout life and not influenced by age, sex, calendar year, or disease status. Although this makes SNPs insensitive to confounding and reverse causation, there are other factors that may obscure its findings. In studies with different ethnic groups confounding (or bias) can be introduced by population stratification if the phenotype of interest varies with ancestry (28). Throughout this thesis, confounding by population stratification was avoided because both cohorts examined were homogeneous in terms of race and ethnicity.

Another potential source of confounding in SNP studies are characteristics that are genetically driven (at least to some extent) and associated with the investigated phenotypes. The influence of this potential form of confounding was reduced by including such characteristics in different adjusting models. However residual confounding due to potential measurement error in the variables adjusted for can never be excluded in observational.

Also, by testing multiple hypotheses, false-positive associations may arised due to chance statistical fluctuations. The possibility of chance findings was lowered through the use of a correction for multiple testing by means of the Bonferroni method (29). However, it may have resulted in too conservative estimates (false-negatives), given that the assumption of independency among tests was not strictly fulfilled (30)

Measurement error in genotyping also can be a source of biased results (31). In order to minimize this threat, biological samples from all the study participants investigated in each cohort were drawn and stored under identical conditions. Further, genotyping was conducted by standardized procedures, with validated assays and according to suppliers' protocols. In the special case of rs4988235 (Chapter 2), a commercially available assay was not available. Thus, it was necessary to design a customized assay and to use additional quality control measures to ensure that the new assay was valid. Accordingly, the SNP was additionally genotyped with a standard RFLP method in a random subset consisting of 50% of the study population. Concordance between techniques was higher than 95%, and discrepant samples were sequenced. As a result of quality control measures, all the SNPs investigated in this thesis showed a high call-rate and followed

Hardy-Weinberg equilibrium (HWE), which otherwise often points at a possible genotyping problems (31).

While results from genetic-association studies are a valuable tool to generate hypothesis concerning a given gene in the etiology of a particular phenotype, functional studies are subsequently needed to infer causality. Also, even though the genome has a static secuence, there are other genetic factors that may play an important role in the modulation of intermediate and final phenotypes of CVD. For instance, the complex compaction of the DNA molecule inside the cell nucleus has an influence on which genes become accessible for transcription (32). Also epigenetic changes, can also regulate gene expression. Examples are DNA methylation and histone modification (33)

Need of replication studies in independent cohorts

Given that variants contributing to complex traits are likely to have modest effects, and that testing multiple hypotheses requires corrections (such as Bonferroni), which limit the power of the studies, small sample sizes may result in insufficient statistical power to detect minor contributions of the alleles to the modulation of the investigated phenotypes. Thus to evaluate the results of this work, further studies in independent cohorts analyzing the same SNPs, or those in perfect linkage disequilibrium, are indispensable (34). This may be done in collaborative research consortia, which combine several large studies. One example is the METASTROKE collaboration, which included 12389 IS cases and 62004 controls of European ancestry from 15 case-control studies. The novel associations identified, were further replicated in the same study in an independent set of 13347 cases and 29083 controls from 18 case-control studies (35).

Use of food-frequency questionnaires

In the study of Chapter 2, food consumption was determined by a validated semi-quantitative food-frequency questionnaire (FFQ). FFQs are widely used in epidemiological studies because they are easy to administer, relatively cost effective, and typically cover habitual diet of the past year. However, they also have limitations, including the inability to assess absolute intakes and

their susceptibility to measurement error. It is well-known that FFQs generally tend to over-estimate food consumption. Further, misreporting of dietary intake may also differ according to age, sex, BMI, and educational attainment. Further inaccuracies may result from incomplete food lists contained in the FFQ. Due to these limitations, FFQ are better used to rank participants according to food consumption rather than for estimating absolute levels of intake, like it was done for lactose intake for the study in Chapter 2. The performance of FFQs in terms of validity and reproducibility always needs to be assessed against other assessment methods, such as the 24-h dietary recall. Advantages of the 24-h recall over the FFQ, include that it collects data based on short-term memory and that it provides absolute intake information rather than consumption ranges. Further, it doesn't consist of a closed list of foods and it does not require adaptation to specific populations. However, a single 24-h recall is not representative of an individual's usual diet, therefore multiple 24-h recalls are needed to assess diet. Details on how the FFQ used in the study of Chapter 2 was validated have been described elsewhere (36, 37), briefly: the FFQ was administered twice to explore reproducibility at 1 year. Four 3-days dietary records were used as reference to explore validity. This FFQ was designed to be interviewer-administered on an optically readable paper. It is likely that the elderly population of this study had limited access to the Internet and reduced computer skills. Thus, a self-administered computerized or Web-based questionnaire, could have translated into a drop-out of the study and therefore in participation bias. The scanning of the questionnaire for its translation into nutrients, most likely reduced transcription errors, data entry expenses and speeded processing. However, regardless of the type of automatization method, computer-aid questionnaires share the same methodological limitations as their source tools. These limitations may be to some extent overcome by integrative methodological approaches that could combine dietary information from multiple measures (38), and also by the use of dietary biomarkers, which assess dietary consumption without the bias of self-reported dietary intake errors (39).

