Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN:
SCIENZE FARMACOLOGICHE E TOSSICOLOGICHE, DELLO SVILUPPO E DEL MOVIMENTO UMANO

Ciclo XXVIII

Settore Concorsuale di afferenza: 06/B1

Settore Scientifico disciplinare: MED/09

TITOLO TESI:
CARDIOMETABOLIC DISEASES' RISK THROUGH POPULATION GENETIC STUDIES: HISTORICAL, PRESENT AND FUTURE RESOURCES OF THE BRISIGHELLA BIOBANK

Presentata da:
Dott. ssa Martina Rosticci

Coordinatore Dottorato

Chiar.ma
Prof.ssa Patrizia Hrelia

Relatore

Chiar.mo
Prof. Claudio Borghi

Esame finale anno 2016

Contents

List of Tables

List of Figures

Keywords

Lipids, statins, cardiometabolic, SNPs, association, *KIF6*, epidemiological study, mendelian randomization

Abstract

Cardiovascular diseases (CVD) comprise the most common chronic disease worldwide. High lipid levels are a strong risk factor, making lipid-lowering statin therapy an important preventive measure.

Here we explore the effects of common variants at the *KIF6* and *HMGCR* loci on a range of cardio-metabolic traits and on response to statin therapy. While *HMGCR* is a well-established lipid-related locus, the role of *KIF6* in response to statin therapy is controversial, and its contribution to related phenotype variability has not been clarified.

We genotyped a coding *KIF6* variant (p.W719R, rs20455) and two intronic ones in high LD to the former (rs9462535,rs9471077), as well as two non-coding variants in *HMGCR* (rs3761740 and rs3846662). Effects on 14 quantitative and 5 categorical cardiometabolic phenotypes including lipid-lowering therapy response were tested in a sample of 1,645 individuals from the Genetics in Brisighella Health Study (GBHS) from Italy and replicated in 10,662 individuals from the Estonian Genome Center (EGCUT).

In GBHS the established *HMGCR* variant rs3846662 affects LDL cholesterol levels ($\boldsymbol{P}$=8.5×10^{-4}) while the intronic *KIF6* variant rs9471077 modifies APOB levels (P=8.2×10^{-4}). The latter association was confirmed in EGCUT. No significant association between *KIF6* variants and response to statin therapy was observed.

In the first genetic study involving GBHS we confirm the *HMGCR* effect on LDL-Cholesterol and demonstrate a novel *KIF6* effect on APOB. The latter association needs to be evaluated for its predictive value for overall CVD risk and its potential contribution to stratified patient care.

1. Epidemiologic studies in Italy and throughout the world

1.1 Epidemiological approaches to heart disease

The use of the word "epidemiology" and the concept of what epidemiology as a discipline may encompass has varied widely since the days of Peter Panum and John Snow. There are today many differing definitions of the word, but nearly all workers in the field will agree on one element of the definition: The word " epidemiology " by etymology refers to the study of something " which is thrust upon the people." There are still some who insist that epidemiology deals only with epidemics of infectious diseases, but current usage suggests that most workers would now agree that epidemiology deals with "the fundamental questions as to where a given disease is found, when it thrives, where and when it is not found in other words it is the ecology of disease"[1]without regard to whether the disease is believed to be infectious. Frost gave an analytical definition when he wrote that epidemiology "includes the orderly arrangement of facts into chains of inference which extend more or less beyond the bounds of direct observation. His definition might be called the essence of the" epidemiological method" except for the fact that it has been used by the physician since the time of Hippocrates to arrive at his clinical diagnosis. Thus, today, the epidemiological approach is used to explore certain relationships in health and disease which, with present technological methods, cannot be observed directly. In the field of cardiovascular diseases, studies using the epidemiological method have led to findings of considerable practical importance for prevention and treatment.

1.2 The Brisighella Heart Study

The Brisighella Heart Study is a prospective, population-based longitudinal epidemiological cohort involving 2939 randomly selected subjects, aged 14 to 84 years, free of cardiovascular disease at enrolment, resident in the Northern Italian

rural town of Brisighella. The study was promoted in 1972 by Professor G. Descovich[2]**.** Subjects were clinically evaluated at baseline and every 4 years following enrolment when extensive clinical and laboratory data were obtained in addition to the assessment of morbidity and mortality. In 1986, the study became part of the WHO European Risk Factors Coordinated Analysis, and in 1990, it became part of the Risk Factors and Life Expectancy Project[3]. Throughout the duration of the entire study, all-cause mortality and morbidity, as well as the incidence of the main cardiovascular risk factors, were recorded. Every three months, the study design included an update of the database with regard to fatal and nonfatal new events and every four years, a complete medical check-up comprised a nutritional habits record and fasting blood sample was performed. From 1986 to 1988, several programs started to check efficacy, cost, and reliability of primary and secondary cardiovascular prevention, including school children and general population nutritional education programs and general practitioner training concerning therapeutic guidelines**.** Physical activity and nutritional habits have been recorded throughout the study and encoded as previously reported**.** The study was carried out in agreement with the Declaration of Helsinki. It was approved by the Ethical Committee, and all subjects gave their written consent to be involved in the study

1.3 Framingham Heart Study

Cardiovascular disease (CVD) is the leading cause of death and serious illness in the United States. In 1948, the Framingham Heart Study - under the direction of the National Heart Institute (now known as the National Heart, Lung, and Blood Institute or NHLBI) - embarked on an ambitious project in health research. At the time, little was known about the general causes of heart disease and stroke, but the death rates for CVD had been increasing steadily since the beginning of the century and had become an American epidemic. The Framingham Heart Study became a joint project of the National Heart, Lung and Blood Institute and Boston University. The objective

of the Framingham Heart Study was to identify the common factors or characteristics that contribute to CVD by following its development over a long period of time in a large group of participants who had not yet developed overt symptoms of CVD or suffered a heart attack or stroke. The researchers recruited 5,209 men and women between the ages of 30 and 62 from the town of Framingham, Massachusetts, and began the first round of extensive physical examinations and lifestyle interviews that they would later analyze for common patterns related to CVD development. Since 1948, the subjects have continued to return to the study every two years for a detailed medical history, physical examination, and laboratory tests, and in 1971, the Study enrolled a second generation - 5,124 of the original participants' adult children and their spouses - to participate in similar examinations.

In 1994, the need to establish a new study reflecting a more diverse community of Framingham was recognized, and the first Omni cohort of the Framingham Heart Study was enrolled.

In April 2002 the Study entered a new phase, the enrollment of a third generation of participants, the grandchildren of the Original Cohort. In 2003, a second group of Omni participants was enrolled. Over the years, careful monitoring of the Framingham Study population has led to the identification of the major CVD risk factors - high blood pressure, high blood cholesterol, smoking, obesity, diabetes, and physical inactivity - as well as a great deal of valuable information on the effects of related factors such as blood triglyceride and HDL cholesterol levels, age, gender, and psychosocial issues. Although the Framingham cohort is primarily Caucasian, the importance of the major CVD risk factors identified in this group have been shown in other studies to apply almost universally among racial and ethnic groups, even though the patterns of distribution may vary from group to group. In the past half century, the Study has produced approximately 1,200 articles in leading medical journals. The concept of CVD risk factors has become an integral part of the modern medical curriculum and has led to the development of effective treatment and preventive strategies in clinical practice.

The Framingham Heart Study continues to make important scientific contributions by enhancing its research capabilities and capitalizing on its inherent resources. New diagnostic technologies, such as echocardiography (an ultrasound examination of the heart), carotid artery ultrasound, magnetic resonance imaging of the heart and brain, CT scans of the heart and its vessels and bone densitometry (for monitoring osteoporosis), have been integrated into past and ongoing protocols.

While pursuing the Study's established research goals, the NHLBI and the Framingham investigators are expanding their research into other areas such as the role of genetic factors in CVD. One project under way will utilize genetic material from immortalized cell lines of all of our cohorts. Framingham investigators also collaborate with leading researchers from around the country and throughout the world on projects in stroke and dementia, osteoporosis and arthritis, nutrition, diabetes, eye diseases, hearing disorders, lung diseases, and genetic patterns of common diseases. The unflagging commitment of the research participants in the NHLBI Framingham Heart Study has made more than a half century of research success possible. We continue with an ambitious research agenda and look forward to new discoveries in the decades to come.[4]

1.4 KORA Study

Referent subjects were drawn from the KORA S4 study, with ages ranging from 25-74 years. The KORA S4 study is a population-based epidemiological survey of persons living in or near the city of Augsburg, Southern Germany conducted between 1999 and 2001. The survey population consisted of German nationality residents born between July 1, 1925 and June 30, 1975 identified through the registration office. A sample of 6640 participants was drawn with ten strata of equal size according to sex and age, and 4261 individuals (66.8%) agreed to participate.

KORA - The Kooperative Gesundheitsforschung in der Region Augsburg Study (KORA) is a series of population-based epidemiological surveys of persons between

25 and 75 years old at the time of enrollment living in or near the city of Augsburg, Southern Germany. KORA F3 was conducted between 2004 and 2005, and KORA S4 was conducted between 1999 and 2001. In 1644 randomly selected individuals from F3 and 1027 from S4, genome-wide genotyping was performed. In KORA F3 and S4, 12-lead resting ECGs were recorded with digital recording systems (F3: Mortara Portrait, Mortara Inc., Milwaukee, USA, S4: Hörmann Bioset 9000, Hörmann Medizinelektronik, Germany). PR interval was measured automatically. In addition all ECGs were visually inspected for technical errors and inadequate quality. The Mortara portrait determines PR intervals by a proprietary algorithm, PR intervals from Hörmann Bioset were determined using the Hannover ECG analysis software (HES-Version 3.22-12) by computerized analysis of an averaged cycle computed from all cycles of the 10 second recording after exclusion of ectopic beats. PR intervals determined by this algorithm represents the earliest begin of atrial depolarization until the earliest deflection of ventricular depolarization between any two leads. In an international validation study the HES-software was among the best performing digital ECG systems. Reproducibility of HES measurements over short and long term time intervals has been investigated and KORA data have been used in several genome-wide association studies of quantitative EKG traits.[5]

1.5 Chris Study

Background: The meaning of CHRIS (Cooperative Health Research In South Tyrol) captures the essence and purpose of the project itself: it is a prospective epidemiological research study centered around the health of people living in South Tyrol. It represents a true partnership between the people participating, the staff working in the health care system and the research personnel of the Centre for Biomedicine at EURAC.

The prevalence of chronic diseases is slowly but constantly increasing, due both to the aging of the population and to better diagnostic procedures. At the same time,

diagnosis is often made after the illness is established, when preventative medicine has less chance of success. In many cases no effective preventative measures are known. Common diseases such as cardiac or respiratory disorders, diabetes or neurological conditions have many unanswered questions: What causes such diseases? What role does lifestyle, genes or a person's environment play in the onset and progression? Are there interactions between any of the influential factors for positive or negative outcomes?

Prospective cohorts are perfect models for epidemiological studies: they are large groups of people that are followed over time (every 4/5 years) and from whom clinical parameters, medical history, and life-style information, such as diet, physical activity, and exposure to environmental factors (e.g. coffee drinking and smoking habits), are progressively collected. With such information we can begin to assess the extent to which some of these factors, alone or together with genetic susceptibility, can explain some of these common diseases, their severity or lack thereof.

The CHRIS Study is an ambitious project that raises from the collaboration between the Local Health System and EURAC, and that aims to involve individuals aged 18 or older by taking part of an interview, enrolling in a quick free clinical assessment, and by providing a small sample of blood.

Participation in the CHRIS Study is completely voluntary. Individuals participating will get a feedback on some of the health parameters being measured, taking the first step towards proactive prevention and improved personal health. Equally important, their participation will provide the local health system with insights for the creation of preventive medicine plans, laying the foundation for improved health care for all of the people of South Tyrol. Finally, participants will aid in the creation of what we hope will be one of the most important research centers in the medical field focusing on neurological and cardiovascular diseases. Epidemiological studies need several hundreds or thousands of people to confirm the goodness of their results.

For this reason the CHRIS Study will also bring its contribute to a much larger epidemiological study that will be ongoing in Germany (German National Cohort),

which aims to answer the same questions by collecting information from approximately 200,000 individuals. The CHRIS Study started in August 2011.[6]

2. Biobanks of epidemiological studies

2.1 The Biobanks

Over the past two decades, the explosion in biomedical research since the 1950s, together with important technical advances in high-throughput analysis, has resulted in the creation of an increasing number of facilities for the long-term storage of human cell and tissue samples for research. Despite significant differences in approach, purpose, scale, and scope, these facilities are now collectively known as "biomolecular resource collections" or simply "biobanks".

When "biobanks" were first reported in the European science press almost 15 years ago, they were greeted with a healthy dose of scepticism and, on occasion, with outright fear. At least to some, the proposal by the Icelandic biotechnology firm deCODE to build an exhaustive database of genetic linkages between all Icelanders raised fears of a new kind of "big brother", with access to our family history and genetic makeup and funded by US American venture capitalists.

Since then, much has changed. For example, large biobank projects throughout the world are now aiming at sequencing the entire genome of tens of thousands of individuals. At about the time deCODE started its endeavour, the private company Celera Genomics raised around US$300 million to assemble the first full sequence of a single human genome, while a parallel public sector project spent probably billions of US$ on the same task. Over the past years, the cost of genome sequencing has dropped considerably and, if we are to believe the announcements from a number of companies developing new sequencing equipment, in only a few years the cost of sequencing the entire genome of a person will cost less then a few hundred US$.

But, while inexpensive genome sequencing may well be a necessary prerequisite for a new area of genome-informed, personalized medical care, cheap sequencing alone is far from sufficient when it comes to extracting real medical benefits from genome research. Thus, apart from a few singular cases, for many years to come the main

beneficiaries of inexpensive genome sequencing will not be patients but, rather, biomedical scientists.

In this new world of clinical genomics, where the entire genomes of tens of thousands of individuals will be readily available, the role and importance of biobanks which link physiological samples to medical and biomedical information, such as whole genome sequence data will become ever more important. But, future biobanks are no longer simply organizations that collect, and store, peripheral blood samples, and then make them available to scientists and, in this fashion, limit the potential risks ranging from a breach of confidentiality to the violation of a informed consent agreement that an investigator faces when dealing with human-derived materials or genetic information. Rather, we believe biobanks will increasingly turn into strong and trusted partners of both medical scientists and the public and, especially, the invisible community of biobank "donors" or "participants". * In this report we advocate a novel approach toward biobank governance; an approach where biobanks are no more simply service institutions for the scientific community, but active, and reliable partners of both the scientific community and the public in the pursuit of genomic medicine. As we have seen, biobanks have a long history, starting with the pathology collections of the late 18th and early 19th century. Further, with the rise of biomedical research during the second half of the 20th century, first in the United States and more recently in Europe, numerous collections of human materials for specific research purposes were created.

There exist several hundred human materials collections in Europe today, and these number does not include smaller collections that individual scientists have gathered during the course of their research. The vast majority of these collections is linked to a single research laboratory, or department, and has been built over time, often as part of a long-term research objective. The users of these collections are typically limited to a small number of scientists, and rarely extend beyond the immediate collaboration network of those collecting the resources. Most of these small tissue banks were started by scientists in areas such as cancer research, where human tissues are used

extensively. Many of these collections remain limited in size, and investment, and are typically known only to insiders, or the agencies or institutions that fund them. And, until very recently, few would actually label themselves as “biobanks”.[7]

2.2 BBMRI-ERIC

Essential for the understanding of the diversity of human disease, biological samples and corresponding data are required for the development of any new drug or diagnostic assay and are therefore critical for the advancement in health research, ultimately leading to personalized medicine. Biobanks also will provide key information on the influence of environment and lifestyle on health constituting a basis for disease prevention programmes and the improvement of public health.

Hence a close collaboration between researchers, biobankers, patient advocacy groups, and biotech and pharma industry is essential in addressing both common and rare disease. Keeping in mind the need for better prevention, diagnostics, and therapy for all we are aware that every single sample impacts our ability to comprehend disease and, thus, achieve our goal for healthier life. Sixteen Member States and one International Organisation have thus joined forces in establishing BBMRI-ERIC, which is one of the largest health Research Infrastructure in Europe today and has recently been joined by the United Kingdom. BBMRI-ERIC primarily aims at establishing, operating, and developing a pan-European distributed research infrastructure of biobanks and biomolecular resources. This will facilitate the access to biological resources as well as biomedical facilities and support high-quality biomolecular and medical research.

The mission

BBMRI-ERIC will increase efficacy and excellence of European bio-medical research by facilitating access to quality-defined human health/disease-relevant biological resourced through.

- the inclusion of associated data in a efficient and ethically and legally compliant manner
- by reducing the fragmentation of the bio-medical research landscape through harmonisation of procedures, implementation of common standards and fostering high-level collaboration
- by capacity- building in countries with less developed biobanking communities thereby contributing to Europe's cohesion policy and strengthening the ERA.

The History

The ESFRI Roadmap: 2006

The European strategy Forum on Research Infrastructures developed and published the European Roadmap on Research Infrastructure to propose research facilities of pan-European interest, among them BBMRI.

The Preparatory Phase: 2008-2011

In 2008, BBMRI was one of the first projects entering the Preparatory Phase (PP)of the ESFRI Roadmap for Research Infrastructures. Funded by the European Commission's Framework Programme 7, BBMRI-PP was granted 5M Euro to conceptualise and secure funding for the construction of BBMRI. Within three years, BBMRI-PP grew into a 54 member consortium with more than 225 associated

organisations from over countries, making it one of the largest research infrastructure projects in Europe.

The project content and action plan of BBMRI-PP were defined in the Grant Agreement with the European Commision. The seven Work Packages of BBMRI-PP were responsible for the specific deliverables, with the goal of integrating the existing quality-controlled biobanks, biomolecular resources and enabling technologies into a novel pan-European biomedical research infrastructure. The major achievements of BBMRI-PP are summarised in the Final Report to the European Commision, including its operational concept as outlined in the Business Plan.

The interim Phase: 2011-2013

After three years, the BBMRI-PP project came to its end 31 January 2011, including the Governance and Management Structures, which were based on the Grant Agreement. In its final teleconference on 25 January 2011, the BBMRI-PP. Steering Committee agreed that the Coordinator, the Executive Manager as well as the steering Committee of Preparatory Phase should continue to function as interim bodies of BBMRI until the preparatory governance body was established.

Building on achievements of BBMRI-PP, the FP7-project BBMRI_LPC aims at helping scientist to access large prospective study sets with the vision that many parts will eventually integrate in BBMRI-ERIC.

Awarded EU Legal Status: 3 December 2013

After finalising the negotiations among interested member states and the ERIC application process, BBMRI-ERIC was founded with the publication the BBMRI-ERIC statutes in the Official Journal of the European Union and entered into force three days after publication on 3 December 2013*(Figure 1*)

Therewith BBMRI was officially awarded the community legal framework for a European research infrastructure consortium (ERIC). This specific legal form is

designed to facilitate the joint establishment and operation of research infrastructures of European interest. The Eric status allows pulling together biobanks and biomolecular resources into a pan-European facility and providing access to collections of partner biobanks and biomolecular resources, their expertise and services on a non-economic basis. On 21 January 2014, on the occasion of the European conference dedicated to the launch of Horizon 2020 in Vienna.