REFFERENCES

1. Almon R, Alvarez-Leon EE, Serra-Majem L. Association of the European lactase persistence variant (LCT-13910 C>T polymorphism) with obesity in the Canary Islands. PloS one. 2012;7(8):e43978. Epub 2012/09/01.

2. Povel CM, Boer JM, Onland-Moret NC, Dolle ME, Feskens EJ, van der Schouw YT. Single nucleotide polymorphisms (SNPs) involved in insulin resistance, weight regulation, lipid metabolism and inflammation in relation to metabolic syndrome: an epidemiological study. Cardiovascular diabetology. 2012;11:133. Epub 2012/10/30.

3. Franceschini N, Carty C, Buzkova P, Reiner AP, Garrett T, Lin Y, et al. Association of genetic variants and incident coronary heart disease in multiethnic cohorts: the PAGE study. Circulation Cardiovascular genetics. 2011;4(6):661-72. Epub 2011/11/02.

4. Angelakopoulou A, Shah T, Sofat R, Shah S, Berry DJ, Cooper J, et al. Comparative analysis of genome-wide association studies signals for lipids, diabetes, and coronary heart disease: Cardiovascular Biomarker Genetics Collaboration. European heart journal. 2012;33(3):393-407. Epub 2011/08/02.

5. Sampath H, Ntambi JM. The role of stearoyl-CoA desaturase in obesity, insulin resistance, and inflammation. Annals of the New York Academy of Sciences. 2011;1243:47-53. Epub 2012/01/04.

6. Warensjo E, Ingelsson E, Lundmark P, Lannfelt L, Syvanen AC, Vessby B, et al. Polymorphisms in the SCD1 gene: associations with body fat distribution and insulin sensitivity. Obesity. 2007;15(7):1732-40. Epub 2007/07/20.

7. Liew CF, Groves CJ, Wiltshire S, Zeggini E, Frayling TM, Owen KR, et al. Analysis of the contribution to type 2 diabetes susceptibility of sequence variation in the gene encoding stearoyl-CoA desaturase, a key regulator of lipid and carbohydrate metabolism. Diabetologia. 2004;47(12):2168-75. Epub 2005/01/22.

8. Gong J, Campos H, McGarvey S, Wu Z, Goldberg R, Baylin A. Genetic variation in stearoyl-CoA desaturase 1 is associated with metabolic syndrome prevalence in Costa Rican adults. The Journal of nutrition. 2011;141(12):2211-8. Epub 2011/11/04.

9. Stryjecki C, Roke K, Clarke S, Nielsen D, Badawi A, El-Sohemy A, et al. Enzymatic activity and genetic variation in SCD1 modulate the relationship between fatty acids and inflammation. Molecular genetics and metabolism. 2012;105(3):421-7. Epub 2012/01/03.

10. Martinez-Gonzalez MA, Corella D, Salas-Salvado J, Ros E, Covas MI, Fiol M, et al. Cohort profile: design and methods of the PREDIMED study. International journal of epidemiology. 2012;41(2):377-85. Epub 2010/12/22.

11. Kettunen J, Silander K, Saarela O, Amin N, Muller M, Timpson N, et al. European lactase persistence genotype shows evidence of association with increase in body mass index. Human molecular genetics. 2010;19(6):1129-36. Epub 2009/12/18.

12. Boeing H, Korfmann A, Bergmann MM. Recruitment procedures of EPIC-Germany. European Investigation into Cancer and Nutrition. Annals of nutrition & metabolism. 1999;43(4):205-15. Epub 1999/12/11.

13. Rundle AG, Vineis P, Ahsan H. Design options for molecular epidemiology research within cohort studies. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2005;14(8):1899-907. Epub 2005/08/17.

14. Rothman KJ, Greenland, S., & Lash, T.L. (2008). Modern Epidemiology, 3rd Edition. Philadelphia, PA: Lippincott, Williams & Wilkins.

15. Prentice RL, Self SG. Aspects of the use of relative risk models in the design and analysis of cohort studies and prevention trials. Statistics in medicine. 1988;7(1-2):275-87. Epub 1988/01/01.

16. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. Journal of clinical epidemiology. 1999;52(12):1165-72. Epub 1999/12/02.

17. Onland-Moret NC, van der AD, van der Schouw YT, Buschers W, Elias SG, van Gils CH, et al. Analysis of case-cohort data: a comparison of different methods. Journal of clinical epidemiology. 2007;60(4):350-5. Epub 2007/03/10.

18. Simmons RK, Alberti KG, Gale EA, Colagiuri S, Tuomilehto J, Qiao Q, et al. The metabolic syndrome: useful concept or clinical tool? Report of a WHO Expert Consultation. Diabetologia. 2010;53(4):600-5.

19. Reaven GM. The metabolic syndrome: time to get off the merry-go-round? J Intern Med. 2011;269(2):127-36.

20. Isomaa B. A major health hazard: the metabolic syndrome. Life Sci. 2003;73(19):2395-411.

21. Eckel RH, Alberti KG, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet. 2009;375(9710):181-3.

22. Reaven GM. Syndrome X. Blood Press Suppl. 1992;4:13-6.

23. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med. 1998;15(7):539-53.

24. Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation. 2002;106(25):3143-421.

25. Alberti KG, Zimmet P, Shaw J. The metabolic syndrome--a new worldwide definition. Lancet. 2005;366(9491):1059-62.

26. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, et al. Diagnosis and management of the metabolic syndrome: an American Heart Association/National Heart, Lung, and Blood Institute scientific statement: Executive Summary. Crit Pathw Cardiol. 2005;4(4):198-203.

27. Alberti KG, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation. 2009;120(16):1640-5.

28. Attia J, Ioannidis JP, Thakkinstian A, McEvoy M, Scott RJ, Minelli C, et al. How to use an article about genetic association: B: Are the results of the study valid? JAMA : the journal of the American Medical Association. 2009;301(2):191-7. Epub 2009/01/15.

29. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. Bmj. 1995;310(6973):170. Epub 1995/01/21.

30. Perneger TV. What's wrong with Bonferroni adjustments. Bmj. 1998;316(7139):1236-8. Epub 1998/05/16.

31. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nature reviews Genetics. 2005;6(2):95-108. Epub 2005/02/18.

32. Marian AJ, Belmont J. Strategic approaches to unraveling genetic causes of cardiovascular diseases. Circulation research. 2011;108(10):1252-69. Epub 2011/05/14.

33. Schnabel RB, Baccarelli A, Lin H, Ellinor PT, Benjamin EJ. Next steps in cardiovascular disease genomic research--sequencing, epigenetics, and transcriptomics. Clinical chemistry. 2012;58(1):113-26. Epub 2011/11/22.

34. Studies N-NWGoRiA, Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, et al. Replicating genotype-phenotype associations. Nature. 2007;447(7145):655-60. Epub 2007/06/08.

35. Taylor BV, Oudit GY, Kalman PG, Liu P. Clinical and pathophysiological effects of active and passive smoking on the cardiovascular system. The Canadian journal of cardiology. 1998;14(9):1129-39. Epub 1998/10/21.

36. Martin-Moreno JM, Boyle P, Gorgojo L, Maisonneuve P, Fernandez-Rodriguez JC, Salvini S, et al. Development and validation of a food frequency questionnaire in Spain. International journal of epidemiology. 1993;22(3):512-9. Epub 1993/06/01.

37. Fernandez-Ballart JD, Pinol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. The British journal of nutrition. 2010;103(12):1808-16. Epub 2010/01/28.

38. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. International journal of epidemiology. 2012;41(4):1187-203. Epub 2012/08/31.

39. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutrition journal. 2012;11:109. Epub 2012/12/15.

8

CONCLUSIONS

CONCLUSIONS

A general conclusion of the work described in this thesis is, that within the candidate genes analyzed in the different populations, a wide genetic variability exists. Furthermore, although some of these genetic variants were not directly associated with intermediate or final phenotypes of cardiovascular diseases, other variants were. Thus, genetic factors, as measured by means of polymorphisms analysis, contribute to explain a portion of the variability of the intermediate or final phenotypes investigated. However, this association does not seem to be deterministic, given that we were also able to observe the existence of statistically significant gene-diet interactions, meaning that the association between the genetic polymorphism with the cardiovascular phenotype varied with diet. Awareness of these interactions will be very important for a more effective primary and/or secondary prevention, affirming the multifactorial etiology of cardiovascular diseases. Accordingly, the large amount of data generated when studying the factors involved in multifactorial diseases, requires a specific and combined treatment that largely profits from the usage of informatic tools for the storage, recovery and analysis of data.

Specific conclusions:

1. In the EPIC-Potsdam study population, the functional genetic polymorphism 164T> C of the *MTTP* gene showed an interaction with total cholesterol levels, which predisposed risk allele carriers (allele C) with low cholesterol levels, to an increased risk of developing a cardiovascular event. However, this excess risk could be abolished among risk allele carriers with high total cholesterol levels. Further studies are warranted to confirm these results and, when confirmed, to clarify the underlying mechanisms of this interaction.

2. Common variants of the *SCD1* gene within the EPIC-Potsdam cohort do not appear to be relevant genetic factors given their lack of direct statistical associations with metabolic factors that may be affected by the enzymatic activity of SCD1 (triglycerides, body mass index, waist circumference, glycated hemoglobin, high-

sensitivity C-reactive protein, gamma-glutamyltransferase, alanine aminotransferase and fetuin-A). However, given its biological relevance, together with the very limited number of studies available and the inconsistency of their results, genetic heterogeneity of human *SCD1* in relation to impaired metabolism needs further investigation in independent study populations, in particular with regard to rare variants of *SCD1*. The inconsistent results of the few studies currently available may be due to interactions with dietary factors. Genetic associations may only be detectable when certain dietary factors are present.