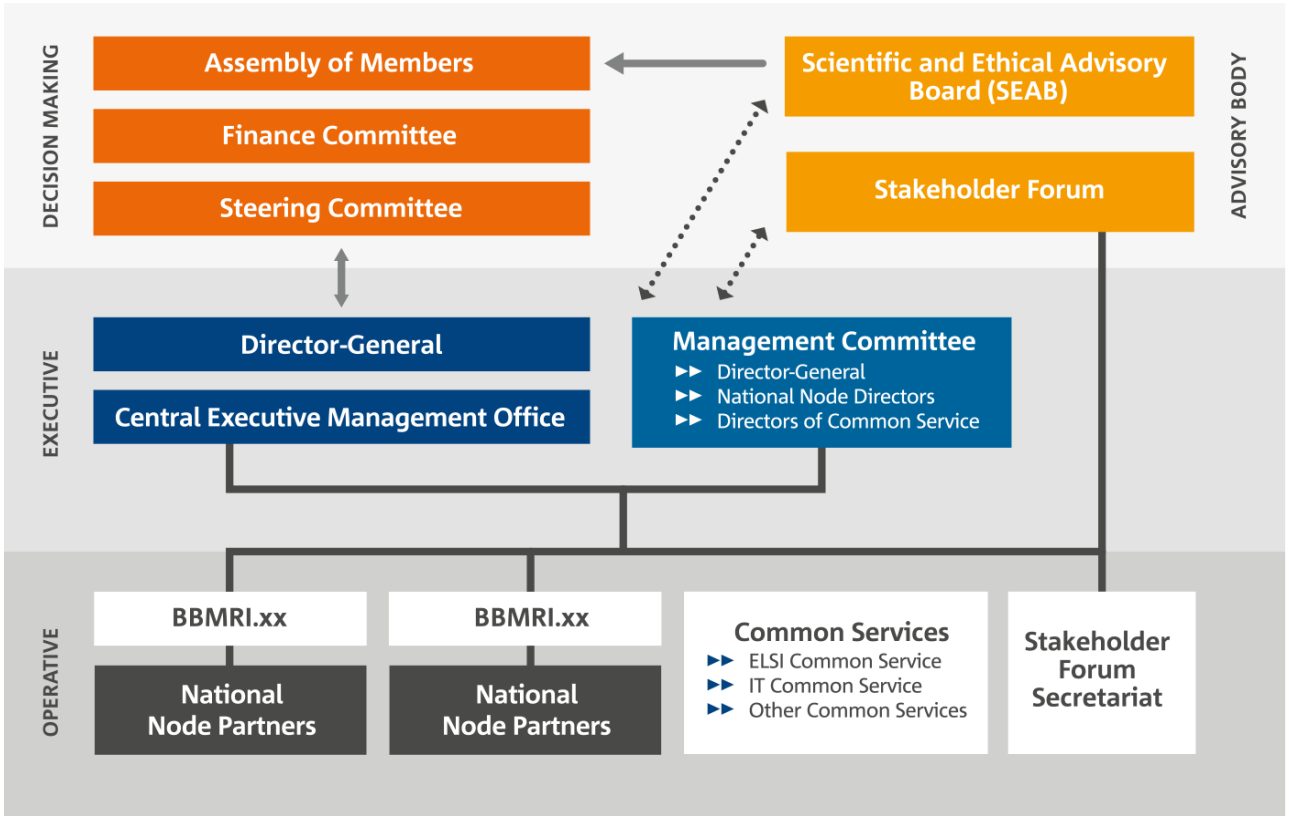

Figure 1.The governance structure of bbmri-eric

3. Genetic polimorphisms modifying susceptibility to metabolic diseases and pharmacological therapy response

3.1 Lipids and lipid metabolism

The liver plays a key role in lipid metabolism. Depending on species it is, more or less, the hub of fatty acid synthesis and lipid circulation through lipoprotein synthesis. Eventually the accumulation of lipid droplets into the hepatocytes results in hepatic steatosis, which may develop as a consequence of multiple dysfunctions such as alterations in b-oxidation, very low density lipoprotein secretion, and pathways involved in the synthesis of fatty acids. In addition an increased circulating pool of non-esterified fatty acid may also to be a major determinant in the pathogenesis fatty liver disease.

Lipid metabolism involves several pathways that are at least in part, inter-dependent and 'cross-regulated'. The focus of the present discussion will be fatty acids and triacylglycerols. Fatty acids are the most commonly stored and circulating forms of energy, and triacylglycerols are the most common non-toxic form of fatty acids. Fatty acids/triacylglycerols may originate from four sources (pool input): De novo lipogenesis, cytoplasmic triacylglycerol stores, fatty acids derived from triacylglycerols of lipoprotein remnants directly taken up by the liver, and plasma non-esterified fatty acids (NEFA) released by adipose tissue. The relative importance of these sources depends on species differences (e.g. in ruminants, only modest amounts of hepatic De novo lipogenesis occurs com- pared with adipose tissue; the inverse is true in birds, with liver being the main site of De novo lipogenesis), and on short- and long-term nutritional status and energy balance. Fatty acids and triacylglycerols may also be used in different ways (pool out- put). Triacylglycerols may accumulate in hepatocytes (while NEFA or activated forms of NEFA may not) unless NEFA are oxidized (more or less completely) or triacylglycerols are exported as constituents of very low density lipoproteins (VLDL). Two examples of inter-connection may be cited: (i) a low rate of esterification when the oxidation rate is high in response to

energy demand, (ii) a strong relationship between VLDL secretion and fatty acid/triacylglycerol availability; this is especially the case in species where De novo lipogenesis is very active, but not in those species where high triacylglycerol concentration may be present and where the liver is not the 'usual' site of De novo lipogenesis. The triacylglycerol content of hepatocytes is regulated by the activity of cellular molecules that facilitates hepatic fatty acid uptake, fatty acid synthesis, and esterification ('input') and hepatic fatty acid oxidation and triacylglycerol export ('output'). Moreover, and interestingly, fatty acids regulate overall lipid metabolism by binding nuclear receptors that modulate gene transcription.

Lipids are a group of natural molecules that include fats, waxes, sterols, of which the most important is cholesterol, soluble vitamins (such as A, D, E and K), monoglycerides, diglycerides, triglycerides, phospholipids. Dietary lipids of physiological and especially pathological importance include triglycerides, which account for 90% of dietary lipid, cholesterol esters, phospholipids and the fat soluble vitamins. Lipids represent the most concentrated source of calories in the diet.

The human diets contain variable amount of triglyceride and cholesterol provided by diet. Typical Western diets provide as much as 40% of the total calories in the form of triglyceride, but, on a worldwide scale, this type of diet is geographically and historically unusual. Most dietary triglyceride is absorbed in the duodenum and proximal jejunum after undergoing partial hydrolysis in the gut lumen. Triglycerides are mechanically mixed with the aqueous secretions of gastrointestinal tract to form large fat droplets. Then a smaller emulsion of smaller fat droplets is formed as bile acids, and phospholipids from the diet and bile become associated with the droplet surface. These substances are amphophilic, i.e., partly hydrophilic and partly hydrophobic, and thus promote the formation of a stable oil[8]water interface. These fat droplets are continuously exposed to the enzymatic activity of lipases. This enzymatic activity releases monoglycerides and free fatty acid (FFAs). At the same time, action of a pancreatic phospholipase on the phospholipid associated with the droplet surface leads to the formation of lysophospholipids and FFAs. The FFAs and

monoglycerides that are converted back to triglycerides are those that have chains at least 14 carbon atoms long. Shorter chain fatty acid largely pass through the epithelial cells and are transported via the portal vein to the liver[9],[10]. Once reached the bloodstream, chylomicrons are hydrolysed by lipoprotein lipase (LPL) releasing triglycerides, which in turn will be absorbed to the luminal surface of capillaries. Most of FFAs released in this way are taken up by adipose cells, and stored as adipose tissue triglyceride. Over the passages chylomicrons change in composition; they lose phospholipid and apo-lipoproteins AI, AII and AIV and take up cholesterol and apo-lipoproteins CI, CII, CIII and apo E. Apo CII it is known to be particularly important because it affects the clearance of chylomicron triglyceride from the plasma, by activating LPL. Thus, patients with a familial apo CII deficiency develop hypertriglyceridemia and hyperlipidaemia when they ingest fat because chylomicrons accumulate in the plasma[11],[12]. The other lipid particularly important is a key component of cell membrane.

Its role in the membrane composition seems to be due apparently to its amphophilic character and its unique, rigid structure, that allows it to intercalate between membrane phospholipids. This intercalation markedly decreases the permeability of membranes to water-soluble molecules and also decreases membrane fluidity[13]. Although a dietary requirement for cholesterol doesn't exist, cholesterol introduced with diet significantly participates to a pool of cholesterol and its bile acid products that circulate several times each day between the intestine and the liver. Cholesterol introduced daily by diet can range up to 0.5/1.0 g/day and most part of this cholesterol is in the form of cholesteryl ester and probably is not directly absorbed.

Cholesterol absorption is incomplete and only 30-60% seems to enter body pools[14].It seems to be likely that cholesterol absorption is realized by soluble, lipid-carrier protein that bring it from the microvillus membrane to the intracellular site of lipoprotein synthesis. The cholesterol absorbed, once within the epithelial cells, mixes with the intracellular pool of cholesterol, is esterified again with fatty acid and enters the chylomicrons, which also contain small amount of free cholesterol. After

the chylomicrons enter the plasma and are attacked by LPL, much of this cholesterol, particularly the cholesterol ester, becomes associated with chylomicron remnants. The remainder of the cholesterol and also the remaining phospholipid apparently become associated with high density lipoprotein (HDL). Chylomicron remnants contain apo- E, which interacts with its receptors localized on the surface of hepatocytes, and thanks to this interaction chylomicrons are rapidly removed from plasma into hepatocytes[15].The ingested particles are hydrolysed within secondary lysosomes to yield amino acids, FFAs, and unesterified cholesterol. One effect of the influx of chylomicron remnant cholesterol into the liver is the decreased synthesis of endogenous cholesterol because of reduced hydroxyl- methylglutaryl-CoA reductase (HMGCoAreductase) activity. This is the rate-limiting enzyme of cholesterol biosynthesis and catalyses the conversion of HMG CoA to mevalonic acid, the first committed metabolite in the biosynthesis of cholesterol. It is subject to multivalent feedback suppression by sterols and by non-sterol products of mevalonic acid[16],as well as to phosphorylation and dephosphorylation[17].Down regulation and inactivation of HMG-CoA reductase by these mechanisms leads to diminished formation of hepatic cholesterol, which limits the tendency of dietary cholesterol to increase hepatic cholesterol levels. The most part of cholesterol synthesised ex novo by the liver or introduced with the diet is either converted into bile acids or secreted directly into bile[18]. One type of lipoprotein secreted is very low density lipoprotein (VLDL), also involved in the triglyceride transport; indeed, VLDL of hepatic origin consists of a core of triglyceride stabilized by a thin film of phospholipid, unesterified cholesterol, apo- lipoproteins B100, C, and E. Into the peripheral tissue VLDLs are metabolized by LPL, that catalyses the partial hydrolysis of the triglyceride, producing remnant lipoproteins. These cholesterol-rich remnants may be taken up directly by the liver or may continue to circulate in the plasma and be gradually converted into small lipoproteins known as low-density lipoproteins (LDLs) that delivery cholesterol to peripheral cells. LDL show a diameter that is one fourth to one third of the diameter of the parent VLDL, and show a core that is made chiefly by

cholesteryl ester surrounded by phospholipid, unesterified cholesterol, and apo-B. The net effect is the formation of a cholesterol-rich lipoprotein that is small enough to be transported across the endothelial cells of peripheral capillaries. The conversion of VLDL to LDL usually requires about 12 h, after which LDLs are gradually cleared from the bloodstream, principally by two mechanisms, one receptor-dependent, and another nonspecific mechanism such as pinocytosis. The best understood mechanism is that mediated by LDL receptor, which bind lipoproteins that contain apo-B and/or apo-E with a high affinity. Once interaction between receptor and lipoprotein is realized, complex is internalized by adsorptive endocytosis. Vesicles formed by this process subsequently fuse with lysosomes, whereupon the lipoproteins are degraded by lysosomal hydrolases and unesterified cholesterol is released into the cytosol. As cytosolic cholesterol accumulates, it is esterified by an intracellular acyl CoA-cholesterol acyltransferase (ACAT) or used to form membranes. At the same time, increased cytoplasmatic cholesterol levels, activates feedback mechanisms that reduce intracellular HMG-CoA reductase activity and down-regulate the LDL receptor. These feedback mechanisms clearly limit both the intracellular synthesis of cholesterol from acetyl CoA and the uptake of excessive amounts of LDL cholesterol. In the LDL clearance process an important role is played by LDL-receptor and apo-B and E, whose importance is emphasized by the strikingly high concentrations of LDL that are typically found in the plasma of patients affected by familial hypercholesterolemia[19].

Familial hypercholesterolemia may be caused by incapacity of our body to form functional apo-B and E receptors and/or inability to internalize lipoproteins. Cells that normally show LDL receptors include fibroblast, smooth muscle cells, adrenocortical cells, luteal cells from the ovary. Thus, these receptors are widely distributed over tissues and organs, accounting for about two thirds of the removal of LDL particles from plasma[19]. The LDL receptors are usually localized into cytoplasmatic vesicles and are subject to a recycling process, according to the intracellular requirements for cholesterol. For example, since LDL it has been seen to

play a key role in delivering cholesterol to endocrine cells that synthesise steroid hormones, under corticotropin stimulation the LDL receptor number increases highly[20].

In addition to secreting VLDL, the liver secretes lipoproteins referred to as nascent high density lipoprotein (HDL), a heterogeneous class of lipoproteins, which have in common a high density (>1.063 g/mL) and a small size. These lipoproteins contain phosphatidilcoline, unesterified cholesterol, and apo-AI or E. Mature HDL3 and HDL2 are synthesised from lipid-free apo-AI or lipid-poor pre-β1-HDL as the precursor. These precursors are released from lipolysed VLDL and chylomicrons or are produced as a nascent HDL by the liver orintestine. After being secreted into plasma, nascent HDLs interact with a plasma enzyme that also is synthesised and secreted by the liver. This enzyme, lecithin-cholesterol acyltransferase (LCAT), forms cholesteryl esters in plasma by transferring fatty acids from HDL phosphatidilcoline to HDL unesterified cholesterol. The LCAT enzyme is activated by apolipoproteins, in particular by apo-AI, that is the principal apo-lipoprotein component of mature, circulating HDL. Differences in the quantitative and qualitative content of lipids, apolipoproteins, enzymes, and lipid transfer proteins result in the presence of various HDL subclasses, which are characterized by differences in shape, density, size, charge, and antigenicity[21].Lipoproteins not only deliver cholesterol to cells but apparently also contribute to reverse cholesterol transport (RCT). Cholesterol in peripheral cells in the form of cholesteryl ester is hydrolysed by an intracellular cholesteryl ester hydrolase[22]. Then, on interaction of cells with HDL that is transiently bound to HDL receptors on the cell surface[23], the liberated unesterified cholesterol transfers to the plasma membrane and becomes associated with the HDL. RCT describes the metabolism and an important anti-atherogenic activity of HDL, namely, the HDL-mediated efflux of free cholesterol from non-hepatic cells and its subsequent delivery to the liver and steroidogenic organs, in which cholesterol is used as a precursor for the synthesis of lipoproteins, bile acids, vitamin D, and steroid hormones[24],[25].

Approximately 9 mg cholesterol per kg body weight is synthesized by peripheral tissues every day and must be moved to liver for effective catabolism[26]. Alteration of RCT can lead to a deposition of free cholesterol (FC) within the arterial wall and thereby contributes to the development of arteriosclerosis. Evidence that HDLs contribute importantly to reverse cholesterol transport is provided by the abnormalities that accompany Tangier disease. This inborn error of metabolism is characterized by abnormally low level of HDL in the plasma and by the presence of cholesteryl ester-rich foam cells (macrophages) in peripheral lymph nodes. Several lines of evidence suggest that cholesterol efflux may be realized by three different mechanisms: aqueous diffusion, SR-BI-mediated FC efflux, and ABCA1-mediated efflux. Cholesterol molecules are sufficiently water-soluble to transfer from cell membrane to an acceptor by the aqueous diffusion mechanism[27]. This process involves desorption of cholesterol molecules from the donor lipid-water interface and diffusion of these molecules. through the intervening aqueous phase until they collide with and are absorbed by an acceptor. The aqueous diffusion mechanism involves a simple diffusion process, and, as such, cholesterol transfer is passive and driven by the cholesterol concentration gradient. In this mechanism process, the unstirred water layer surrounding cells creates a significant diffusion barrier[28]. The other mechanism of cholesterol efflux is based on the interaction between acceptor and SR-BI receptors of donor. These receptors are members of the CD36 family of proteins and share about 30% sequence homology with the other members of this family[2930]. Besides stimulating the efflux of free cholesterol, expression of SR-BI also facilitates the influx of free cholesterol, thus, the movement of free cholesterol is bidirectional and the net movement depend on the cholesterol concentration gradient, like to the aqueous diffusion. Furthermore, besides mediating the bidirectional flux of free cholesterol, SR-BI is able to induce a selective uptake of other lipoprotein lipids, including cholesteryl ester, phospholipid, and triglyceride[31].This movement is unidirectional, and by promoting the net flux of HDL cholesteryl ester and triglyceride, SR-BI induces depletion of HDL core lipids. SR-BI receptors interact

with a wide range of acceptors, including HDL, LDL, oxidized LDL, acetylated LDL, and small unilamellar vesicles[32].

The last mechanism of free cholesterol efflux is mediated by ABCA1. This receptor is a member of a large family of ATP-binding cassette transporters that have common structural motif and use ATP as an energy source to transport a variety of substrate, including ions, lipid, and cytotoxins. In contrast to SR-BI, the preferred cholesterol-acceptors for ABCA1 is a lipid-poor apo-lipoproteins. All of exchangeable apo-lipoproteins, such as apo AI, Apo AII, Apo AIV, Apo E, and Apo C can act as phospholipid and cholesterol acceptors for ABCA1[3334]. ABCA1 promotes the unidirectional efflux of cholesterol and phospholipids to lipid-free or lipid-poor apo-lipoprotein, and is also involved in the lipidation of apo-lipoprotein AI in the formation of nascent HDL. It has been demonstrated by several studies that a direct interaction between cholesterol transporter and apo-lipoprotein play an important role in ABCA1-mediated efflux. ABCA1 has been seen to cross-link with the Apo-AI, indicating a very specific association between them. These results are confirmed by recent studies conducted by Fitzgerald et al., that show as four different mutant ABCA1 transporters with missense mutation exhibited little or no Apo-AI-induced efflux or cross-linking to Apo-AI. This result suggest that a direct interaction between Apo-AI and ABCA1 is necessary for efflux. Plasma lipoprotein can also be removed by less specific mechanisms. Macrophages and endothelial cells possess a scavenger receptor that recognizes modified LDL (i.e. oxidized LDL). The uptake of modified lipoproteins by this scavenger pathway has been hypothesized to contribute to the deposition of cholesterol in atherosclerotic plaques[35].

3.2 *Dyslipidaemia: hypercholesterolemia and hypertriglyceridemia*

The term dyslipidaemia encompasses abnormalities of lipoprotein transport associated with a decrease of lipids in plasma, hypolipidaemia, as well as those causing an excess, hyperlipidaemia. The definition of dyslipidaemia has gradually

evolved as a result of advances in the understanding of the underlying mechanisms. Many monogenically-inherited disorders can now be defined in terms of the specific mutation(s) responsible for encoding the dysfunctional receptor, ligand or enzyme causing dyslipidaemia, whereas most polygenic and secondary forms of dyslipidaemia are still defined by arbitrary cut-offs such as the 5th and 95th percentile of the distribution of the lipid or lipoprotein variable in question.

In a similar manner the classification of dyslipidaemia has evolved from the Fredrickson and World Health Organization classifications of lipoprotein phenotypes devised over 30 years ago to the simpler system now in use. This includes both hypolipidaemia and hyperlipidaemia and differentiates the latter into hypercholesterolaemia, hypertriglyceridaemia and mixed hyperlipidaemia.

Speaking of hyperlipidaemia or dyslipidaemia, we are referring to hypercholesterolemia, hypertriglyceridemia or both. These hyperlipidaemias may be primary or secondary. Secondary hyperlipoproteinaemias are the complication of underlying, metabolic disturbances, drug-induced effects, or the result of dietary excesses. Otherwise, primary hyperlipoproteinaemias are genetically determined and may be classified by lipoprotein phenotype or genotype. The genetic classification may be subdivided into monogenic (caused by single-gene inheritance) and polygenic-multifactorial (caused by multiple subtle genetic factors that act together or in combination with environmental factors)[36].Primary hyperlipidaemia has been classified into six types, based on the specific patterns of the various lipoproteins in plasma. (*Table 2*)

Phenotype	Lipoprotein(s) elevated	Serum cholesterol concentration	Serum triglyceride concentration	Relative frequency, %
I	Chylomicrons	Normal to ↑	↑↑↑↑	<1
IIa	LDL	↑↑	Normal	10
IIb	LDL and VLDL	↑↑	↑↑	40
III	IDL	↑↑	↑↑↑	<1
IV	VLDL	Normal to ↑	↑↑	45
V	VLDL and chylomicrons	↑ to ↑↑	↑↑↑↑	5

Table 2.Primary Hyperlipidaemia classification based on lipoprotein phenotype

However, these types of pattern are not specific, and the plasma lipoprotein pattern may change with the time in any individual. Abundant evidence supports the importance of abnormal lipoprotein metabolism in atherosclerotic disease and more in general in cardiovascular disease. Several epidemiological studies have documented that diets reach of saturated fats and cholesterol accelerate atherogenesis. The prominent alterations that are consistently related to atherogenesis include hypercholesterolemia (reflecting increased concentration of LDL), hypertriglyceridemia (reflecting increased concentration of VLDL and/or remnants and triglyceride enrichment of LDL and HDL), increased apo-B levels, and reduced levels of HDL and its apo-AI[37]. Since different animals are able to form lesion of atherosclerosis similar to human when they develop hypercholesterolemia, it has been possible to understand the cellular changes that come into play in the atherosclerotic plaque formation. Faggiotto et al. focused on this particular issue adding new important informations[38]. These data show that the first and most striking event occurs after 7 to 14 days of diet-induced hypercholesterolemia. The Lipid Research Clinical Trials results it seems quite clear as lowering of plasma LDL levels would be extremely beneficial. Furthermore, it has been seen either in animal models either in human patients aggressively treated with lipid-lowering regimens a regression of atherosclerotic plaque, fatty streaks, significant reduction in the size of the smooth

muscle proliferative lesions, with a consequent benefits in atherosclerosis, myocardial infarction, and coronary artery disease. Brown and colleagues have demonstrated a statistically significant regression of semiocclusive lesions of human coronary atherosclerosis in association with decreases in plasma cholesterol and LDL levels[39].These results provide clear evidence of pathological role of cholesterol and more in general of lipids.

3.3 *Lipid pathway: role of HMGCR and KIF6*

HMGCR (3-Hydroxy-3-Methylglutaryl-CoA Reductase) is a Protein Coding gene. Diseases associated with HMGCR include Alzheimer disease 17 and mevalonicaciduria. Among its related pathways are regulation of cholesterol biosynthesis and metabolism. HMG-CoA reductase is the rate-limiting enzyme for cholesterol synthesis and is regulated via a negative feedback mechanism mediated by sterols and non-sterol metabolites derived from mevalonate, the product of the reaction catalyzed by reductase. Normally in mammalian cells this enzyme is suppressed by cholesterol derived from the internalization and degradation of low density lipoprotein (LDL) via the LDL receptor. Competitive inhibitors of the reductase induce the expression of LDL receptors in the liver, which in turn increases the catabolism of plasma LDL and lowers the plasma concentration of cholesterol, an important determinant of atherosclerosis. Alternatively spliced transcript variants encoding different isoforms have been found for this gene.

Three major superfamilies' of microtubule motor proteins have been identified, kinesin, myosin, and dynein. Kinesin is a wide superfamily constituted by 45 members expressed in mammalian cells, which share analogies in the motor domain but differ considerably in their cargo-binding tail domain[40]. Although KIFs are conserved between species, many different types exist within a single organism, suggesting that KIFs might have diverged through molecular evolution to mediate different cellular functions. Conventional kinesin, also known kinesin-1, and most of

the other vesicles motors in the kinesin superfamily move unidirectionally toward the plus end of the microtubule[41]. Thus, kinesins are likely to be involved in the in trafficking events directed toward the cell periphery, such as motility from the Golgi to the plasma membrane. However, other minus end-directed kinesins also contribute to intracellular trafficking events, such as the minus end-directed transport of early endosomes[42]. The intracellular transport is essential for appropriate cellular morphology and function. Kinesin proteins have been seen to transport organelles, protein complexes, and mRNA to specific site along the microtubule while hydrolysing ATP for energy. Kif6 is a member of the kinesin 9 family but the precise molecular function of kif6 is still not known, and in particular its association with the circulating cholesterol level is not understood. It is likely to play a role in the cellular transport of proteins along microtubules and could be involved in cellular transport in the cardiovascular system. To date, all studies have focused on investigating polymorphism in the KIF6 gene within the context of coronary artery disease even if its molecular pathway needs to be clarified. Indeed, there has been considerable debate regarding the role of KIF6 in coronary artery disease. Several lines of evidence that used a candidate gene based approach in several atherosclerosis population cohorts have observed an increased risk of adverse coronary events in carriers of the rs20455 C variant in the KIF6 gene which lead to a Trp719Arg substitution in the Kif6 protein. 129,130 However, these results were refuted by a large part of Genome-Wide-Association-studies metanalysis using 19 case control studies where cases were defined either by a history of prior myocardial infarction and/or the presence of coronary artery disease at angiography. Furthermore, no association between the KIF6 variant and this pathological condition has been found in the most recent large meta-analysis. Nevertheless, evidence from international randomized controlled trials with statins associated carriage of Tpr719Arg genotype with a greater clinical response to statin therapy. Another study, JUPITER trial, has been realized on a wide population of patients, men and women, without prior cardiovascular disease or diabetes, were randomly allocated to rosuvastatin 20

mg/day or to placebo and followed for first major vascular events and for all causes of mortality. The effects of Tpr719Arg polymorphism have been evaluated. Outcomes obtained from this study didn't show any correlation between genotype and major cardiovascular risk factors, as well as no clinically meaningful differences in pharmacological response have been seen between carriers and non-carriers.

3.4 *Pharmacology of lipid metabolism*

Lipid regulating drug therapy is indicated in patients with coronary heart disease (CHD) and high-risk individuals in whom dietary and lifestyle measures have failed to control dyslipidaemia. Because this is usually a life-long commitment, the use of these drugs should be the exception rather than the rule in asymptomatic patients, being restricted to those with severe hyperlipidaemia of genetic origin, such as familial hypercholesterolaemia (FH), or in whom the presence of other risk factors results in an unacceptably high risk. The opposite applies, however, to patients with clinically manifest CHD, in whom even mild dyslipidaemia requires vigorous drug therapy aimed at achieving target levels stipulated in guidelines.

The rationale for the use of lipid regulating drugs is based on the large body of evidence from epidemiological and clinicpathological studies that points to the central role of cholesterol in atherosclerosis. Confirmation that the association is causal comes from numerous angiographic and clinical end- point studies showing that lipid-lowering therapy slows the rate of progression of atherosclerotic lesions in coronary, carotid and femoral arteries, and reduces the frequency of associated cardiovascular events. The best proof of causality derives from the atherogenicity of low density lipoprotein (LDL) and the evidence that lowering LDL cholesterol arrests or reverses the process. Loss of the protective effect of high density lipoprotein (HDL), resulting from a decrease in plasma levels, also has strong epidemiological support, but the evidence that raising HDL cholesterol is beneficial is less compelling. Likewise, although increasing epidemiological and angiographic

evidence suggests that triglyceride-rich remnant particles play a role in promoting the progression of mild-to-moderate lesions in coronary arteries, more data are needed on whether lowering triglyceride prevents CHD events. Hence, it is not surprising that the current therapeutic emphasis is on lowering LDL cholesterol. That said, there is increasing interest in developing compounds that raise HDL cholesterol and could be used as an adjunct to LDL-lowering drugs.

As it has been mentioned above, several lines of evidence from epidemiological and biochemical studies have demonstrated the importance of reducing to physiological level lipids concentration and in particular total cholesterol and LDL concentration. Indeed, it has been well demonstrated over the years that a linear relationship between LDL-C levels and relative risk for CHD exists, as showed in

Figure 2 below:

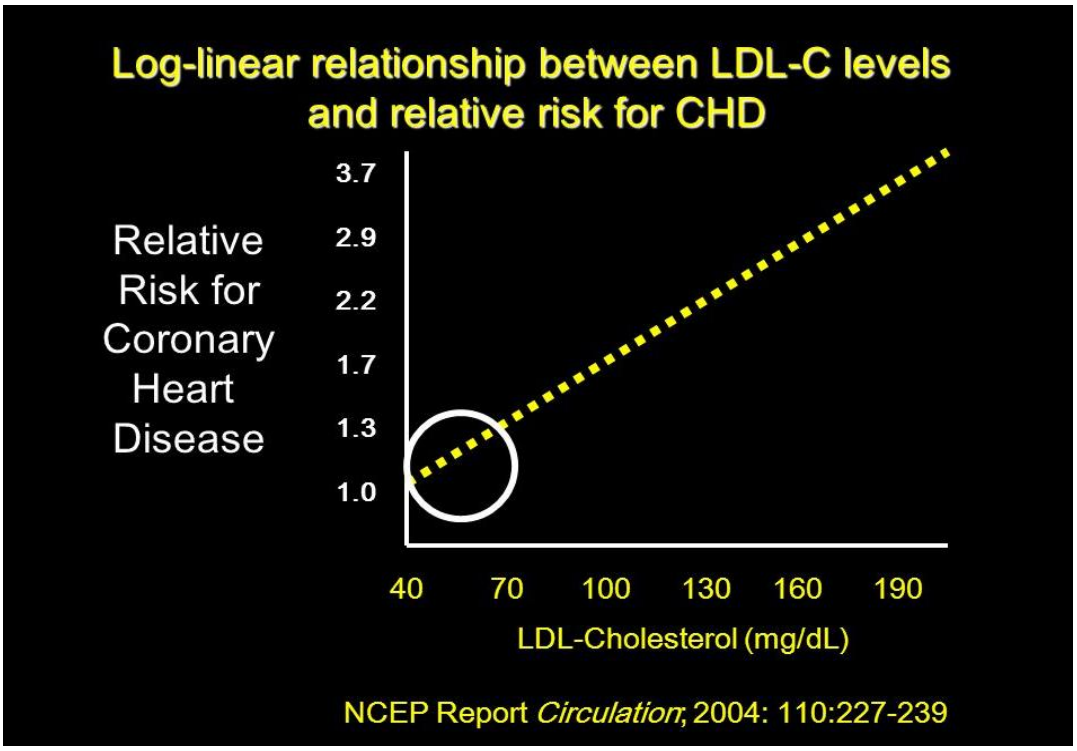

Figure 2 .Linear correlation existing between LDL-C and Relative Risk for CHD demonstrates the high importance to keep total cholesterol level and LDL within a physiological range suggested from the international guidelines for cholesterol level (Grundy et al. Circulation2004; 110(2):227-239).

According to the World Health Organization (WHO), LDL cholesterol level ranging from 100-130 is considered optimal, >160 high, >190 very high; Total Cholesterol is considered desired < 200, borderline between 200-239, and high ≥240; HDL cholesterol <40 is believed low, ≥60 high (*Table 3*).

LDL Cholesterol level (mg/dl)	*Acceptance Degree*
<100	**Optimal**
100-129	**Near optimal/above optimal**
130-159	**Borderline high**
160-189	**High**
≥190	**Very high**
Total Cholesterol	
<200	**Desirable**
200-239	**Borderline high**
≥240	**High**
HDL Cholesterol	
<40	**Low**
≥60	**High**

Table 3.Classification of lipids level in the blood and acceptance risk according to the international guidelines.

Risk factors such as diet, smoking, physical inactivity, obesity, high blood pressure, lipids (high LDL and low HDL cholesterol and raised triglycerides), diabetes, family history of coronary heart disease are responsible of a huge number of cases of coronary heart disease. The World Health Organization (WHO) believes that 60% of coronary heart disease and 40% of strokes are due to elevated LDL cholesterol levels. Reducing LDL-C has long been the primary target of cholesterol policy and this remains the case today. The two major approaches of LDL-lowering therapy are

therapeutic lifestyle changes (TLC) and drug therapy. Dietary therapy should be the first approach in all form of lipid disorders. Although there is a great variation in individual response to a low-fat, low-cholesterol diet, changes of alimentary habits can be expected to play a significant role in improving most lipid profile. The minimal goals of dietary therapy should be to reduce LDL cholesterol to under 160 mg/dl and lower the total cholesterol to under 240 mg/dl in patients without CAD or two additional risk factors and to less than 130 mg/dl if any of these conditions apply. Drug treatment of lipid disorders should not be used until a precise diagnosis. Before to start the drug lipid-lowering therapy, maximum dietary effort should be done. To date, there are several classes of lipid-lowering drugs which act at a different levels of cholesterol biosynthesis, such as bile acid sequestrants (Cholestyramine, Cholestipol), Nicotinic acid, Fibric acid derivatives (Gemfibrozil, Clofibrate), HMGCoA-reductase inhibitors (Statins).(*Table 3*) Cholestyramine and Cholestipol are quaternary ammonium salts that act as anion-exchange resins, binding bile salts in the intestine tract. This sequestration leads to a reduction of reabsorption of bile salts and consequently to a decrease of their enterohepatic recycling; in the same time there is an increase of bile salts excretion in the stool. The positive effects on lipid-lowering of this class of drugs is represented by the fact that cholesterol will be used for an ex-novo synthesis of bile salts rather than be used for LDL or VLDL synthesis. Cholestyramine and Cholestipol are considered first line agents and are highly effective in decreasing LDL cholesterol[43]. The niacin is able to reduce the LDL levels by about 25% and VLDL levels by about 75%.Furthermore, this drug is able to increase the HDL levels by 20 to 40%. Similarly to bile acid resins, nicotinic acid's activity, is accompanied by frequent side effects, that include cutaneous flushing and gastrointestinal (GI) symptoms. Gemfibrozil and clofibrate are the fibric-acid derivatives widely used in USA; in Europe, fenofibrate, benzafibrate, and ciprofibrate are also used. The fibric-acid derivatives are used in VLDL and triglycerides lowering. Their mechanism of action is not completely clarified, even if they seem to act by increasing lipoprotein lipase activity via peroxisome proliferator activated

receptors-α; this action leads to an increase of VLDL clearance[44].The last class, not for importance, of lipid-lowering drug is represented by 3-hydroxyl-3- methyl-glutaryl-CoA-reductase inhibitors, the so-called statins: atorvastatin, pravastatin, lovastatin, rosuvastatin, and simvastatin. This class of drugs is the most extensively used in the treatment of hypercholesterolemia. The efficacy, safety, and benefits of statins make them the drugs of choice. Statin are able to lower the serum low-density lipoprotein and triglycerides levels, and in the same time they increase the HDL levels. Furthermore, it has been seen that statins offer an important contribution to the reduction of the risk for the major cardiovascular events, such as MI and death for CVD in selected population. Side effect of statins are principally muscle-related (such as muscle weakness) and in general mild in nature. More serious side effects, such as myopathy and fatal rhabdomyolysis are rare and often can be seen when statins are used in combination with other medicines that affect statin pharmacokinetics[45]. Usually statins are used singularly, even if there are several cases in which statins are used in combination with other lipid-lowering drugs. However, doubts about the benefits of statins, alone or in combination with other lipid-lowering drugs, remain. In patients with a very high LDL levels or mixed dyslipidaemia that fail to reach the desired lipids levels on statin monotherapy, other classes of drugs are often added in order to improve clinical outcomes. Besides the co-administration, another aspect of statin therapy that remains difficult for prescribers is the ability to predict individual patient's response to therapy, in terms of efficacy and genetic predisposition to possible adverse effects[46]. Plasma concentration of statins as well as the pharmacological efficacy can vary widely amongst patients that receive the same dose of the same statin. Recent results obtained from pharmacogenomics and pharmacogenetics studies have demonstrated a wide number of candidate genes (40 gene) that are involved in the modulation of statin pharmacokinetics and pharmacodynamics. Doubtless, the most important is the gene expressing for the HMGCoA-reductase enzyme. Statins act binding to the catalytic domain of HMGCoA-reductase enzyme that, as above said, catalyses the rate-limiting step of

cholesterol biosynthesis. All statins share rigid, hydrophobic groups that are covalently linked to the HMG-like moiety. All statins are competitive inhibitor of the natural substrate HMGR, and it has been seen that they act occupying the HMG-binding pocket and part of the binding surface for CoA, thus sterically preventing substrate from binding. Comparing the complex obtained between the natural substrate and the enzyme or between a statin and the enzyme, a clear rearrangement of the substrate-binding pocket has been seen after the statin interaction. The strong efficacy of statin's interaction with the catalytic site of HMGCoA-reductase is probably due to the large number of van der Waals interaction, which in turn reflects the importance of each amino acid that takes part in the catalytic site's constitution. Several studies have shown as structural changes of catalytic site of HMGCoA-reductase, principally caused by genetic alteration, lead to a loss of statin treatment efficacy. (*Table 3)*

Daily dose	**Mean change (%)**		
	LDL-C	HDL-C	TG
Atorvastatin 40 mg	-51	+5	-32
Nicotinic acid 4 g	-9	+43	-34
Gemfibrozil 1.2 g	-18	+12	-40
Ezetimibe 10 mg	-18.5	+3.5	-4.9
Colestyramine 24 g	-23	+8	+11
TC, total cholesterol; LDL-C, low density lipoprotein cholesterol; HDL-C, high densitylipoprotein cholesterol; TG, triglyceride.			