3. In the EPIC-Potsdam Study, the minor allele of the rs2943634 SNP (allele A), was associated with lower risk to develop an ischemic stroke, but not by means of the biological risk factors investigated in the study (body mass index, waist circumference, blood pressure, total cholesterol, HDL-cholesterol, LDL-cholesterol, triglycerides, blood glucose, glycated hemoglobin, hs-C-reactive protein, adiponectin, creatinine). Although rs2943634 was associated with plasma levels of HDL-cholesterol and adiponectin these factors also did not explain the lower risk associated with the rs2943634. Further investigations are needed to confirm these results and to clarify the mechanisms underlying the association.

4. The previously reported association between the genetic variant *LCT* −13910 C/T and body mass index was confirmed in a Mediterranean population at high cardiovascular risk (PREDIMED-Valencia). Furthermore we reported for the first time upon a potential modulation of the effects of this SNP by lactose intake. These findings also suggest that some of the inconsistencies observed among results from previous studies investigating the association of dairy products intake and obesity, may be explained by the potential heterogeneous effects of these products on lactase non persistence and lactase persistence individuals.

5. In recent years, the tedious task of assessing a person's habitual dietary intake, has greatly benefited from the rapid development of information and communication technologies. This has enabled the automation of dietary intake assessment by means of computerized questionnaires and also, since the expansion of the Internet, by means of Web-based questionnaires. Some of the available tools allow self-administration of

questionnaires, while others are designed for administration by a healthcare professional. Other innovative technologies lately becoming available include audio, pocket-computers, touch screens, digital cameras and mobile phones. The accessible variety of tools enables the study of the diet as a risk factor for various chronic diseases in large epidemiological studies at lower costs and with greater validity and precision.

CONCLUSIONES

Como conclusión general del trabajo compilado en esta tesis, podemos afirmar que existen una amplia variabilidad genética en los genes candidatos analizados en las distintas poblaciones. Además, algunas de estas variantes genéticas no muestran asociación directa con fenotipos intermedios o finales de enfermedad cardiovascular, pero otras sí que han mostrado su relevancia. Ello nos lleva a afirmar que los factores genéticos, medidos a través del análisis de polimorfimos, sí que contribuyen a explicar de manera significativa una parte importante de la variabilidad de los fenotipos intermedios o finales analizados. Pero esta asociación no parece determinista, ya que también hemos podido observar la existencia de interacciones gen-dieta estadísticamente significativas de manera que la asociación del polimorfismo genético con el fenotipo cardiovascular varía según la dieta. El conocimiento de estas interacciones será muy importante para una prevención primaria y/o secundaria más eficaz reafirmando el carácter multifactorial de las enfermedades cardiovasculares. De acuerdo con ello, la gran cantidad de datos que se genera en el estudio de los factores implicados en las enfermedades multifactoriales, requiere un tratamiento agregado y específico que se beneficia enormemente de la aplicación de herramientas informáticas, tanto en el almacenaje como en su recuperación y análisis.

Conclusiones específicas:

1. En la población del estudio EPIC-Potsdam, la variante genética funcional -164T> C del gen *MTTP*, mostró una interacción con niveles plasmáticos de colesterol total que predisponía a los portadores del polimorfismo de riesgo (alelo C) con niveles bajos de colesterol, a un mayor riesgo de desarrollar un evento cardiovascular. Sin embargo este exceso de riesgo podría verse anulado para aquellos portadores de la variante con niveles altos de colesterol. Son necesarias investigaciones adicionales para confirmar estos resultados y clarificar los mecanismos que subyacen a esta interacción.
2. En la población del estudio EPIC-Potsdam, variantes genéticas comunes del gen *SCD1* no parecen ser factores genéticos relevantes en su asociación estadística directa con factores metabólicos que podrían estar relacionados con la actividad de la enzima SCD1 (triglicéridos, índice de masa corporal, circunferencia de cintura, hemoglobina

glucosilada, proteína-C-reactiva de alta sensibilidad, gamma-glutamyltransferasa, alanina aminotransferasa y fetuina-A). Sin embargo, dada su importancia biológica, el número todavía escaso de estudios disponibles y la inconsistencia de sus resultados, la heterogeneidad genética del gen humano *SCD1* en relación con alteraciones metabólicas requiere investigación adicional en poblaciones de estudio independiente, en particular con respecto a variantes genéticas raras. Estos resultados también nos hacen concluir que pueden existir importantes interacciones con la dieta, y en función de la composición de la misma que se lleguen a magnificar o no las posibles asociaciones genéticas.