Table 4. Comparative effects of lipid regulating drugs

3.5 *Modulation of statin pharmacokinetics and pharmacodynamics and individualized therapy*

Individual variability in drug efficacy and drug safety is one of the most important challenge in current clinical practice of this century and the personalized therapy seems not to be a dream anymore. In a large patient population, a medication that is demonstrated efficacious and safe in many patients, it is proven to be inefficacious and sometimes lethal in some other peoples. Although large individual variability in drug efficacy and safety has been well documented amongst peoples since the beginning of the medicine, its understanding still represents an unmet need. The demonstration of individual variation in drug response has proven difficult, even if the demand to overcome such variation has received more attention and become the goal of modern medicine. It is well known that large variability of drug efficacy and adverse drug responses in patients is a major determinant of the clinical use, regulation, and withdrawal from the market of clinical drugs. Drug response is influenced by several factors, some of which involve important aspects that are intrinsic to the biology itself. Genetic variation in humans was recognized as the most important determinant of individual variability of drug response from clinical observations in late 1950s[47],[48]. From these studies it has been observed amongst a certain population differences in plasma concentration of specific drugs and in the similar way also differences of metabolites concentration in urine, demonstrating that variances in pharmacodynamics and pharmacokinetics are present and exert a very important role. It has been seen that in some patients a very low or a very high concentration could be present and the biochemical traits leading to the variation of drug concentration were found inherited. These clinical findings fostered the formation of pharmacogenetics. Differences of drug response and behaviour reflect sequence variations of specific genes coding for drug targets, such as proteins, drug metabolizing enzymes, receptors, drug-transporters[49],[50]. Starting from these outcomes, the availability of the complete human genome sequence has made it

possible to analyse the impact of variations of human genome sequence on the pathogenesis of important diseases and the response to drug therapy. The amount of knowledges and informations acquired over the years on genome-disease and genome-drug interaction has brought to a new area so-called pharmacogenomics starting from the previous pharmacogenetics. This goal provided a rationale for the hope that the personalized medicine can be achieved in the near future. Both pharmacogenomics and individualized drug therapy are increasingly influencing medicine, biomedical research, in many areas, including clinical medicine, drug development, drug regulation, pharmacology, toxicology. As already said drug efficacy and adverse drug reactions dose-dependently determine the clinical outcome of clinical therapy. To date, it is well known as these parameters change widely over the cohort and the personalized therapy, based firstly on a genetic screening and then on using of the right drug and concentration, may represent the only possible way to avoid to use an inefficacious drug and increase adverse drug reactions. We know that a higher dose boosts drug therapeutic effect but in the same time it increases the possibility to have undesirable side effects. The difference between the drug therapeutic effect and its side effects defines the so-called therapeutic window (*Figure 3*)

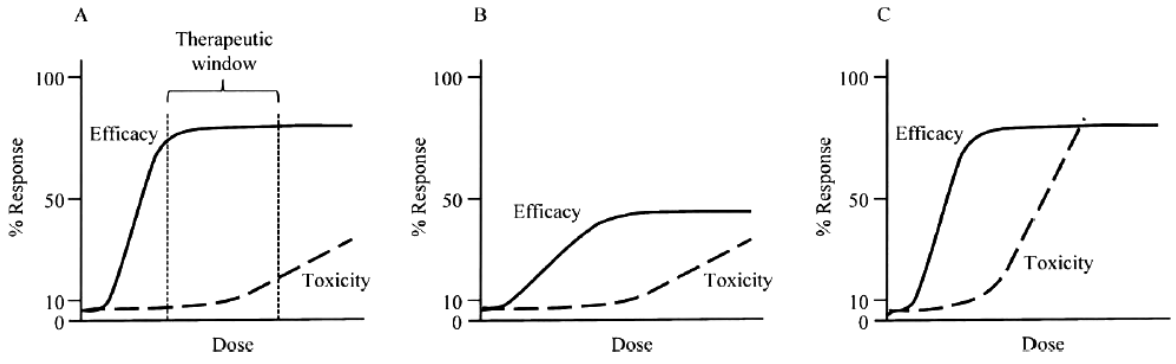

Figure 3.Representation of dose-response curve. Panel A shows the therapeutic window, that is the difference between the efficacy and the toxicity. This window may vary from patient to patient, B and C, requiring dose-adjustment (Qiang Ma and Anthony Y. H. Lu, Pharmacogenetics, Pharmacogenomics, and Individualized Medicine, Pharmacol Rev 63:437–459, 2011).

For many drugs, the optimum dose required for effective and safe therapy varies significantly from patient to patient, because the minimum therapeutic dose can be too low or too high. Such a situation can be characterized by an atypical therapeutic window, dose response curve, and toxicity resulting in unexpected and undesirable outcomes. Genetic and non genetic factors affect individual variability of a drug response by modulating the dose response curves of drug efficacy and drug toxicity of patients. Clinical outcome is altered if drug dose is not adjusted accordingly.

3.6 *Genetic variation: SNPs*

Inherited differences in DNA sequence contribute to phenotypic variation, influencing an individual's anthropometric characteristics, risk of disease, and response to the environment. A central goal of genetics is to localize the DNA variants that contribute most significantly to population variation in each trait[51]. The idea that genes control some drug responses was suggested for the first time in the 1950 thanks to a correlation between inheritance or ethnicity and aberrant drug responses. This idea was strengthened by family and twin studies in 1960s 1970s, extended by biochemical studies in 1980s, and solidified by molecular genetics in the 1990s[52,53,54]. Cloning and characterization of the first human gene containing DNA sequence variations, singular nucleotide polymorphisms, that influence drug pharmacodynamics and pharmacokinetics did not take place until the end of 1980s[55].Although there are a number of different types of polymorphic markers, most attention recently has focused on single nucleotide polymorphism (SNPs) and the potential for using these to determine the individual drug response profile. Polymorphisms are genetic variations of human genome that occur at a frequency of 1% or greater in the population. These variations can appear as insertion or deletion of short fragments of nucleotide, even if the most common variations are SNPs. A SNP is a DNA sequence variation characterized by a singular nucleotide substitution among members of the same species or between chromosomes in an individual. SNPs

are the most abundant genetic variants in mammalian genomes and specifically in the human genome they account for 3.2 million and are responsible for the 90% of genetic human differences[56]. SNPs are classified in three different groups depending on where they are located in the genome: 1. c-SNPs, variations located in coding region, exons, whose presence could modify or not the amino acid sequence in the protein structure, defined as synonymous and non synonymous respectively 2. p-SNPs are located in perigenic region, i.e., in regulatory regions such as promoter, enhancer, etc… Such polymorphism can modulate both the expression level and the stability of the protein, by influencing the mRNA stability. 3. r-SNPs, random SNPs that are located in the intergenic region, i.e., in those regions that do not show genes but that constitute the 98% of our genome. These variations are not able to influence the genetic expression. However they may alter the DNA structure by influencing the tertiary structure of DNA, and its capacity of interaction with chromatin or with enzymes such as topoisomerase. For these reasons, to date, we cannot exclude their implication in the pharmacogenetics. Currently there is a growing list of polymorphisms found in genes encoding of drug transporters and targets, drug-metabolizing enzymes, as well as disease-modifying genes, that have been linked to drug effects in humans. Pharmacogenetics has indeed proven to be a potential source of biomarkers able to predict drug response and adverse drug reaction. For this reason, during the last decade a rapid development of techniques in the area of genome analysis has been seen and recognized as needful for the identification of new pharmacogenomics biomarkers. Such biomarkers mainly originate from genes encoding drug metabolizing enzymes, drug transporters, and drug targets. Some of these are now integrated by the USA Food and Drug Administration (FDA) and the European Medicines Agency (EMA) into drug label inserts[57].Although the numerous potential benefits of pharmagenomics and pharmacogenetics have been announced and sometimes also demonstrated, to date, this innovative area of research is so rarely used in clinical practice. This seems to be a particularly relevant question that in the next year needs to be answered. It seems that the failure to use of this type of

approach in the clinical practice is related not only to scientific issues but also to cultural limits. For example, individualizing dosages, even based on easily assessed patients characteristics, such as sex, age, renal functionality, has not embraced by medical or pharmaceutical communities. Furthermore, there is a strong resistance to relying on tests for every medical decision. Not only does pharmacogenomics require a laboratory test, it also requires an interpretation of genotypes, which will probably require clinicians to receive further training in molecular biology or genetics. Lastly, it also difficult to demonstrate that individualization of drug therapy on the basis of genetics improves clinical outcomes, given the multigenic nature of most drug effects. These represent just some of limits that can hinder the clinical use of this discipline.

4 Role of *KIF6* and *HMGCR* polymorphisms in variability of cardiometabolic phenotypes and response to statin therapy in the Brisighella cohort

4.1 *Introduction*

Cardiovascular diseases (CVD) encompass a wide range of pathological conditions, including raised blood pressure, hypertension, ischaemic and coronary heart disease, cerebrovascular disease, heart failure[58]. To date, CVDs represent the number one leading cause of mortality and morbidity with an estimated number of 17 million of deaths/year in the world, projected to reach 23.3 million by 2030[59,60]. Several lines of evidence from epidemiological and biochemical studies have widely established that high level of low-density lipoprotein cholesterol (LDL-C), low concentration of high-density lipoprotein cholesterol (HDL-C), and high level of total cholesterol (TC) play a pivotal role for the development of CVD, and moreover can be considered as important factors in the neurodegenerative pathology. In particular, high LDL-C is causally associated with coronary artery disease (CAD).Triglyceride levels have been associated with coronary artery disease (CAD)[61]. The diagnostic and predictive value of LDL-fractions as apolipoprotein B (APOB) has so far only been investigated in smaller studies.

To date, amongst the great body of lipid-lowering drugs, statins are the most used worldwide and represent the cornerstone of lipid lowering strategy. However, reduction of statins prices, related to wide availability of generic drugs, further adds to an increase in their prescription specially in subjects at lower estimated CVD risk[62]. Despite large randomized controlled trials have well-documented the effectiveness of statins in reducing circulating LDL-C levels and CVD events, several issues have been noted in their use. In particular, response to the same dosage of similar types of statins varies among patients; in addition, to date, there are no established criteria to prevent their adverse effects, such as myalgia and

rhabdomyolysis. Genetic factors are thought to come into play in the inter-individual variations of response to statins[63,64]. Thus, elucidating genetic variation in patients undergoing statin therapy may enable tailored treatment, thereby optimizing efficacy as well as minimizing costs and adverse effects. Over the past decade, more than 40 loci in human genome have been described with respect to the differential effect of statins on decreasing the risk of clinical endpoints including cardiovascular death and myocardial infarction (MI), and on modulating lipid levels[59,60,65,66].

Amongst these, common variants at Kinesin Like-Protein 6 (*KIF6*) gene have seen the most controversial observations. KIF6 is a homodimeric molecule belonging to the KIF9 family of kinesins, a big superfamily of motor proteins that are involved in variety of processes in the eukaryotic cells, such as cell division, cytoplasm organization and intracellular microtubule transportation of protein complexes, organelles and mRNA, *KIF6* is a relatively recent candidate for CVD, initially suggested as contributing to cardiovascular risk within the European population[67]. Specifically, a common non synonymous SNP, rs20455 resulting in a missense Tryptophan to Arginine substitution at position 719 of the codified polypeptide chain (p.W719R, Trp719Arg) has been widely studied since 2007[68,69]. In a meta-analysis of seven prospective studies it has been demonstrated that heterozygous and homozygous carriers of the *KIF6* 719Arg variant were at an increased risk to develop CVD and received a significant benefit from the statin therapy compared to 719Trp homozygous wild type individuals[70]. However, contradictory results have been reported by subsequent studies for this variant role in statin therapy response[59,71,72,73]. Moreover, variants within *KIF6* locus have never been reported in association with lipids levels in large genome-wide association studies (GWAS), therefore its role in regulating statin therapy response is not defined.

An important and well established factor in statin-responsive lipid and lipoprotein concentrations is the 3-Hydroxy-3-Methylglutaryl-Coenzyme A reductase (HMG-CoA reductase)[74]. *KIF6* and *HMGCR* are potentially both involved in the pathway of lipid-lowering statin drugs. Genetic effects of HMG-CoA reductase (*HMGCR*) and

KIF6 genetic polymorphisms have been extensively studied individually, however no studies have investigated their combined effects on bloodstream lipid levels modulation.

In this study, we aimed to investigate the role of *KIF6* variants in the body metabolism through analysis of association with cardiometabolic phenotypes, including14 quantitative traits and 5 categorical disease outcomes, including categorical response to statin therapy, in the GBHS and EGCUT population-based samples of up to 12.307 individuals. We used established lipid locus *HMGCR* variant as proof-of-principle analysis for this first genetic study from GBHS.

4.2 Materials and Methods

4.2.1 Study populations

The GBHS is a prospective, population-based longitudinal epidemiological investigation of genetic factors influencing a range of cardiometabolic phenotypes involving 2,939 randomly selected Caucasian subjects (1,491 men and 1,448 women), aged 14 to 84 years, free of cardiovascular disease at enrolment, resident in the northern Italian rural town of Brisighella. The study started in 1972 and it is still ongoing. The town of Brisighella was originally selected as the site for the study, given a homogeneous life-style of its residents, with a very low rate of migration. Subjects were clinically evaluated at baseline and every four years thereafter by collecting an extensive amount of clinical and laboratory data. The GBHS protocol and its sub-studies, largely described elsewhere[75],[76]have been approved by the Ethical Board of the University of Bologna and all volunteers involved gave their signed consent to participate in the study.

For the purpose of this study, we randomly selected 1,645 subjects from the 2008/2012 GBHS population survey, with available blood samples.

The Estonian Genome Center University of Tartu (EGCUT) is a prospective, volunteer-based sample of the Estonian resident adult population (aged ≥ 18 years). The current number of participants – close to 52,000 – represents a large proportion, 5% of the Estonian adult population, making it ideally suited to population-based studies. General practitioners (GP) and medical personnel in the special recruitment offices have recruited participants throughout the country. At baseline, the GPs performed the standardized health examination of the participants, who also donated blood samples for DNA, white blood cells and plasma tests and filled out a 16 – module questionnaire on health-related topics such as lifestyle, diet and clinical diagnoses described in WHO ICD-10. Two independent samples of 2,589 and 8,073 individuals from the EGCUT genotyped on Illumina370CNV and Illumina OmniExpress arrays, respectively and imputed to 2,5 Mio SNPs from HapMap were used for replication of results in follow-up.

4.2.2 *DNA extraction , Genetic analysis and quality control*

GeneCatcher™ gDNA Blood Kits (Invitrogen) The GeneCatcher™ Technology is a novel magnetic bead-based technology that is designed to work on a wide range of blood samples including archived or poorly stored blood samples to facilitate genomic DNA purification.

The procedure is designed for isolating gDNA using theGeneCatcher™ Magnetic Beads procedure. *(Figure 4*)

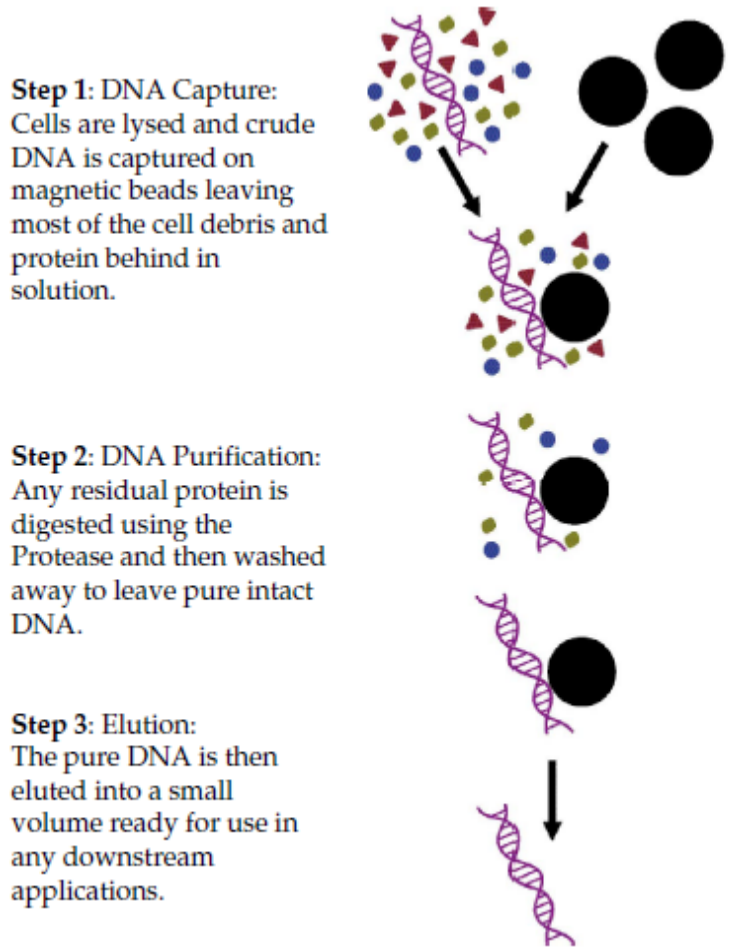

Figure 4 .Procedure designed for isolating gDNA using theGeneCatcher™ Magnetic Beads procedure.

The DNA quantification was done with the Quant-iT ™ dsDNA Broad-Range Assay Kit (Invitrogen) with a fluorescence microplate reader (Victor PerkinElmer) equipped with excitation and emission filters 510/527 nm appropriate for fluorescein. For genotypes analysis, we genotyped five selected SNPs: three in the *KIF6* gene on chromosome 6, one non synonymous missense (rs20455) and two intronic (rs9462535, rs9471077, linkage disequilibrium from 1000 Genomes CEU database between the three SNPs $0.8<r^2<1$), and two in *HMGR* gene on chromosome 5, one upstream (rs3761740) and another intron 12(rs3846662, LD $r^2=0.215$).

Genotyping was performed by real-time using the 5′-nuclease allelic discrimination assay (TaqMan®, Applied Biosystems, Foster City, CA), according to manufacturer instruction. Negative controls were included in each reaction as quality control.
All five genotyped SNPs satisfied our quality control criteria with a call rate ≥95% and a Hardy-Weinberg equilibrium test p-value $P>1x10^{-4}$. All minor allele frequencies (MAF) were >0.01(*Table 6*). We estimated linkage disequilibrium (LD) between genotyped SNPs (rs20455, rs9462535 and rs9471077 for *KIF6*, rs3761740 and rs3846662 for *HMGCR*) in our sample using Haploview software and compared it to LD reported for 1000Genome CEU data .

4.2.3 *Phenotypes*

We analysed a total of 19 cardiometabolic phenotypes. Of these 14 were quantitative traits and 5 were binary outcomes including type 2 diabetes (T2D), coronary artery disease (CAD), hypertension (HTN), hypercholesterolemia (HTC) and categorical response to lipid-lowering therapy. Diseases were defined following the standard international guidelines[77,78,79].Response to lipid-lowering therapy was defined only for patients that underwent lipid-lowering therapy, most of whom used statins at the time of sampling. These subjects were categorised as “responder” and “non-responder” to the therapy based on their value of for this study, we defined statin-resistant those subjects regularly taking a statin and who experienced an LDL-C reduction of no more than 30% of the expected reducing efficacy of that drug at the given doses.
Quantitative phenotypes included total cholesterol (TC), high-density lipoprotein cholesterol (HDL-C), low-density lipoprotein cholesterol (LDL-C), triglycerides (TG), apoliprotein A1 (APOA1), apolipoprotein B (APOB), fasting glucose (FG), body mass index (BMI), waist circumference (WC), wrist circumference (WrC), height, systolic and diastolic blood pressure (SBP, DBP), heart rate (HR), serum uric acid (UR), creatinine (CREA).Following indication reported elsewhere[80,81,] for lipids

we excluded individuals undergoing lowering-lipids therapy while for those individuals undergoing anti-hypertension therapy we added 15 mm Hg to SBP and 10 mm Hg to DBP.

To check the assumption of normally distributed quantitative traits, we performed a Shapiro-Wilk test on all such variables. Since normality test significantly deviated from normal distribution (p-value<0.05) for all phenotypes we applied an inverse normal transformation of residuals on all quantitative phenotypical variables. Residuals were calculated using linear regression analysis adjusted for specific covariates: all lipids and BMI were adjusted for sex, age and age^2; FG, CREA, UR and height were adjusted for sex and age; anthropometric traits, DBP, SBP and HR were adjusted for sex, age, age^{2}and BMI. Normality test and transformation of data were run using R software[82].

4.2.4 *Association analysis*

For quantitative traits we assumed an additive genetic model of effects and applied a linear regression to transformed residuals after covariate adjustment. For categorical phenotypes we assumed log-additive genetic effects and implemented logistic regression. We adjusted the trait models for covariates as follows: T2D was adjusted for sex, age and BMI, CAD was adjusted for sex and age, HTC for sex, age and age^2 and HTN for sex, age, age^2 and BMI.

All single variant association analyses were run using R and PLINK software in parallel[83,84].Association analysis in two Estonian samples was run using SNPTEST software using maximum likelihood genotype estimates on imputed SNP array data[85]. Meta-analysis of results from three studies was performed using GWAMA software[86].

Association tests were Bonferroni corrected for multiple testing with 42 independent tests accounting for 14 non-highly correlated phenotypes ($r<|0.5|$, $r^2<0.25$) and 3 independent SNPs (as the three genotyped *KIF6* variants are highly correlated

($0.8<r^2<1$) and were counted as one test) resulting in a study wide significance threshold of $P<0,05/42$ i.e. $P<1,2x10^{-3}$.

Gene-gene interaction between *KIF6* and *HMGCR* loci variants in their effects on cardiometabolic phenotypes was tested using regression analysis with an interaction term to estimate effect between each pair of tested SNPs.

Haplotype estimation was performed for KIF6 SNPs ($0.8<r^2<1$) using PLINK software, followed by association analysis of haplotypes and diplotypes (i.e. haplotype genotypes) with cardiometabolic phenotypes.

4.2.5 *Phenotype correlation and multi-phenotype analysis*

We evaluated the correlation between analysed phenotype using the *polycor* package in R. *hetcor* function allow to perform a Pearson product-moment correlations test between quantitative variables, a polyserial correlation test between quantitative-discrete variables and a polychoric correlation test between discrete-discrete variables. For those phenotypes that resulted highly correlated ($|r|>0.5$), we selected only the phenotype with less missing data to be included in the multi-phenotype analyses.

To check for pleiotropic effects of analysed variants on multiple cardio-metabolic phenotypes we run ordinal regression analysis using an implementation of the Multiphen algorithm[87] in C++ code. In Multiphen the genotype of a marker is used as outcome variable, and the set of multiple phenotypes as predictors: for joint analysis of K phenotypes, we modelled the genotype, G_{ij}, of the ith individual, at the jth variant, coded as 0, 1 or 2, according to the number of minor alleles it carries, as a linear function of phenotype values, yi, in a logistic regression framework. Specifically,

$$g^{-1}(G_{ij}) = \alpha_j + \beta_j y_i,$$

Where g^{-1} was the logit link function, $\alpha_j$was the intercept, and β_j was a vector of phenotype regression coefficients for the jth variant. Under this model, we obtained

maximum-likelihood estimates (and standard errors) of the phenotype regression coefficients and the corresponding deviance D_j defined as:

$$D_j = 2 \times [\![(l]\!]_j - l_0)$$

with an approximate chi-squared distribution with *n* degrees of freedom, where l_j is the log-likelihood of the *j*th logistic regression model and l_0 is the log-likelihood for the null model.

The Bayesian Information Criterion (BIC) score was selected as the optimal model fit statistic. BIC is defined as:

$$BIC_j = 2l_j + (s_j + 1) \times \log(n)$$

Where l_j is the log-likelihood of the *j*th logistic regression model, s_j is the number of phenotypes in the model and *n* is the sample size (note that for a null model with intercept only, $BIC_0 = -2l_0 + log\ (n)$, where l_0 is the log-likelihood for the null model).

We run a logistic regression analysis to test for association with *KIF6* polymorphisms, correcting for age and BMI. As positive control, we tested also *HMGCR* polymorphisms.

4.3 *Results*

In the GBHS all the three variants in the *KIF6* gene are in high LD (*Figure* 5). This is consistent with data reported for European samples from the 1000 Genome (1000G) data with $r^2_{rs20455\text{-}rs9462535}$= 0,86 (0.80 in 1000G), $r^2_{rs20455\text{-}rs9471077}$= 0,85 (0.80 in 1000G) and $r^2_{rs9471077\text{-}rs9462535}$=0,96 (1,00 in 1000G). The two variants in the *HMGCR* gene, instead, have a smaller correlation (LD) (*Figure 6*), lower than that reported in the1000 Genome data with $r^2_{rs3761740\text{-}rs3846662}$= 0,12 (0.22 in 1000G).

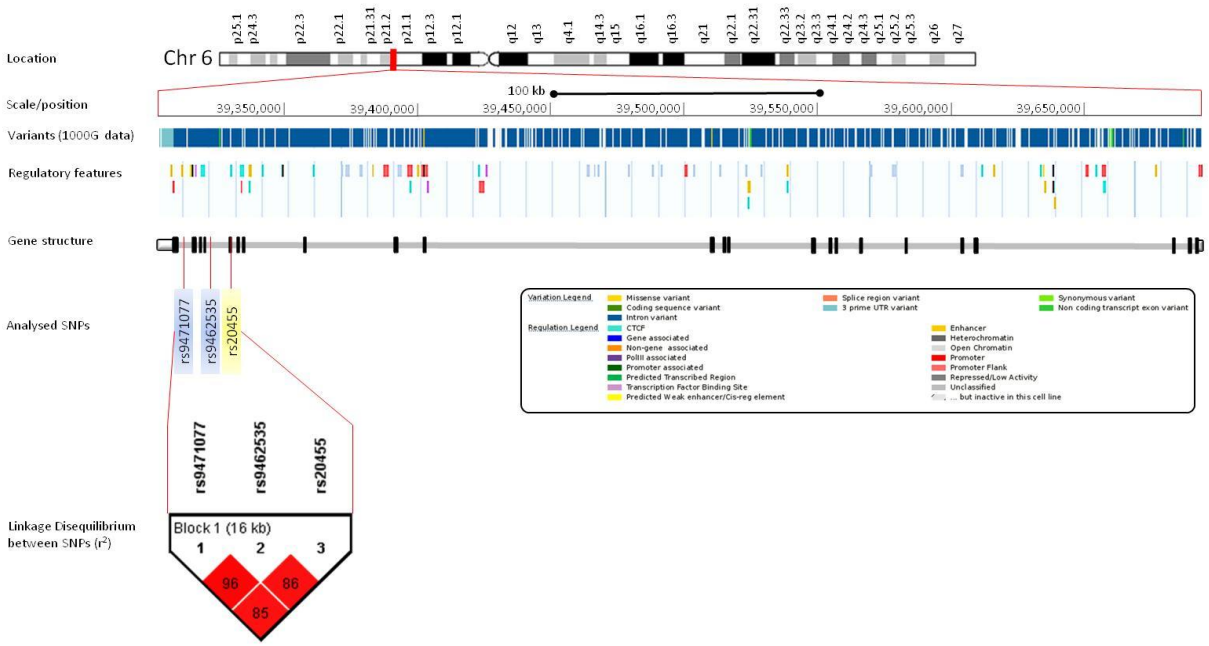

*Figure 5.**Locus view of the three investigated variants in the KIF6 gene**. The view includes the gene structure of KIF6 and the linkage disequilibrium (r^2) between the three SNPs*

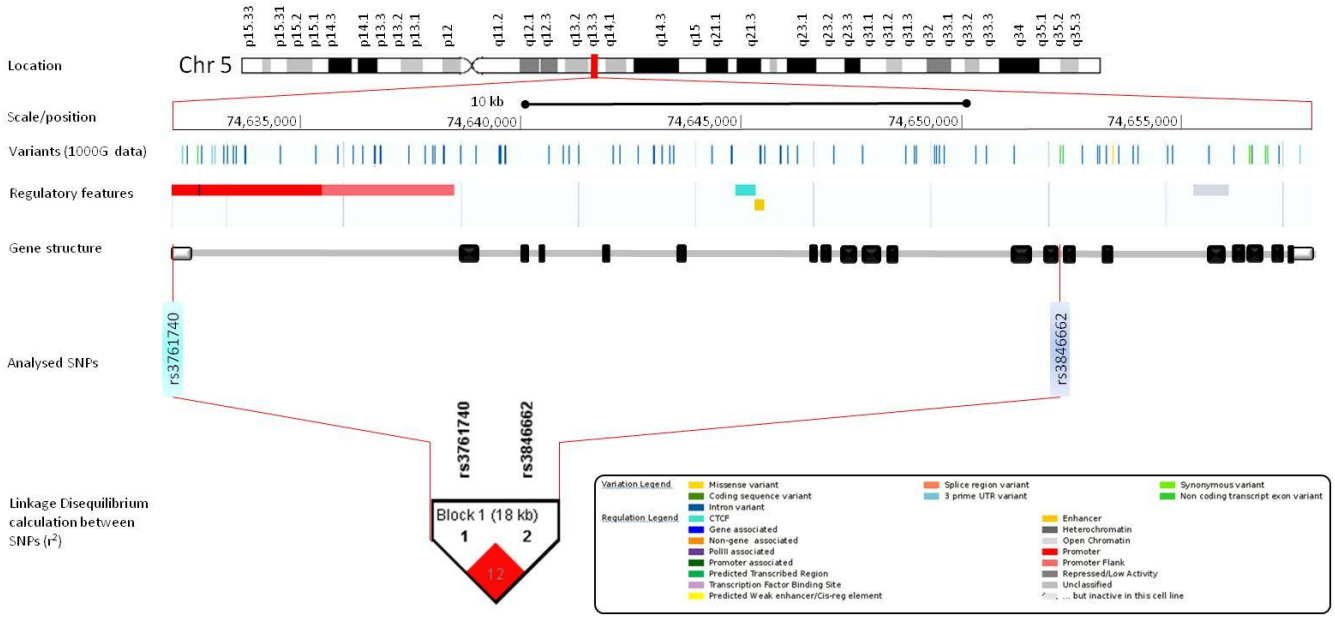

*Figure 6. **Locus view of the two investigated variants in the HMGCR gene**. The view includes the gene structure of HMGCR and the linkage disequilibrium (r^2) between the two SNPs*

lipid-lowering therapy (i.e. the categorical phenotype "statin response" as defined in materials in methods) in 230 subjects from GBHS, including 134 responders and 94 non-responders (*Table* 54).

In the GBHS we replicated the association at *HMGCR* locus rs3846662 with LDL-C levels already reported in literature [88] (P=0.00085, and after Bonferroni correction for 42 multiple tests (14 traits * 3 independent SNPs) $P_{corrected}$=0.0357) (*Table 5*).

Interestingly, *HMGCR* rs3761740 variant showed a significant association with SBP and DBP (p-value$_c$= 0.01285 and 0.0357, respectively) but the association was not significant after Bonferroni correction for 42 multiple tests.

			KIF6			***HMGCR***	
	SNP	rs20455	rs9462535	rs9471077	haplotype	rs3761740	rs3846662
	risk allele	T	C	A	ACT	C	G
	other allele	C	A	G	GAT/GAC	A	A
	risk allele freq.	0.635	0.619	0.616	0.606	0.904	0.457
TC	beta	0.05365	0.06225	0.08123	0.0855	0.06065	0.1131
	se	0.04018	0.04033	0.04022	-	0.06555	0.03735
	pval	0.182	0.123	**0.0436**	**0.0319**	0.355	**0.00251**
TG	beta	0.09048	0.12101	0.11634	0.113	0.06269	-0.02634
	se	0.04014	0.04024	0.04015	-	0.06555	0.03743
	pval	**0.0243**	**0.00268**	**0.00382**	**0.00459**	0.339	0.482
LDL-C	beta	0.08837	0.08154	0.10358	0.106	0.035838	0.1254
	se	0.04035	0.04053	0.04038	-	0.066263	0.0375
	pval	**0.0287**	**0.0444**	**0.0104**	**0.0079**	0.589	***0.00085***
HDL-C	beta	-0.12974	-0.10938	-0.1051	-0.101	-0.0051489	0.01972
	se	0.04008	0.04028	0.0402	-	0.0655845	0.0375
	pval	**0.00124**	**0.0067**	**0.00903**	**0.011**	0.937	0.599
APOA1	beta	-0.10598	-0.09433	-0.09094	-0.0863	0.037209	-0.007389
	se	0.04123	0.04151	0.0414	-	0.066898	0.038384
	pval	**0.0103**	**0.0232**	**0.0282**	**0.0357**	0.578	0.847
APOB	beta	0.1091	0.12246	0.13144	0.13	0.013194	0.06559
	se	0.04085	0.04096	0.04079	-	0.066228	0.03797
	pval	**0.00766**	***0.00086***	***0.00082***	**0.00137**	0.842	0.0843
FG	beta	0.02586	0.01042	0.01486	0.0391	-0.10199	-0.005246
	se	0.03941	0.03944	0.03933	-	0.06283	0.036599
	pval	0.512	0.792	0.706	0.316	0.105	0.886
BMI	beta	0.04548	0.03599	0.02809	0.00314	-0.0027122	-0.06371
	se	0.03678	0.03674	0.03669	-	0.0603682	0.03462
	pval	0.216	0.327	0.444	0.931	0.964	0.0659
WC	beta	0.011598	0.03176	0.03877	0.0369	-0.05687	-0.02493

	se	0.036959	0.03683	0.03677	-	0.06042	0.03476
	pval	0.7537	0.3886	0.2918	0.313	0.3468	0.4734
WirC	beta	-0.009856	0.0151	0.005258	0.00412	0.0007305	-0.005438
	se	0.037937	0.03785	0.03777	-	0.0615324	0.03555
	pval	0.795	0.6899	0.8893	0.912	0.9905	0.8784
SBP	beta	0.012364	0.02692	0.02037	0.0175	0.18191	-0.05053
	se	0.036922	0.03682	0.03676	-	0.06025	0.03475
	pval	0.738	0.465	0.58	0.632	**0.01285**	0.146
DBP	beta	-0.004714	0.004176	-0.007108	-0.0144	0.1623	-0.03039
	se	0.036968	0.036832	0.036788	-	0.06025	0.03479
	pval	0.899	0.91	0.847	0.693	**0.0357**	0.383
HR	beta	-0.01559	0.004266	0.01921	0.012	0.0877	0.06831
	se	0.03796	0.037885	0.0378	-	0.06169	0.03567
	pval	0.6813	0.9104	0.6114	0.75	0.1554	0.05571
UR	beta	-0.03218	-0.03477	-0.02557	-0.0216	0.028311	-0.05321
	se	0.03679	0.03672	0.03667	-	0.059787	0.03451
	pval	0.382	0.3439	0.4857	0.552	0.6359	0.1233
CREA	beta	-0.04357	-0.0558	-0.06174	-0.0637	0.0533	0.009019
	se	0.03674	0.03664	0.03658	-	0.05967	0.034541
	pval	0.2358	0.128	0.09164	0.0797	0.3719	0.794
T2D	beta	-0.06142	-0.055759	-0.030537	0.0099503	0.162563	0.032447
	se	0.15141	0.149095	0.148631	-	0.253219	0.146279
	pval	0.8255	0.9523	0.8908	0.973	0.7104	0.9661
HTN	beta	0.021284	0.029566	-0.003233	0.0099503	0.402659	-0.1361429
	se	0.088038	0.087797	0.087840		0.1436035	0.082461
	pval	0.3255	0.4475	0.8477	0.878	0.05767	0.1891
CAD	beta	-0.29514	-0.30918	-0.28848	-0.2484613	0.07424	0.02221
	se	0.24145	0.23918	0.23962	-	0.41132	0.23632
	pval	0.2972	0.2242	0.24	0.299	0.9151	0.7764
HTC	beta	0.138029	0.148877	0.173474	0.1906203	0.0097577	0.1433676
	se	0.080282	0.080229	0.080170		0.1302524	0.0756111
	pval	**0.04539**	**0.058776**	**0.039708**	**0.0164**	0.9891	**0.02806**
Statin	beta	0.004711	-0.105805	-0.106027	-0.1031407	-0.202124	0.2445135
response	se	0.32397	0.30384	0.30382	-	0.50637	0.33915
	pval	0.3579	0.2649	0.2649	0.734	0.9569	0.3522

*Table 5: P-values reported in **bold** are significant before Bonferroni correction (P<0.05) while P-values reported in **bold and italic** are significant also after Bonferroni correction (42 independent tests, $P<1.2x10^{-3}$)*

In the GBHS the *KIF6* rs20455 (p.W719R) variant was associated with APOB levels (P=0.00766, $P_{corrected}$=0.322) and HDL-C (P=0.00124, $P_{corrected}$=0.0521) but both associations were no longer significant after Bonferroni correction for 42 tests.

The other two intronic and highly correlated *KIF6* variants were both associated with APO-B levels, even significant after Bonferroni adjustment. (rs9471077-G:P=0.00082, $P_{corrected}$=0.0344 and rs9462535-A: P=0.00086, $P_{corrected}$=0.0361).

In GBHS both SNPs were also associated with higher levels of TG (p-value$_{rs9471077c}$=0.0191, p-value$_{rs9462535c}$=0.0134), and lower HDL-C (p-value$_{rs9471077c}$=0.04515, p-value$_{rs9462535c}$=0.0335) but not significant after multiple testing correction.

Of these newly discovered associations, only the one between APOB levels and variants rs9471077 and rs9462535 was confirmed in replication analysis in two Estonian independent samples (Table 6). *HMGCR* association with blood pressure showed a similar trend towards association in replication cohorts. Results were significant when we applied a meta-analysis of GBHS results with results in the Estonian cohorts.

Testing models that took into account also the interactions between *KIF6* and *HMGCR* polymorphisms, we did not find any significant association with analysed cardiometabolic phenotypes. This indicates that there is no gene-gene interaction between *KIF6* and *HMGCR* in our data.

SNP	Chromosome	Position (hg19)	Hardy-Weinberg p-value	Genotype call rate (%)	MAF	Alleles (Major/Minor)
rs3761740	5	74632133	0,6962	99,9	0,096	C/A
rs3846662	5	74651084	0,2	99,9	0,457	A/G
rs9471077	6	39308742	0,1127	100	0,384	A/G
rs9462535	6	39315792	0,1045	100	0,381	C/A
rs20455	6	39325078	0,2982	99,9	0,365	T/C

Table 6 Genotyping results: Statistics and quality control of the five SNPs investigated in this study

We performed haplotype estimation at the *KIF6* locus we observed three possible haplotypes comprising the three SNPs rs9471077-rs9462535-rs20455 : GAT (frequency=0.606), ACC (frequency=0.355) and ACT (frequency=0.023).

Interestingly, GAT haplotype was significantly associated with an increase of lipids values ($0.001<P<0.03$) confirmed also in the Estonian cohort only for APOB (Table 6*)*. Significant association of ACT haplotype was also observed with risk of HTC (*Table 5*).

From the correlation analysis between considered phenotypes we identify three groups of highly correlated phenotypes (*Figure 7*): (1) SBP, DBP and HTN; (2) APOB, LDL and TC; (3) APOA1 and HDL. For each of these groups we retained only the phenotype with less missing data in our sample: SBP for the first group, LDL for the second one and HDL for the third one.