3. En la población europea del estudio EPIC-Potsdam, el alelo menos frecuente del SNP rs2943634, (alelo A) se asoció con menor riesgo de desarrollar ictus hisquémico, pero no por medio de los factores de riesgo investigados en el estudio (índice de masa corporal, circunferencia de cintura, presión arterial, colesterol total, colesterol-HDL, colesterol-LDL, triglicéridos, glucemia, hemoglobina glicosilada, proteína-C-reactiva de alta sensibilidad, adiponectina y creatinina). El polimorfismo rs2943634 se asoció también con niveles plasmáticos de colesterol-HDL y de adiponectina. Son necesarias investigaciones adicionales para confirmar estos resultados y aclarar los mecanismos que subyacen a la asociación.

4. Replicación en población mediterránea de alto riesgo cardiovascular (PREDIMED-Valencia), de la asociación previamente sugerida en otras poblaciones, de la variante genética LCT-13910C/T con índice de masa corporal. Además, es la primera vez que se aporta información acerca de una posible modulación de los efectos de este SNP por mediación de ingesta de lactosa. Los resultados de este estudio también sugieren que algunas de las inconsistencias observadas entre resultados de estudios previos sobre la asociación de consumo de productos lácteos y obesidad, podrían deberse a la heterogeneidad del efecto de estos productos en individuos con y sin persistencia del enzima lactasa.

5. En los últimos años, la tediosa tarea de evaluar la dieta de una persona, se ha beneficiado del rápido desarrollo de las tecnologías de la información y la comunicación. Este desarrollo ha permitido la automatización de la medida de la dieta mediante cuestionarios computarizados y también, desde la expansión de Internet, mediante cuestionarios basados en Web. Algunas de las herramientas disponibles permiten la autoadministración

de cuestionarios, mientras que otras están diseñadas para la administración por un profesional sanitario. Otras tecnologías innovadoras que comienzan a estar disponibles incluyen audio, ordenadores de bolsillo, pantallas táctiles, cámaras digitales o teléfonos móviles. La variedad de herramientas permite el estudio de la dieta como factor de riesgo para diferentes enfermedades crónicas en grandes estudios epidemiológicos a costes inferiores y con mayor validez y precisión.

9

Investigación predoctoral adicional

Chapter 9: Investigación predoctoral adicional

Durante los años de formación predoctoral, además de los trabajos presentados en el compendio de publicaciones que conforman esta tesis, la doctoranda tuvo la oportunidad de participar en proyectos de investigación adicionales que complentaron su formación en diversas áreas de la epidemiología genética y nutricional.

Así, en colaboración con el Departamento de Lenguajes y Sistemas Informáticos de la Universitat Jaume I de Castellón, participó en proyectos del ámbito de la informática biomédica aplicada a la investigación en nutrigenética (proyectos FITUVEROLES y OBENUTIC) y en colaboración con el Departamento de Departamento de Medicina Preventiva y Salud Pública, Ciencias de la Alimentación, Toxicología y Medicina Legal, participó en estudios de asociación genética de polimorfismos en genes candidatos (*SCD1*, *INSIG2*, *MCM6*, *NPPA*, *TAGLN2*, *THRα*) con fenotipos intermedios de riesgo cardiovascular en población mediterránea de alto riesgo cardiovascular.

Como resultado de este trabajo, la doctoranda obtuvo el título de *Diploma de Estudios Avanzados* expedido por la Universitat de València. El trabajo de investigación fue titulado "*Efecto hipolipemiante de los fitoesteroles de la dieta y su modulación por genes candidatos. Diseño y validación de un cuestionario para su medición, modelado bioinformático para el proceso integrado de datos y resultados de un estudio de intervención*". Asímismo fue coautora de dos artículos de revista sobre las características de la Bioinformática Clínica y de diversas comunicaciones en congresos nacionales e internacionales.

La descripción de los resultados obtenidos en el marco de estos proyectos, queda fuera de los objetivos de este documento de tesis, sin embargo, dado que el trabajo fue realizado en periodo de formación predoctoral, en este capítulo se ennumeran las tareas realizadas dentro de los dos proyectos principales, FITUVEROLES y OBENUTIC y se ofrece el listado de las publicaciones obtenidas.

9.1 TAREAS REALIZADAS EN EL MARCO DEL PROYECTO FITUVEROLES

El objetivo general del proyecto FITUVEROLES consistió en diseñar y validar los instrumentos necesarios para conocer la ingesta de esteroles vegetales en población mediterránea española y estimar los efectos de la suplementación con esteroles vegetales sobre los principales parámetros del metabolismo lipídico y otros fenotipos intermedios de riesgo cardiovascular relacionados, estudiando también su modulación por polimorfismos genéticos. Este objetivo general se abordó a través de las siguientes actividades específicas:

- Diseño y validación de un cuestionario de frecuencia de consumo de alimentos (CFCA) específico para conocer la ingesta de fitoesteroles en población mediterránea española, utilizando y comparando distintas tablas de composición de alimentos (TCA).
- Diseño de una infraestructura de soporte informático para la automatización de la medida de la ingesta de fitoesteroles en estudios epidemiológicos.
- Estudio de intervención para conocer el efecto de la ingesta de alimentos enriquecidos con fitoesteroles sobre parámetros del metabolismo lipídico, hidrocarbonado e inflamación en población general mediterránea española normocolesterolemia o con hipercolesterolemia moderada.
- Estudio de la modulación genética de los efectos de los fitoesteroles en el metabolismo lipídico analizando polimorfismos en genes candidatos relacionados con estos compuestos (ATP-binding cassette, sub-family G (WHITE), members 5 and 8, *ABCG5* y *ABCG8*).

9.2 TAREAS REALIZADAS EN EL MARCO DEL PROYECTO OBENUTIC

El objetivo general del proyecto OBENUTIC consistió en diseñar y validar los instrumentos necesarios para facilitar la medida de la dieta y de ingesta de nutrientes en grandes estudios epidemiológicos. Este objetivo general se abordó a través de las siguientes tareas específicas:

- Desarrollo de «NutriOntología», ontología para el alineamiento de alimentos y de sus componentes en cuatro TCAs, estableciendo una jerarquía de clases, subclases e instancias de alimentos y nutrientes. Para cada instancia de NutriOntología se

determinó el peso en gramos de tres raciones estándar, de tamaño grande, mediano y pequeño.

- Diseño puesta a punto y validación de «OntoReceta», soporte informático Web para la especificación y almacenamiento de recetas estándar.
- Diseño y validación de «OBENUTIC-R24h», infraestructura de soporte informático para administración electrónica on-line de recordatorios 24h, que ofrece un soporte computacional para el cálculo de la ingesta de alimentos y nutrientes.
- Diseño y validación de «OBENUTIC-CFCA», infraestructura de soporte informático para la administración electrónica on-line de CFCA a la carta, que ofrece un soporte computacional para el cálculo de la ingesta de alimentos y nutrientes.
- La validación y puesta a punto de las herramientas diseñadas incluyó el reclutamiento de 200 participantes de la Comunidad Valenciana, a los que se administró tres R24h y dos CFCA. OBENUTIC-R24h y OBENUTIC-CFCA se sirven del conocimiento almacenado en NutriOntología y OntoReceta para transformar la información de ingesta alimentaria de R24h y CFCA en ingesta de nutrientes.

9.3 PUBLICACIONES ADICIONALES DERIVADAS DEL TRABAJO PREDOCTORAL

Las publicaciones ennumeradas en este apartado se divididen en tres secciones: publicaciones relacionadas con la bioinformática clínica, publicaciones en el ámbito de la informática biomédica aplicada a la epidemiología nutricional y publicaciones derivadas de estudios de asociación genética.

Publicaciones relacionadas con la bioinformática clínica

Coltell O., **Arregui M**., Falomir Z., Puig A. La Bioinformática Clínica: aplicación en el ámbito de la sanidad. Todo Hospital, Julio-Agosto 2005; 218: 426-435. ISSN 0212-19721.

Coltell O, **Arregui M**, Fabregat A, Portolés O. Integration of clinical and biological data in clinical practice using bioinformatics. Rev Med Chil. 2008 May;136(5):645-52.

Coltell O, **Arregui M**; Fabregat A, Causanilles M, Vázquez R, Portolés O. Las Redes Temáticas

de Investigación Cooperativa: INBIOMED y el Nodo UJI – IRIS. I Jornadas de seguimiento del convenio UJI – UO. Santiago de Cuba. 16-20 de enero de 2006. ISBN: 978-84-8021-610-4.

Publicaciones relacionadas con la informática biomédica aplicada a la epidemiología nutricional

Proyecto FITUVEROLES

Arregui M, Coltell O, Vázquez R, Fabregat A, Portolés O, Corella D. FITUVEROLES: un portal Web piloto para la determinación de fitoesteroles ingeridos en la dieta mediante cuestionarios digitalizados. Public Health Nutrition 2006; 9(7A): 255. ISSN 1368-9800

Izquierdo-Galbis A, Alegría-Torán A, Lagarda M.J, Farré-Rovira R, **Arregui M**, Corella D. Diseño y validación de un cuestionario para estimar el aporte de esteroles vegetales a través de la dieta en población española. Public Health Nutrition 2006; 9(7A): 284. ISSN 1368-9800

Portolés O, **Arregui M**, Fabregat A, Farré R, Corella D, Coltell O. Estimación de la Ingesta de Fitoesteroles en Población Española Mediante Cuestionario Informatizado: Diferencias de Consumo egún el Origen de las Tablas de Composición Utilizadas. V Congreso de la Sociedad Española de Nutrición Básica y Aplicada SENBA. Bilbao, 25-27 de abril de 2007.

Fabregat A, Portolés O, **Arregui M**, Núñez M, Barrera E, Corella D, Coltell O. Servicios Web para la Investigación Biomédica: Aspectos de Soporte a la Gestión en el Proyecto Fituveroles. Actas del X Congreso de Informática y Salud, INFORSALUD 2007, Madrid; 2007: 133-138. ISBN: 84-610-2927-4.