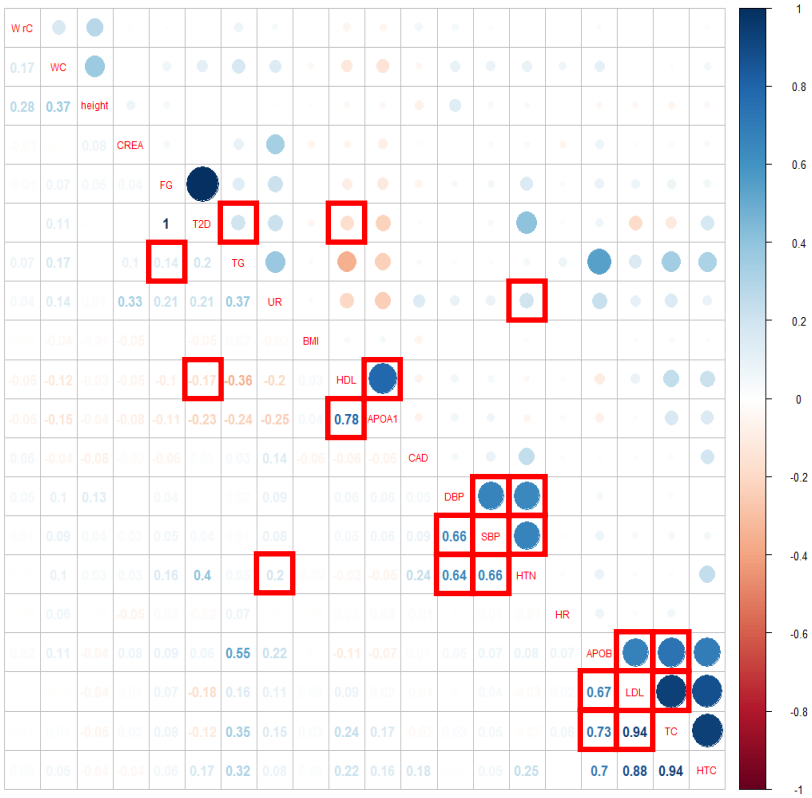

*Figure 7.**Correlation analysis between considered phenotypes.** Given are the values of the correlation coefficient r. Our analysis identified three groups of highly correlated phenotypes (r>|0.5|): (1) SBP, DBP and HTN; (2) APOB, LDL and TC; (3) APOA1 and HDL. For each of these groups we retained only the phenotype with less missing data in our sample: SBP for the first group, LDL f or the second one and HDL for the third one.*

In multi-trait effect phenotype association analysis we tested the complete set of all 14 quantitative and 5 discrete phenotypes included in our study. The multi-phenotype analysis showed the strongest association with individual phenotypes and did not identify any significant multi-trait effect for tested variants (*Table 7):* single phenotype models were most significant included HDL-C for rs20455, TG for rs9462535 and rs9471077, LDL-C for rs3846662 and SBP for rs3761740 with p-values as low as *P*=0.00085.

rs20455							
Phenotypes Set	LogLikelihood	nullLogLikelihood	LikelihoodRatio	P-value	BIC	BICnull	Model
set 1	-1647,2	-1651,39	8,37702	**0,00379995**	3308,68	3309,91	HDL-C
set 2	-2152,88	-2155,64	5,50874	0,0189216	4320,57	4318,68	HTC
set 3	-1647,2	-1651,39	8,37702	**0,00379995**	3308,68	3309,91	HDL-C
rs9462535							
Phenotypes Set	LogLikelihood	nullLogLikelihood	LikelihoodRatio	P-value	BIC	BICnull	Model
set 1	-1669,49	-1674,07	9,14531	**0,00249357**	3353,27	3355,27	TG
set 2	-2183,45	-2185,66	4,42991	0,0353145	4381,7	4378,73	HTC
set 3	-1669,49	-1674,07	9,14531	**0,00249357**	3353,27	3355,27	TG
rs9471077							
Phenotypes Set	LogLikelihood	nullLogLikelihood	LikelihoodRatio	P-value	BIC	BICnull	Model
set 1	-1671,37	-1676,04	9,33939	**0,0022428**	3357,02	3359,22	TG
set 2	-2186,7	-2189,02	4,62811	0,0314522	4388,21	4385,44	HTC
set 3	-1671,37	-1676,04	9,33939	**0,0022428**	3357,02	3359,22	TG
rs3761740							
Phenotypes Set	LogLikelihood	nullLogLikelihood	LikelihoodRatio	P-value	BIC	BICnull	Model
set 1	-768,509	-772,454	7,8897	**0,0049717**	1551,3	1552,05	SBP
set 2	-1034,87	-1035,78	1,82159	0,177124	2084,54	2078,96	HTC
set 3	-768,509	-772,454	7,8897	**0,0049717**	1551,3	1552,05	SBP
rs3846662							
Phenotypes Set	LogLikelihood	nullLogLikelihood	LikelihoodRatio	P-value	BIC	BICnull	Model
set 1	-1734,16	-1740,02	11,7256	**0,000616467**	3482,59	3487,18	LDL-C
set 2	-2262,15	-2265,74	7,17982	**0,00737281**	4539,1	4538,88	HTC
set 3	-1734,16	-1740,02	11,7256	**0,000616467**	3482,59	3487,18	LDL-C

Table 7.P-values reported in bold are significant also after Bonferroni's corrections

4.4 *Discussion*

In this genetic study in up to 12.307 individuals from GBHS and EGCUT studies we identified and validated a novel association of *KIF6* variants with APOB levels, proved with several statistical approaches, including haplotype, single- and multi-phenotype analyses. This relationship is reported here for the first time and haven't previously been explored either in lipid GWA studies, or in individual analyses.

Apolipoprotein B is a carrier of lipids through the body to the tissues and is representative of the total number of atherogenic particles in the blood, since it binds, not only LDL-C, but also very low-density lipoprotein (VLDL) and intermediate-density lipoprotein, large buoyant LDL-C, small dense LDL-C and lipoprotein(a)[90]. While APOB levels are rarely measured within large-scale studies, in contrast to other major lipid groups; many studies affirm that APOB level is a better measure than LDL and others to predict the risk of CVD and to evaluate the efficacy of statin treatment, than any other cholesterol index[91,92,93,94,95]. Based on these premises, the fact that *KIF6* has never been associated with classic lipid (LDL-C, HDL-C, TG or TC), its debated effect on efficacy of statin therapy, and the significant association we found with APOB level may underlie the hypothesis under which genetic variants on *KIF6* have an effect on susceptibility to statin treatment (in terms of risk of developing CVD) through an effect on APOB. This hypothesis, if confirmed by further studies, may help in setting up novel approaches of diagnosis and treatment for hyperlipidaemia and CVD.

KIF6 has been reported as associated with differential benefit, in terms of CVD events, from lipid-lowering therapy with statins. In particular, the presence of rs20455-G variant had greater benefit[96,97] suggested to act through an increased vulnerability to LDL cholesterol[98]. However, other studies have not confirmed such findings[99,100,101]

In our study, we did not find any significant association of *KIF6* genotype at the three most studied polymorphisms with the response to lipid-lowering therapy, in terms of risk of fatal CVD outcomes in ten years. Additionally, we did not find any effect of interaction between the two studied genes, suggesting that *KIF6* and *HMGCR* in our sample act on lipid levels through different pathways.

We confirmed the association of *HMGCR* locus with lipids, already reported in large-scale GWA meta-analyses[102,103]. Moreover, other unexpected associations were observed in GBHS, for example that of *HMGCR* with blood pressure. The replication of this association was obtained in a meta-analysis of our cohort with other two Estonian cohort, but was not significant in the two replication cohorts taken separately. These results may be because of more little effect of the gene on blood pressure in the Estonian cohort then in GBHS that might derive from exposition to particular influencing environmental factors in the Brisighella population, or from interaction with variants in other genes, which have bigger inter-population differences.

The null findings from our study about role of KIF6 variants in statin therapy susceptibility may be due to a real lack of power, given the sample size. It can also be biased by the type of data used for the analyses. In fact, due to unavailability of longitudinal data, we used an estimation of the ten-year CVD risk instead of a real registration of CVD events. Moreover, most of the significant results reported in the literature are from studies of response to the therapy with an individual statin molecule[64,65,104]; in our data, instead, lipid-lowering therapy included a mix of several statin molecules (simvastatin, atorvastatin, rosuvastatin, fluvastatin) and that could have masked the effect on response to only one particular type of them. Additionally, other studies have suggested that *KIF6* gene may reveal unexpected associations, when studied in single cohorts. Further studies, based meta-analysis of the literature and stratification for type of statin would be better suited to overcome these limitations.

Our study has some additional limitations. Firstly, the sample size of our discovery and replication sample is relatively small, further replication analyses in larger samples is required to confirm our novel findings. Moreover, the biological validity and clinical usefulness of the obtained results has to be verified using functional approaches, such as the one proposed by Shen and colleagues[105]. Finally, the set of cardiometabolic diseases used in our analysis has been chosen arbitrarily, even if the diagnostic cut-off applied are those suggested by the main specific international guidelines.

Beyond that, the data are reliable and obtained from two European population-based studies, well-characterised and defined[106,107].

In conclusion, this study highlighted a novel effect of *KIF6* variants on APOB levels, which might be predictive of CVD risk, but wasn't sufficiently empowered to see the effect on statin therapy susceptibility. Further investigation of these findings would be important from additional studies. Additionally, evaluation of biological mechanisms underlying this mechanistic relationship will be better suited to uncover the biological relationship between the *KIF6* gene and lipids metabolism, to advance our knowledge and possibly implement findings from this studies in more personalised approaches for CVD prevention and therapy.

5 Mendelian Randomization Analysis of associations between Serum Uric Acid and cardiovascular disease

5.1 Serum Uric Acid

In most mammals Serum Uric Acid (SUA), also known as Urate, is largely considered to be waste product of purine metabolism. In humans the kidneys excrete 70 percent of the metabolized SUA, yet circulating plasma SUA has also been shown to be associated with many physiological pathways in contrast to its waste product denomination[108]. Purine metabolism is an essential biological pathway that degrades excess purines such as Adeninemonophosphate (AMP), Adenosine, Guanine monophosphate (GMP), Guanosine, Inosine, and Hypoxanthine to SUA[109] and is catalyzed by the enzyme Xanthine Oxidase (XO) (Figure 8). Most mammals further metabolize SUA to Allantoin, via the Uricase enzyme Urate Oxidase (UO) encoded by the Uox*gene. UO expression in man and higher primates has evolutionarily been suppressed by the inheritance and positive selection of several nonsense mutations that render the* Uox*gene silenced[110]. Some evidence suggests that loss of UO expression may have imparted an evolutionarily advantage, and an interesting postulation by Orowan (1955), suggests a similar function of SUA to cerebral stimulants such as caffeine and theobromine in the rise of man. The* Uox *gene is considered to be truly ancient as it is observed and expressed in prokaryotes and eukaryotes*[110]*.Deregulation of SUA excretion and an inability to metabolize SUA, can lead to elevated levels of SUA (hyperuriceamia), which has been epidemiologically observed in many metabolic diseases such as Hypertension, Type 2 Diabetes (T2D), and Obesity, and elevated SUA is known to be the main cause of Gout[111]. Whether, SUA is a true risk factor, or just a compensatory response in these complex diseases still remain to be determined*

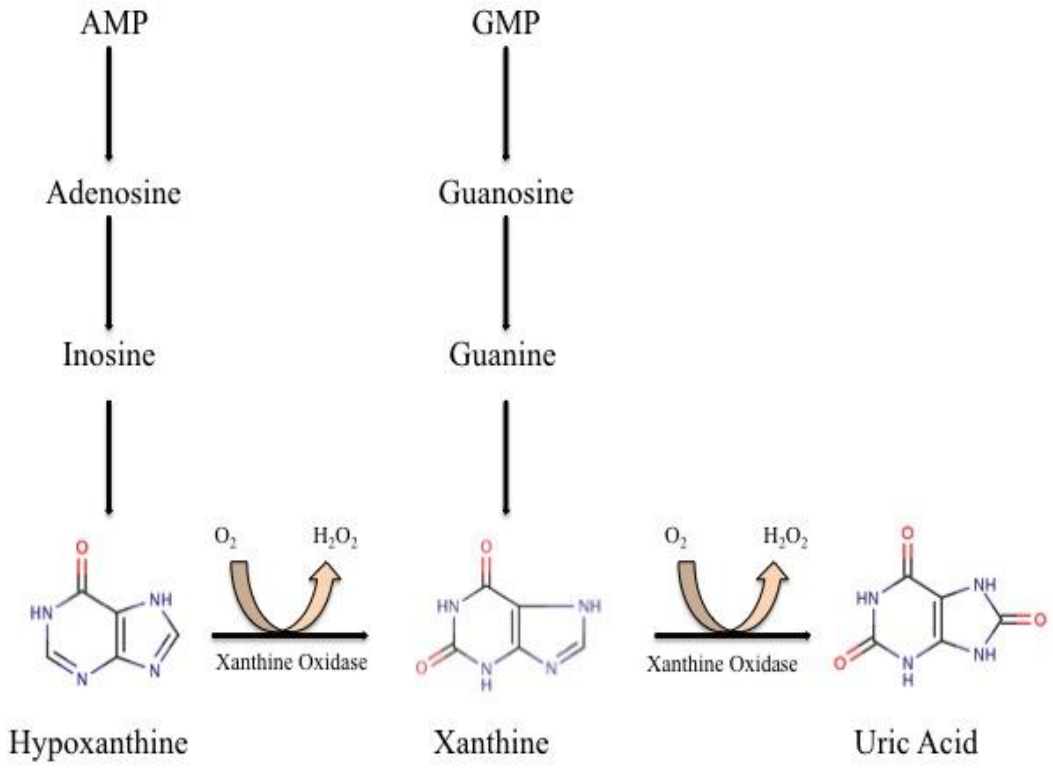

Figure 8. The purine metabolism pathway. AMP is degraded into Hypoxanthine, and GMP into Xanthine. Hypoxanthine and Xanthine are oxidized by Xanthine Oxidase to produce Uric Acid. Adapted from George & Struthers, 2009.

5.2 Association of Serum Uric Acid and Cardiovascular Disease

The association between elevated SUA and CVD has been observed epidemiologically in many independent and archaic studies[112]. More recent evidence also suggests a potential causative role of elevated SUA in CVD, although there is no direct biological interaction defined for the association; and as of such SUA is not yet recognized as a true risk factor in CVD (World Heart Federation, 2014). Other previous studies have sought to define the relationship of SUA to CVD, via interactions of SUA to the many known risk factors associated with CVD. One such

interaction is the activation of the renin-angiotensin system, loss of the macula densa nitric oxidase system, and development of micro vascular disease through the production of reactive oxygen species facilitated by increased prooxidant activity of elevated SUA[113]. Another suggested interaction is the decrease in size of Low Density Lipoprotein (LDL) and High Density Lipoprotein (HDL). Abnormal lipid biology has long been recognized as a clear risk factor in Atherosclerosis, and evidence shows that elevated SUA promotes *in vitro* oxidation of LDL and HDL particles, attenuating their antioxidant properties and increases the likelihood of CVD[114].An interesting observation in regards to current treatments of hyperuriceamia is the reduction in systolic blood pressure (SBP) and diastolic blood pressure (DBP) in chronic hyperuriceamia patients recently diagnosed with hypertension and who are treated with the XO inhibitor Allopurinol. Patients showed a decrease of 6.9 mm Hg in systolic BP, and a decrease of 5.1 mm Hg diastolic BP when treated with Allopurinol. In contrast, the placebo group showed a decrease of 2.0 mmHg systolic BP, and a decrease of 2.4 mm Hg diastolic BP[115]. In spite of this, no definitive biological mechanism has been reported, but activation of the renin-angiotensin system is a potential mechanism of action. Nonetheless, this interesting relationship requires further exploration to define Allopurinol as a therapeutic drug for hypertensive patients due to proposed adverse effects.

5.2.1 **Genetics of Serum Uric Acid levels**

Hyperuriceamia, and in general abnormal SUA levels, has been widely associated with variants in several genes through the implementation of Genome Wide Association Studies (GWAS). Abnormal SUA levels have an estimated heritability of 40-70% [116,117]. At present 46 loci have shown association with abnormal SUA levels, including the 18 most recent loci identified in a large meta-analysis on >140,000 individual by Kottgen et al., 2013 (*Table 8 and Figure 9*). Of these 46, a Single Nucleotide Polymorphism (SNP) within the gene *SLC2A9* has shown the most

independently replicated association in many studies[118,119,120]. SLC2A9 is a member of the SLC2A facilitative glucose transporter family, and has been described in its role as a urate transporter[121]. Another highly replicated association is within the *ABCG2* gene, which encodes an ATP Binding Cassette protein, and has recently also been described as an urate transporter[122]. Defective urate transport and clearance is a clear mechanism in hyperuriceamia, and association with urate transport genes is not surprising. Nonetheless, whether variants in these genes are also associated with CVD remains to be elucidated.

Loci	**Chromosome**	**Sample Size**	**Population**	**Author**
PDZK1	1	Meta >140,000	European	Kottgen A (2013)
TRIM46	1	Meta >140,000	European	Kottgen A (2013)
GBAP1	1	Meta 71,149	Chinese	Okada Y (2012)
PPP1R12B	1	1,957	Korean	Son C (2014)
GCKR	2	Meta >140,000	European	Kottgen A (2013)
INHBB	2	Meta >140,000	European	Kottgen A (2013)
ORC4	2	Meta >140,000	European	Kottgen A (2013)
LRP2	2	~14,700	Japanese	Kamatani Y (2010)
MUSTN1	3	Meta >140,000	European	Kottgen A (2013)
AC007953.1	3	1,017	African Americans	Charles BA (2011)
SLC2A9	4	Meta >140,000	European	Kottgen A (2013)
GLUT9	4	6,148	European	Li S (2007)
WDR1	4	1,017	African Americans	Charles BA (2011)
GLUT9	4	868	Old Amish	McArdle PF (2008)
ABCG2	4	Meta >140,000	European	Kottgen A (2013)
SEMA5A	5	1,300	European	Karns R (2012)

TMEM171	5	Meta >140,000	European	Kottgen A (2013)
AL162718.1	6	1,017	African Americans	Charles BA (2011)
RREB1	6	Meta >140,000	European	Kottgen A (2013)
LRRC16A	6	Meta 28,141	European	Kolz M (2009)
SLC17A1	6	Meta 28,283	European	Yang Q (2010)
SLC17A3	6	7699	European	Dehghan A (2008)
VEGFA	6	Meta >140,000	European	Kottgen A (2013)
ARID1B	6	1,017	African Americans	Charles BA (2011)
MLXIPL	7	Meta >140,000	European	Kottgen A (2013)
PRKAG2	7	Meta >140,000	European	Kottgen A (2013)
STC1	8	Meta >140,000	European	Kottgen A (2013)
HNF4G	8	Meta >140,000	European	Kottgen A (2013)
RP11-523018.1	10	1,017	African Americans	Charles BA (2011)
ASAH2	10	Meta >140,000	European	Kottgen A (2013)
SLC16A9	10	Meta >140,000	European	Kottgen A (2013)
SLC22A11	11	Meta >140,000	European	Kottgen A (2013)
NRXN2	11	Meta >140,000	European	Kottgen A (2013)
MAP4K2	11	Meta 21,708	African Americans	Tin A (2011)
LTBP3	11	Meta >140,000	European	Kottgen A (2013)
INHBC	12	Meta >140,000	European	Kottgen A (2013)
PTPN11	12	Meta >140,000	European	Kottgen A (2013)
NRG4	15	Meta >140,000	European	Kottgen A (2013)
IGF1R	15	Meta >140,000	European	Kottgen A (2013)
NFAT5	16	Meta >140,000	European	Kottgen A (2013)
MAF	16	Meta >140,000	European	Kottgen A (2013)

HLF	17	Meta >140,000	European	Kottgen A (2013)
C17orf82	17	Meta >140,000	European	Kottgen A (2013)
PRPSAP1	17	Meta >140,000	European	Kottgen A (2013)
MYO18B	22	1,017	African Americans	Charles BA (2011)

Table 8. Known Loci associated with Serum Uric Acid Levels, ordered by chromosome position. Sample size of discovery cohort, the population and author is provided.

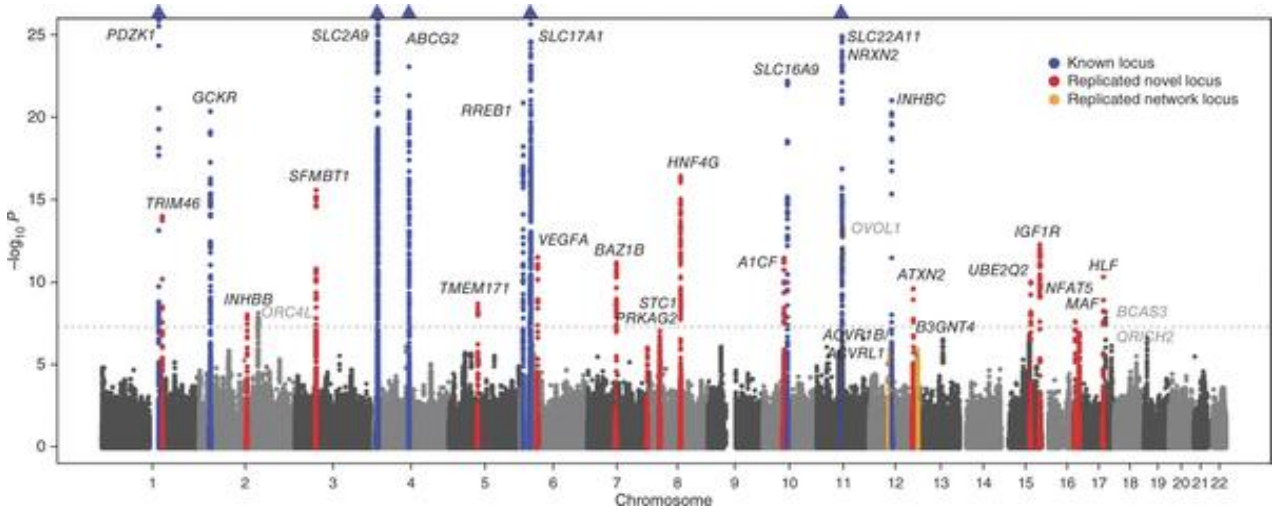

Figure 9. 28 genomic loci contain SNPs associated with SUA concentration. A manhattan plot showing –log$_{10}$ P values of 10 previously known replicated loci, and 18 novel loci discovered in a large meta-analysis of >140,000 individuals. The gene located closest to the SNP with the lowest P value at each locus is listed. Loci in gray met one but not both replication criteria. Blue triangles represent loci containing SNPs with P values below 1 x 10^{-25}.Reproduced with permission from Köttgen et al., 2013.

5.3 Potential usefulness of Mendelian Randomization

To disentangle a potential causal relationship between SUA and CVD, we propose a Mendelian Randomization (MR) approach. MR was first proposed in 1986 by Katanto evaluate whether low levels of LDL cholesterol increased cancer risk (Katan, 1986). Katan noted that observational studies had reported a higher risk of cancer in individuals with low levels of LDL, compared to subjects with normal or elevated LDL levels. However, a major limitation of observational studies that have identified an epidemiological association between an intermediate phenotype and an

outcome is whether these associations are causal or spurious due to problems of confounding and reverse causation. Reverse causality occurs when the outcome affects the levels of the intermediate phenotype, rather than the expected direction of effect. Confounding is when an unknown, or known, factor influences both the intermediate phenotype and the outcome of interest.

MR implements the use of genetic variation, which are assigned randomly at conception, to infer causality of an intermediate phenotype to an outcome or disease[123]. Unlike a classical association study where only a single association is defined, a MR approach relies on the association between 3 distinct factors. First, the association of an instrumental variable (IV), a genetic variant, to the intermediate phenotype has to robust and well established in independently replicated studies. Second, the association of the IV and the outcome of interest, which can be determined by a qualitatively or quantitatively study design. And finally, estimates of these associations can hence, be used to infer causality of the intermediate phenotype to the outcome of interest (Figure 10)[124].

5.3.1 Limitations of Mendelian Randomization

For a MR study to be successful, several limitations need to be addressed. Genotypic effects on phenotype are typically small so MR can require very large sample sizes. And hence, for MR to be an effective method there has to be a well-defined association of the IV to the intermediate phenotype, and in the context of a genetic analyses robust association of a SNP to the intermediate phenotype. Another core consideration of a MR analysis requires that no pleiotropy is present within the IV, and the interaction of the IV to the outcome is only mediated via the intermediate phenotype. Nonetheless, to completely evaluate the pleiotropic effects of an instrumental variable, both the function of the gene and its variants need to be known which is rarely the case and hence, a stringent selection criteria for these instrumental variables are required[125].

MR studies are generally conducted with single IVs, but the use of independent multiple instrumental variables (mIV) to limit the effect of pleiotropy and increase statistical power has also been well implemented in some MR studies (Palmer et al., 2012). However, mIV are more prone to other confounding such as population stratification and linkage disequilibrium LD) (*Figure 10*) Population stratification, a potential problem in all association studies, occurs when ancestral origins determine the prevalence of certain IV in the sample. Nonetheless, there is debate regarding the extent to which this can generate spurious associations, and several statistical techniques, such as Genomic Control, are commonly implemented in association studies to counter effect size inflation[126]. Confounding due to LD can occur when an IV, in high LD with another genetic variant, influences the outcome not through the intermediate phenotype, but through a different mechanism. However, selection of an IV that is shown to be in LD with a known genetic variant that is associated with the outcome but mediates its effect via the intermediate phenotype would increase the likelihood of inferring a true causal relationship. Furthermore, increasing statistical power by using independent mIV in different regions of the genome would significantly dampen confounding effects due to, either pleiotropy and/or LD, as genes in discrete genomic regions would unlikely have the same biological mechanisms[127]. Therefore, in repetition, a stringent IV selection criterion is essential to limit the possible biases imparted by the confounding factors.

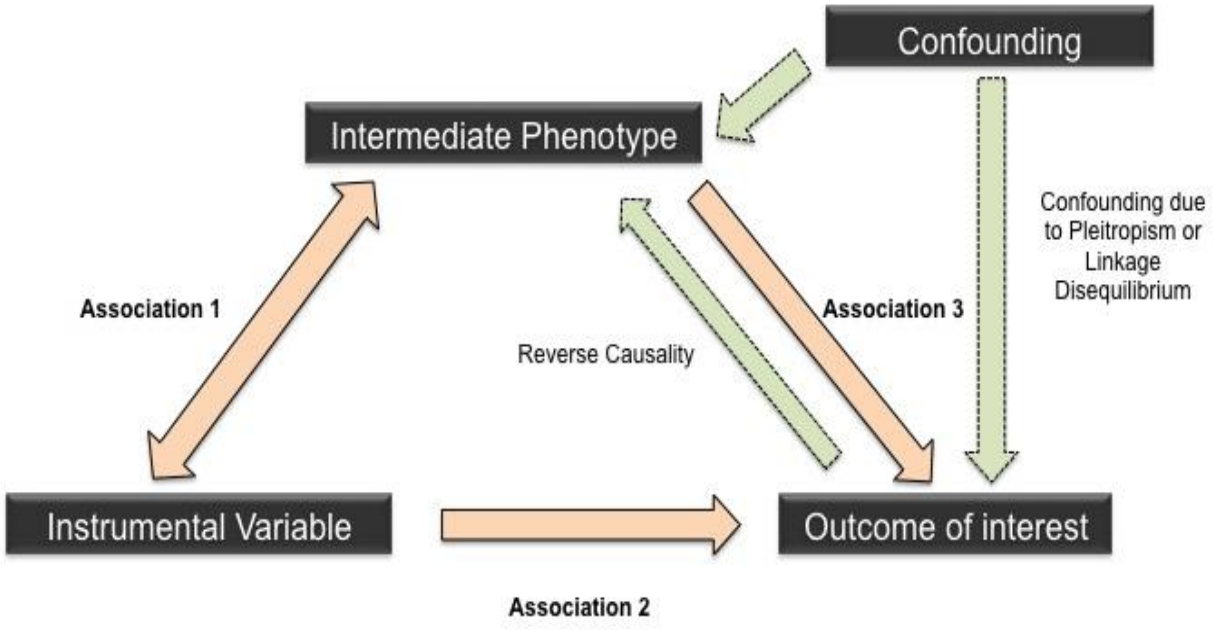

Figure 10. Diagram showing the hypothesis underlying a classic Mendelian Randomization approach. Orange Arrows: Association 1, between the instrument variable (IV) and intermediate phenotype (IP), should be well established in literature. Association 2, between the IV and the outcome of interest (OI), should be mediated through the IP only. Association 3, between the intermediate phenotype and the OI, is the causal relationship inferred from the estimated values of association 1 and 2. Green Arrows: Possible limitations of a Mendelian Randomization study such as reverse causality and confounding via IV pleiotropy or LD.

5.4 Relationship between Serum Uric Acid and Cardiovascular Disease in Brisighella Heart Study: A Mendelian Randomization Analysis

5.4.1 Hypothesis

We hypothesized that increased SUA levels affect susceptibility to CVD and the genetic effects observed in cardiometabolic traits are mediated through their effect on SUA levels - investigating the evidence provided by the population based Brisighella Study.

5.4.2 Aims

The aim of this study was to implement a Mendelian Randomization approach to define a causal relationship between SUA and CVD. To this aim, we used genetic variants associated with SUA, by genotyping 1656 samples from the Brisighella Cohort using Fluidigm Integrated Fluidic Circuit (IFC) and TaqMan systems, as our instrumental variables.

5.4.3 Research Plan

Initially, we performed a comprehensive background literature search to identify SNPs associated with SUA levels allowed for efficient selection of variants to use in MR. Fluidigm IFC system required gDNA to be normalized and pre-amplified for the selected SNPs. Testing of two amplification methods determined the most appropriate. Genotypes from Fluidigm IFC and TaqMan systems were then quality controlled using the R statistical package. Association of the variants to the phenotype was tested using PLINK, and a Genetic Risk Score, using a simple allele counting method, determined causal relationship of the variants to CVD, and or, related traits.

5.4.4 Materials and Methods

Detailed information of the Brisighella Heart Study has been published previously[128]. Briefly, it is a classic epidemiological study where 2939 randomly selected subjects (1491 men and 1448 women) from the northern Italian rural town of Brisighella were enrolled at baseline in year 1972. These subjects were aged 14-84 years, free of cardiovascular disease, and underwent follow-up examinations every 4 years. This included measurements of BMI, SBP, DBP, SUA, and Creatinine. Logistic traits such as Hypertensions (HTN), Hypercholesterolemia (HC) and whether they suffer from CHD was also recorded. Individual mortality and morbidity was also recorded (*Table 9*) This study was conducted using gDNA extracted from blood by a standard extraction protocol from 1656individualsbetween 1973-2012 follow-up periods.

SNP Selection

A comprehensive list of SNPs associated with SUA discovered from a background literature search is provided . Several criteria had to be met for a SNP to be suitable for MR analysis. Firstly, the SNP had to be robustly and independently associated with SUA, which was determined using the published literature for the relevant SNP. Due to the recent publication of their meta-analysis by Kottgen et al., 2013, priority was given to their SNPs with the lowest p-values to ensure appropriate statistical association with SUA. Secondly, using the online resource dbSNP (build 37) and 1000 genomes (CEU), all SNPs were checked for association with CVD and/or CVD risk factors and LD with other SNPs that may be associated with either; CVD directly, or CVD related traits such as HTN or obesity. This was to provide greater scope for a potential causal mechanism in SUA mediated CVD. Priority was given to SNPs that had direct association with CVD or CVD risk factors or were in LD ($r^2 \geq 0.2$) with other SNPs associated with CVD or CVD risk factors. Only independent (approximately 250kb apart and no LD present $r^2 < 0.2$) SNPs were selected to ensure

efficient genome coverage. Lastly, non-synonymous SNPs were given priority over synonymous. The final 24 selected SNPs are listed in (*Table 10*)

Total subjects (n)	**1656**
Men	794
Women	859
Age (years)	
Mean (±)	53.20 (±18.06)
Median	51
Hyprecholestereamiaeamia (n)	
Yes (%)	596
No (%)	886
Diabetes (n)	
Yes (%)	70
No (%)	1505
Hypertension (n)	
Yes (%)	912
No (%)	658
Coronary Heart Disease (n)	
Yes (%)	38
No (%)	1551
BMI (kg/m2)	
Mean (± SD)	26.46 (±4.54)
Median	25.91
Systolic Blood Pressure (mmol/Hq)	
Mean (± SD)	137.31 (±20.30)
Median	135
Diastolic Blood Pressure (mmol/Hg)	
Mean (± SD)	85.90 (±11.29)
Median	86
Fasting Uric Acid (mg/dL)	
Mean (± SD)	5.06 (±1.55)
Median	4.9
Creatinine (mg/dL)	
Mean (± SD)	0.93 (±0.18)
Median	0.9

Table 9. Subject Characteristics for Brisighella Heart Study Cohort

rs ID	Ch	Position (b37)	*Locus Name*	Effect (mg/dL)	P-Value	SNP	AA[b] Variation	Other Associations for SNP	Other associations in the locus (approximate distance between the two SNPs-SNP rsID(r2))	CEU MAF[c] (dbSNP b126)	Study First Author[d] and Date
rs1471633	1	145,723,739	***PDZK1***	0.059	1.2 × 10–29	C>A				0.45	Kottgen A (2013)
rs11264341	1	155,151,493	***TRIM46***	–0.050	6.2 × 10–19	C>T			Obesity(20kb)-rs11264330, Parkinsons(20kb)-rs12726330	0.44	Kottgen A (2013)
rs1260326	2	27,730,940	***GCKR***	0.074	1.2 × 10–44	C>T	Leu to Pro	2hGlu,TG, TC	FI, FG, T2D, WC-TG, PL (10kb)-rs780094(0.933)/rs780093(0.901)/rs1260333(0.676)	0.42	Kottgen A (2013)
rs2307394	2	148,716,428	***ORC4***	–0.029	2.2 × 10–8	A>G	Asn to Ser		Response to statin therapy (220kb)-rs7584099(0.184)	0.38	Kottgen A (2013)
rs6770152	3	53,100,214	***MUSTN1***	–0.044	2.6 × 10–16	G>T			Height (20kb)-rs2336725(0.264), Ventricular contraction (200kb)-rs4687718	0.51	Kottgen A (2013)
rs12498742	4	9,944,052	***SLC2A9***	0.373	0	A>G			CVD Risk factors -rs7671266(0.571)/rs4698036(0.525)	0.24	Kottgen A (2013)
rs2231142	4	89,052,323	***ABCG2***	0.217	1.0 × 10–134	C>A	Gln to Lys		IBD (10kb)-rs4656958, Crohns (10kb+20kb)-rs2274910, Response to statin therapy (10kb)-rs1481012(0.92), Serum Metabolites (150kb)-rs1440581	0.11	Kottgen A (2013)
rs675209	6	7,102,084	***RREB1***	0.061	1.3 × 10–23	C>T			Adipose Tissue (10kb)-rs2842895, FG (100 kb)-rs2714337(0.283), Waist-Hip ratio (100kb)-rs6931262	0.71	Kottgen A (2013)
rs1165151	6	25,821,616	***SLC17A1***	–0.091	7.0 × 10–70	A>C			Fe(50kb+150kb)-rs932316/rs2274089/rs17270561/rs17342717, Platelet(50kb)-rs441460/rs12526480, Cimt(150kb)-rs4712972, CVD(150kb)- rs11754288(0.875),RBC Traits(200kb)-rs17342717/rs1408272	0.48	Kottgen A (2013)
rs729761	6	43,804,571	***VEGFA***	–0.047	8.0 × 10–16	G>T			Adiponectin(100kb)-rs998584, CHD+WHR (150 kb)-rs6905288, CKD(20kb)-rs881858(0.83), Thyroid Hormones(20kb+50kb)-rs9472138/rs11755845, VEGF(50kb)-rs9472155/rs4513773	0.75	Kottgen A (2013)
rs1178977	7	72,857,049	***MLXIPL***	0.047	1.2 × 10–12	A>G			TG, HDL (10kb-150kb)-rs3812316(0.778)/rs2286276(0.666)/rs17145738(0.725)/rs714052(0.725)/rs1178979(1.00)/rs2240466(0.725), Gamma GlutyTransferase(150kb)-rs17145750(0.832)/rs12539316(0.697), Lipid Metabolism(150kb)-rs13226650(0.945)/rs13247874(0.945), Obesity(125kb)-rs12666883	0.18	Kottgen A (2013)
rs10480300	7	151,406,005	***PRKAG2***	0.035	4.1 × 10–9	C>T		RBC Traits	CKD(5kb)-rs7805747(1.00), Heamoglobin(10kb)-rs10224002(0.961)	0.25	Kottgen A (2013)
rs17786744	8	23,777,006	***STC1***	–0.029	1.4 × 10–8	A>G	Asn to Ser		CKD(20kb)-rs10109414(0.935), Adipose(20kb)-rs7833268(0.125), Height (300kb)-rs1013209	0.45	Kottgen A (2013
rs2941484	8	76,478,768	***HNF4G***	0.044	4.4 × 10–17	C>T			Obesity(130kb)-rs4735692, BMI(100kb)-rs2922763(0.201)	0.45	Kottgen A (2013)
rs10821905	10	52,646,093	***ASAH2***	0.057	7.4 × 10–17	A>G				0.2	Kottgen A (2013
rs1171614	10	61,469,538	***SLC16A9***	–0.079	2.3 × 10–28	A>G				0.8	Kottgen A (2013)
rs2078267	11	64,334,114	***SLC22A11***	–0.073	9.4 × 10–38	A>G			Obesity(20kb)-rs7124676(0.837), IBD(150kb)-rs559928, PBC(200kb)-rs538147	0.54	Kottgen A (2013
rs478607	11	64,478,063	***NRXN2***	–0.047	4.4 × 10–11	A>G			Obesity(100kb)-rs7124676, IBD(250kb)-rs559928, PBC(300kb)-rs538147	0.86	Kottgen A (2013)
rs3741414	12	57,844,049	***INHBC***	–0.072	2.2 × 10–25	A>G			TG+HDL(60KB)-rs11613352(0.959)	0.25	Kottgen A (2013)
rs653178	12	112,007,756	***PTPN11***	–0.035	7.2 × 10–12	A>G		DBP, CKD, Celiac Disease	CAD (130kb)-rs3184504(0.873), DBP/SBP (130kb)-rs3184504(0.873), LDL (70kb)-rs11065987(0.552), CHD(150kb)-rs11066015, Stroke(200kb)-rs2238151(0.213), TG(200kb)-rs671	0.56	Kottgen A (2013)
rs6598541	16	79,734,987	***MAF***	–0.032	1.6 × 10–9	A>G			T2D(250kb)-rs17797882,Height(100kb)-rs7498403,Obesity(50kb)-rs1424233,Thyroid(100kb)-rs17767419(0.932)/rs3813579(0.333)/rs3813582(0.932)	0.38	Kottgen A (2013)
rs7188445	17	59,465,697	***C17orf82***	0.043	1.2 × 10–8	A>G			Creatinine(20kb)-rs8068318(0.437), Height (30kb)-rs2079795/rs757608, Obesity(150kb)-rs17513268	0.17	Kottgen A (2013)
rs7224610	15	99,271,135	***IGF1R***	0.043	4.8 × 10–15	A>G			Height (75kb)-rs2871865	0.7	Kottgen A (2013
rs2079742	17	53,364,788	***HLF***	–0.042	5.4 × 10–17	A>C			Obesity(150kb)-rs11653011	0.58	Kottgen A (2013

Table 10 . 24 selected SNPs to be used in MR analysis

Genotyping

All 24 TaqMan SNP genotyping assays were ordered from Life Technologies and were 40x concentrations apart from rs2231142, which was 20x. Once received all assays were stored at -20° C and covered in opaque tape to minimize photo bleaching

of the VIC and FAM fluorophores. Initially, all samples were to be genotyped using Fluidigm 192.24 Integrated Fluidic Circuits (IFC) system (http://www.fluidigm.com/chips-kits.html), which would have allowed for multiplexing of the 24 SNPs. However, after performing genotyping on the first 2 IFCs, rs12498742 did not perform well, and provided spurious genotype calls and poor amplification when multiplexed, but worked as expected when tested as a singleplex on the TaqMan 7900ht system. Therefore, all SNPs were finally genotyped on 192.24 IFC with the exception of rs12498742that was genotyped using a singleplex TaqMan protocol.