Fabregat A, **Arregui M**, Portolés O, Barrera E, Andreu Y, Vázquez R, Causanilles M, Corella D, Coltell O. Servicios Web para la investigación biomédica: aspectos de gestión de seguridad y protección de datos en el proyecto FITUVEROLES. Actas del XI Congreso Nacional de Informática Médica, INFORMED 2006, Murcia, 2006:32-38. ISBN: 690-0102-7.

Coltell O, **Arregui M**, Fabregat A, Barrera E, Causanilles M, Guillén M, Corella D, Portolés O. Nutriontología: Construcción de una Ontología Nutricional Integrando Diversas Tablas de Composición de Alimentos para la Investigación en Genómica Nutricional. VI Conferencia Iberoamericana en Sistemas, Cibernética e Informática: CISCI 2007. Orlando, Florida (EE.UU.), 12-15 de julio de 2007 ISBN 1-934272-03-5.

Arregui M, Fabregat A, Barrera E, Portolés O, Corella D, Coltell O Fitosteroles y síndrome metabólico: diseño y validación de recursos informáticos para la medición de la ingesta de fitoesteroles en población española. IX Simposium Nacional de Obesidad: Aspectos Básicos y Aplicados. Reus, 24-25 abril 2008

Proyecto OBENUTIC

Arregui M, Fabregat A, Añíbarro E, Barrera E, Portolés O, Corella D, Coltell O. Instrumentos para la Investigación en Dieta y Arteriosclerosis: Integración de la Gestión de Recetas en el Recuerdo de 24 Horas por Medio de una Ontología Biomédica. Clínica e Investigación en Arteriosclerosis 2008; 20 (Supl 3): 68. ISSN: 0214-9168.

Coltell O, **Arregui M**, Fabregat A, Barrera E, Añíbarro E, Corella D. OBENUTIC-24H: computer tools for data acquisition of diet and its transformation in nutrients from 24 h recalls in nutricional studies Obesity Metab, 2009; 5(1S): 36.

Fabregat A, Barrera E, **Arregui M**, Portolés O, Corella D, Coltell O. BOGENVI: A Biomedical Ontology for Modelling Gene*Environment Interactions on Intermediate Phenotypes in Nutrigenomics Research. 21st IEEE International Symposium on Computer-Based Medical Systems (CBMS2008).Jyväskylä, Finland; June 2008; 302-307. ISBN: 978-0-7695-3165-6.

Arregui M , Fabregat A, Barrera E, Corella D, Coltell O. Experiencia del empleo de las herramientas tic «OBENUTIC-r24h», «OntoReceta» y «NutriOntología» para la determinación de la ingesta de alimentos y nutrientes en el estudio epidemiológico nutricional «OBENUTIC». XII Congreso Nacional de Informática Médica, INFORMED 2008, Santa Cruz de Tenerife, 2008: 125-132. ISBN: 978-84-691-6687-1.

A. Fabregat, M. Arregui, O. Portolés, E. Barrera1, D. Corella, O. Coltell. NutrigenOntología: ontología biomédica para la investigación en genómica nutricional integrando genotipo, ambiente y fenotipo. Actas del I Symposium del CIBER «Fisiopatología de la Obesidad y Nutrición», Santiago de Compostela; 2007: 113.

Fabregat A, Añíbarro E, **Arregui M**, Portolés O, Barrera E, Corella D, Coltell O. OntoReceta: Integración de la gestión de recetas en una Ontología Biomédica como soporte para la investigación en Genómica Nutricional. Actas del X Congreso de Informática y Salud,

INFORSALUD 2008, Madrid; 2008: 234-239. ISBN: 978-84-691-2468-0.

Coltell O, **Arregui M**, Fabregat A, Barrera E, Añíbarro E, Corella D. OBENUTIC-24H: computer tools for data acquisition of diet and its transformation in nutrients from 24 h recalls in nutricional studies. Proceedings of the II SIMPOSIUM CIBER Fisiopatología de la Obesidiad y Nutrición. Isla de la Toja, 2008: 159. ISBN: 978-84-936158-8-8.

Fabregat A, **Arregui M**, Barrera E, Portolés O, Corella D, Coltell O. NutriGeneOntology: A Biomedical Ontology for Nutrigenomics Research. The International Conference on BioMedical Engineering and Informatics (BMEI2008). International Conference on Biomedical Engineering and Informatics (BMEI), IEEE Conference #13886, Proceedings code PR3118. Sanya, China; March 2008; 915-919. ISBN: 978-0-7695-3118-2China.

Barrera, E., Antonio Fabregat Mundo, María Arregui Rementería, Corella, D., Oscar Coltell Simón. Cálculo de la Ingesta de Nutrientes en Base a Respuestas de CFCA y Ontologías de composición de Alimentos y Recetas.. XII Congreso Nacional de Informática de la Salud. Madrid: 16-03-2009. Nacional. 2009 CEFIC. ISBN: 978-84-691-9561-1.

Sorlí JV; Fabregat A; **Arregui M**; Barrera E; Corella D; Coltell O. BIGR24H: aplicación web para la valoración del aporte de energía y consumo de nutrientes en Atención Primaria. 30 Congreso de la Sociedad Española de Medicina de Familia y Comunitaria (SEMFYC). Valencia, Junio 2010 Spain 2010

Coltell O, **Arregui M**, Fabregat A, Grao E, Beltrán A, Madueño F, Ordovás JM, Corella D. Diseño y validación de Software "online" para la automatización del procesado de recuerdos de 24 h en amplios estudios epidemiológicos XXIV Congreso Nacional de la Sociedad Española de Arteriosclerosis y XIII Congreso Internacional de la Sociedad Iberolatinoamericana de Arteriosclerosis. Publication: Clin Invest Arterioscl. 2011;23(Espec Cong):55-56. Sevilla, 25-27 de mayo de 2011

Publicaciones derivadas de estudios de asociación genética

Corella D, **Arregui M**, Carrasco P, Francés F, Sorlí JV, Saiz C, Ruiz de la Fuente S, Sabater A, Coltell O, Portolés O. Ausencia de asociación del polimorfismo rs7566605 cercano al gen *INSIG2* con parámetros antropométricos y lipídicos en una población mediterránea de alto riesgo

cardiovascular. XX Congreso Nacional de la SEA. 26-29 mayo de 2007, ISSN 02149168.

Arregui M, Ortega-Azorín C, Carrasco P, Sotos-Prieto M, Guillén M, González JI, Corella D. Impact of the C to T substitution at -13910bp upstream the lactase gene (rs4988235) on consumption of dairy products and obesity risk in a Spanish Mediterranean population. Proceedings of the II SIMPOSIUM CIBER Fisiopatología de la Obesidiad y Nutrición. Isla de la Toja, 2008: 159. ISBN: 978-84-936158-8-8.

Ortega-Azorín, C., Godoy, D., Sorlí, J.V., **Arregui M**, Ordovás, J.M., Corella, D.. Elevado desequilibrio de ligameiento de los polimorfismos rs17145738 y rs3812316 del Gen *MLXIPL* y su asociación con triglicéridos plasmáticos en población mediterránea de alto riesgo cardiovalscular. Clínica e investigación en arteriosclerosis. Num. Vol 21. pp. 27. 2009 Internacional.

Corella D, Portolés O, **Arregui M**, Guillen M, Ortega C, Carrasco P, Sotos M, Guillén-saiz P, Ordovas JM. *TAS2R38* Polymorphisms Bitter Perception, Food Intake and Obesity Risk in a Mediterranean Population. Journal of Nutrigenetics and Nutrigenomics 2008: 290-291. ISSN 1661-6499. On-line ISSN: 1661-6758

Corella, D, Carrasco, P, Ortega C, Portoles O, Guillen M, Coltell O, **Arregui M**, Estruch R, Ordovas, JM, Sorli JV. Study of Genetic Variation in *FTO* and *MC4R* with food intake and body mass index in a Mediterranean Population. 3 congress onf the International Society of Nutrigenetics/Nutrigenomics. Bethesda, Maryland: 21-10-2009. Internacional. 2009 . ISBN:

Arregui M, Coltell O, Portolés O, Sorlí JV, Verdú JJ, Corella D. Estudio de la asociación del polimorfismo rs10883463 del gen de la EstearoilCoA-1 desaturasa (*SCD1*) con medidas antropométricas y bioquímicas en población mediterránea española. X Edición del Simposium Nacional de Obesidad: aspectos básicos y aplicados, (Simposium Satélite del II Congreso FESNAD) Nutr Hosp, 2010; 25 (3): 493 Reus, 1-2 Marzo 2010 : Spain 2010

Arregui M, Coltell O, Portolés O, Asensio EM, Olivares L, Corella D. Asociación del polimorfismo rs3754686 del gen *MCM6* con peso y obesidad en población mediterránea española. X Edición del Simposium Nacional de Obesidad: aspectos básicos y aplicados, (Simposium Satélite del II Congreso FESNAD) **Publication:** Nutr Hosp, 2010; 25 (3): 491 Reus, 1-2 Marzo 2010 **Spain** 2010

Arregui M; Estruch R; Salas J; Coltell O; Covas MI; Ros E; Guillem P; Osma R; Corella D. Influencia de las variantes rs198358 (T>C) y rs5068 (A>G) en el gen *NPPA* (precursor del péptido natriuretico A) en la hipertensión arterial en población de alto riesgo cardiovascular. XXVIII Reunión Cientifica de la Sociedad Española de Epidemiología. Gac Sanit, 2010, 24 (Esp Congr 2): 139. Valencia, 27-29 octubre 2010

INDEX OF FIGURES

Chapter 2

INDEX OF TABLES

Chapter 6