5.4.5 Results

Single SNP association to Uric Acid and CVD related traits

Association of genetic variants (n=24) to SUA showed, as expected, a positive effect of risk alleles in increasing levels of SUA (*Table 11*) We confirmed association for the most significantly associated, SNP rs12498742 (β = 0.41 per A allele; SE = 0.052; $p = 1.5 \times 10^{-15}$) at locus *SLC2A9*, and report an effect size was in line with that stated in previous studies (Kamatani et al., 2010; Karns et al., 2012; Köttgen et al., 2013; Tin et al., 2011; B. Yang et al., 2014). We also showed significant associations, before multiple testing correction, for 4 other SNPs; at locus *ABCG2* for rs2231142 (β = 0.23 per T allele; SE = 0.089; $p = 0.0092$), at locus *VEGFA* for rs729761 (β = 0.12 per G allele; SE = 0.55; $p = 0.0298$), at locus *PRKAG2* for rs10480300 (β = 0.10 per T allele; SE = 0.051; $p = 0.0483$), and at locus *SLC16A9* for rs1171614 (β = 0.19 per C allele; SE = 0.054; $p = 0.0005$). After a Bonferroni Correction ($p<0.002$), only rs12498742 and rs1171614 survived (*Table 11).*

Association of genetic variants to creatinine also showed several positive associations ($p<0.05$) before correction. The most significant association was observed at locus *GCKR* for rs1260326 (β = 0.019 per C allele; SE = 0.006; $p = 0019$), and also for

SNP variants; rs12498742 (β = 0.018 per A allele; SE = 0.007; *p* = 0.0072), rs1178977 (β = -0.017 per A allele; SE = 0.009; *p* = 0.0492), rs2078267 (β = 0.014 per C allele; SE = 0.006; *p* = 0.0253), and rs7224610 (β = 0.012 per A allele; SE = 0.006; *p* = 0.0482) at *SLC2A9*, *MLXIPL*, *NRXN2*, and *IGF1R* loci, respectively. After correcting for multiple testing only the association at rs1260326 survived (*Table 11*Table 11*)*.

We detected only a single significant association, before correction, for BMI at locus *SLC16A9*for SNP rs10821905 (β = -0.11 per A allele; SE = 0.048; *p* = 0.0219), but after correction it is not significant. Similarly, we observed a single significant association for SBP at locus *PRKAG2*for SNP rs10480300 (β = 1.36 per T allele; SE = 0.65; *p* = 0.038), but did not survive Bonferroni correction. Interestingly, we observed several associations for DBP with the most significant being SNP rs2941484 (β = -0.86 per T allele; SE = 0.38; *p* = 0.0231) at locus *HNF4G*and others for rs653178 (β = 0.82 per C allele; SE = 0.38; *p* = 0.0321) and rs2079742 (β = -1.13 per C allele; SE = 0.501; *p* = 0.0246), *MAF* and *HLF* loci respectively. Unfortunately, after multiple testing correction none of the SNPs survived (*Table 11*)

We identified 4 significant associations for CAD (*Table 12*), the most significant at locus *RREB1*for SNP rs675209(OR = 1.92 per T allele; 95% C.I = 1.16 -3.18*p* = 0.0098), and also at loci *MUSTN1*, *SLC17A1*, and *SLC16A9*, respectively for SNPs; rs6770152 (OR = 0.54 per G allele; 95% C.I = 0.32 – 0.901; *p* = 0.0017), rs1165151 (OR = 1.74 per G allele; 95% C.I = 1.09 – 2.80; *p* = 0.0198), and rs10821905 (OR = 1.84 per A allele; 95% C.I = 1.092 – 3.101; *p* = 0.0202). None of the 24 genotyped variants showed association with HTN (*Table 12*). For HC, we showed 3 at least nominally significant associations for SNPs; rs12498742 (OR = 0.78 per A allele; 95% C.I = 0.66 – 0.92; *p* = 0.003), rs17786744 (OR = 0.81per G allele; 95% C.I = 0.70 – 0.95; *p* = 0.0084), and rs2079742 (OR = 0.79 per C allele; 95% C.I = 0.64 – 0.96; *p* = 0.0192) at loci *SLC2A9*, *STC1*, and *HLF*, respectively. Finally, we also detected 3 significant associations for DB with the most significant being for SNP rs729761 (OR = 1.92 per G allele; 95% C.I = 1.19 – 3.11; *p* = 0.007) at locus

VEGFA. Others for SNPs; rs478607 (OR = 2.11 per A allele; 95% C.I = 1.10 – 4.06; p = 0.0218) and rs2307394 (OR = 0.66 per C allele; 95% C.I = 0.45 – 0.98; p = 0.038) respectively at loci *INHBC* and *ORC4*. However, after Bonferroni correction none of the above associations remain significant (*Table 12*)

					URIC ACID				CREA			BMI			SBP			DBP		
CHR	SNP	BP	***Locus Name***	Alleles (effect/other)	Sample Size	Effect	SE	*P Value*	Effect	SE	*P Value*	Effect	SE	*P Value*	Effect	SE	*P Value*	Effect	SE	*P Value*
1	rs1471633	145723739	***PDZK1***	A/C	1485	0.0552	0.0496	0.2659	0.0044	0.0063	0.4838	-0.0143	0.0376	0.7045	0.2747	0.6281	0.662	0.1393	0.3962	0.7252
1	rs11264341	155151493	***TRIM46***	C/T	1544	0.0379	0.0494	0.4421	0.0035	0.0063	0.5854	-0.0155	0.0376	0.6807	-0.0263	0.625	0.9665	0.0079	0.3941	0.984
2	rs1260326	27730940	***GCKR***	C/T	1533	0.0164	0.0485	0.7354	**0.0193**	**0.0062**	**0.0019**	0.0156	0.0369	0.6727	0.0251	0.6134	0.9673	0.3023	0.3871	0.4348
2	rs2307394	148716428	***ORC4***	C/T	1499	0.0107	0.0511	0.8344	-0.0009	0.0066	0.8933	-0.0029	0.039	0.9405	-0.9353	0.644	0.1466	-0.6862	0.4066	0.0917
3	rs6770152	53100214	***MUSTN1***	G/T	1526	0.0483	0.0495	0.33	0.0015	0.0064	0.8091	-0.0442	0.0378	0.2421	0.0407	0.6294	0.9485	0.0831	0.3973	0.8343
4	rs12498742	9944052	***SLC2A9***	A/G	1528	**0.4143**	**0.0521**	**1.5×10^{-15}**	**0.0183**	**0.0068**	**0.0072**	-0.0247	0.0401	0.5377	-0.9804	0.6725	0.1451	0.2009	0.4233	0.6352
4	rs2231142	89052323	***ABCG2***	T/G	1532	**0.2307**	**0.0885**	**0.0092**	0.0086	0.0114	0.4499	0.0056	0.0676	0.9341	0.7892	1.124	0.4829	0.046	0.7076	0.9482
6	rs675209	7102084	***RREB1***	T/C	1528	0.0081	0.06	0.8922	0.0052	0.0077	0.497	-0.0496	0.0456	0.2775	-0.738	0.7585	0.3308	-0.1563	0.4784	0.7439
6	rs1165151	25821616	***SLC17A1***	G/T	1532	0.0805	0.0484	0.0964	-0.0002	0.0062	0.971	0.0516	0.0368	0.1606	-0.5442	0.6122	0.3742	-0.3083	0.3861	0.4248
6	rs729761	43804571	***VEGFA***	G/T	1533	**0.1193**	**0.0548**	**0.0296**	0.013	0.0071	0.065	-0.0317	0.0418	0.4481	0.8882	0.6946	0.2012	0.8237	0.4376	0.06
7	rs1178977	72857049	***MLXIPL***	A/G	1522	0.0517	0.0671	0.4415	**-0.017**	**0.0086**	**0.0492**	0.0652	0.0507	0.1991	1.149	0.8455	0.1744	0.6654	0.5351	0.2139
7	rs10480300	151406005	***PRKAG2***	T/C	1529	**0.1013**	**0.0513**	**0.0483**	0.0111	0.0066	0.0917	0.0019	0.0391	0.9612	**1.362**	**0.6489**	**0.036**	0.4837	0.4101	0.2384
8	rs17786744	23777006	***STC1***	G/A	1531	0.0046	0.0498	0.9263	0.0027	0.0064	0.6668	-0.0022	0.0379	0.9534	-0.5857	0.6297	0.3524	-0.4674	0.397	0.2392
8	rs2941484	76478768	***HNF4G***	T/C	1532	0.0101	0.0475	0.8315	0.009	0.0061	0.1403	-0.0104	0.0362	0.7739	-1.102	0.6	0.0665	**-0.8603**	**0.3783**	**0.0231**
10	rs10821905	52646093	***ASAH2***	A/G	1532	0.0628	0.0625	0.3154	0.0039	0.008	0.63	**-0.1091**	**0.0475**	**0.0219**	0.1774	0.7905	0.8224	0.5291	0.4999	0.29
10	rs1171614	61469538	***SLC16A9***	C/T	1524	**0.1896**	**0.0544**	**0.0005**	0.0004	0.007	0.9569	0.0467	0.0415	0.2604	-1.191	0.6887	0.084	-0.2063	0.435	0.6354
11	rs2078267	64334114	***SLC22A11***	C/T	1532	0.0567	0.048	0.2382	0.0138	0.0062	**0.0253**	0.0476	0.0365	0.1922	0.6785	0.6066	0.2636	0.1815	0.3829	0.6356
11	rs478607	64478063	***NRXN2***	A/G	1533	0.0218	0.0691	0.7528	-0.0021	0.0089	0.8108	0.0697	0.0524	0.1834	0.7669	0.874	0.3804	-0.5264	0.5502	0.3389
12	rs3741414	57844049	***INHBC***	T/C	1531	0.0184	0.0672	0.7842	0.0143	0.0086	0.0964	0.0403	0.0512	0.4317	-0.5383	0.8496	0.5264	-0.1109	0.536	0.8362
12	rs653178	112007756	***PTPN11***	C/T	1530	0.085	0.0478	0.0752	0.0037	0.0061	0.5493	-0.0342	0.0364	0.3481	0.8803	0.6053	0.146	**0.8194**	**0.382**	**0.0321**
15	rs6598541	99271135	***MAF***	A/G	1527	0.0887	0.0493	0.0719	-0.0086	0.0063	0.1726	0.0586	0.0375	0.1187	-0.1375	0.6241	0.8256	-0.1326	0.3937	0.7363
16	rs7188445	79734987	***C17orf82***	G/A	1529	0.0047	0.05	0.9257	-0.0078	0.0064	0.2232	0.0126	0.038	0.7398	-0.2134	0.6328	0.736	-0.0256	0.399	0.9489

17	rs7224610	53364788	***IGF1R***	A/C	1531	0.028	0.0472	0.553	0.0121	0.0061	**0.0462**	-0.0058	0.0359	0.8722	0.5462	0.5977	0.3609	0.3186	0.3769	0.3981
17	rs2079742	59465697	***HLF***	C/T	1527	0.0639	0.063	0.3101	0.0149	0.0081	0.0657	0.0159	0.048	0.7398	0.6895	0.7967	0.3869	**-1.129**	**0.5016**	**0.0245**

*Table 11.Association of 24 Uric Acid associated SNPs to Uric Acid and CVD related quantitative traits. Abbreviations: CREA – Creatinine, BMI – Body Mass Index, SBP – Systolic Blood Pressure, DBP – Diastolic Blood Pressure. P values in **black** bold show significantly associated SNPs before Bonferroni Correction, P values in **red** bold show SNPs that survive after multiple testing corrections.*

					CAD		HTN		HC		DB	
CHR	SNP	BP	***Locus Name***	Alleles (effect/other)	OR (95% Confidence Intervals)	P Value	OR (95% Confidence Intervals)	P Value	OR (95% Confidence Intervals)	P Value	OR (95% Confidence Intervals)	P Value
1	rs1471633	145723739	***PDZK1***	A/C	1.3729 (0.8467 - 2.2272)	0.1973	0.9852 (0.8496 - 1.1413)	0.8395	1.15 (0.9862 - 1.3414)	0.0753	1.1261 (0.7924 - 1.6)	0.5072
1	rs11264341	155151493	***TRIM46***	C/T	0.7645 (0.4808 - 1.2164)	0.256	1.0397 (0.8961 - 1.2058)	0.607	0.9346 (0.8006 - 1.0912)	0.3931	0.9276 (0.654 - 1.3149)	0.6725
2	rs1260326	27730940	***GCKR***	C/T	0.9932 (0.6252 - 1.578)	0.977	0.9656 (0.835 - 1.116)	0.6361	0.9195 (0.7904 - 1.07)	0.2768	1.099 (0.7809 - 1.547)	0.5875
2	rs2307394	148716428	***ORC4***	C/T	0.9785 (0.5988 - 1.599)	0.931	0.88 (0.7554 - 1.025)	0.1009	0.9385 (0.7997 - 1.101)	0.4369	**0.6606 (0.4473 - 0.9756)**	**0.036**
3	rs6770152	53100214	***MUSTN1***	G/T	**0.5371 (0.3199 - 0.9018)**	**0.017**	0.9998 (0.8624 - 1.159)	0.9984	1.124 (0.9639 - 1.311)	0.1357	1.21 (0.8579 - 1.706)	0.2767
4	rs12498742	9944052	***SLC2A9***	A/G	0.7446 (0.4572 - 1.212)	0.2337	0.9699 (0.8278 - 1.1366)	0.7069	**0.78 (0.6618 - 0.9191)**	**0.003**	1.0898 (0.7457 - 1.5924)	0.6568
4	rs2231142	89052323	***ABCG2***	T/G	1.445 (0.6852 - 3.046)	0.3312	0.9579 (0.7323 - 1.253)	0.7537	1.022 (0.7688 - 1.357)	0.8834	0.919 (0.4759 - 1.775)	0.8014
6	rs675209	7102084	***RREB1***	T/C	**1.92 (1.16 - 3.177)**	**0.0098**	0.9657 (0.806 - 1.157)	0.7051	1.017 (0.8419 - 1.229)	0.8595	0.8402 (0.5359 - 1.317)	0.4476
6	rs1165151	25821616	***SLC17A1***	G/T	**1.742 (1.086 - 2.793)**	**0.0198**	0.9671 (0.8365 - 1.118)	0.6511	1.098 (0.9443 - 1.278)	0.2238	0.9321 (0.6615 - 1.314)	0.688
6	rs729761	43804571	***VEGFA***	G/T	0.7722 (0.4615 - 1.2928)	0.3244	1.0371 (0.8764 - 1.2277)	0.6723	1.1689 (0.9785 - 1.3961)	0.0848	**1.9201 (1.1862 - 3.1085)**	**0.007**
7	rs1178977	72857049	***MLXIPL***	A/G	0.9542 (0.5102 - 1.7841)	0.8829	1.1534 (0.9452 - 1.4073)	0.1596	0.993 (0.8058 - 1.2244)	0.9488	0.9407 (0.5893 - 1.502)	0.798
7	rs10480300	151406005	***PRKAG2***	T/C	1.025 (0.6226 - 1.687)	0.9231	1.087 (0.9284 - 1.273)	0.3	1.041 (0.8842 - 1.225)	0.6304	1.104 (0.7663 - 1.59)	0.5959
8	rs17786744	23777006	***STC1***	G/A	1.067 (0.6701 - 1.7)	0.7836	0.9612 (0.8298 - 1.113)	0.5974	**0.8133 (0.6974 - 0.9484)**	**0.0084**	0.9239 (0.6517 - 1.31)	0.6564
8	rs2941484	76478768	***HNF4G***	T/C	1.01 (0.6341 - 1.609)	0.9661	0.9705 (0.8385 - 1.123)	0.6877	0.9253 (0.7944 - 1.078)	0.3178	1.375 (0.977 - 1.936)	0.0667
10	rs10821905	52646093	***ASAH2***	A/G	**1.84 (1.092 - 3.101)**	**0.0202**	1.034 (0.8522 - 1.253)	0.7376	1.068 (0.8739 - 1.305)	0.5211	0.9175 (0.5756 - 1.462)	0.7174
10	rs1171614	61469538	***SLC16A9***	C/T	0.8937 (0.5313 - 1.5026)	0.6711	1.0004 (0.846 - 1.1829)	0.996	1.1352 (0.9533 - 1.3521)	0.1552	0.8368 (0.5718 - 1.2237)	0.357
11	rs2078267	64334114	***SLC22A11***	C/T	1.3609 (0.8453 - 2.1915)	0.2033	1.1067 (0.9569 - 1.2794)	0.1708	0.9551 (0.821 - 1.1104)	0.5473	1.0867 (0.771 - 1.5314)	0.6348
11	rs478607	64478063	***NRXN2***	A/G	0.7593 (0.4137 - 1.3945)	0.3735	1.0065 (0.817 - 1.2399)	0.9511	0.9756 (0.7843 - 1.2133)	0.8233	**2.1119 (1.1001 - 4.0552)**	**0.0216**
12	rs3741414	57844049	***INHBC***	T/C	1.202 (0.6552 - 2.206)	0.5515	0.9351 (0.7641 - 1.144)	0.5152	0.8857 (0.7152 - 1.097)	0.2658	0.7937 (0.4728 - 1.332)	0.3811
12	rs653178	112007756	***PTPN11***	C/T	1.002 (0.6318 - 1.59)	0.9921	1.108 (0.9583 - 1.281)	0.166	1.019 (0.876 - 1.185)	0.8087	0.9733 (0.6913 - 1.37)	0.8767
15	rs6598541	99271135	***MAF***	A/G	0.6315 (0.3839 - 1.039)	0.068	0.9352 (0.8071 - 1.084)	0.3723	0.9181 (0.7871 - 1.071)	0.2762	1.105 (0.7829 - 1.561)	0.5692
16	rs7188445	79734987	***C17orf82***	G/A	1.2478 (0.7348 - 2.1195)	0.412	0.956 (0.817 - 1.1187)	0.5748	0.9785 (0.8306 - 1.153)	0.7967	0.9461 (0.6553 - 1.3654)	0.7671
17	rs7224610	53364788	***IGF1R***	A/C	0.9346 (0.5851 - 1.4914)	0.7759	1.047 (0.9058 - 1.2105)	0.5353	0.9728 (0.8361 - 1.1315)	0.7189	1.0858 (0.7698 - 1.5307)	0.6387
17	rs2079742	59465697	***HLF***	C/T	1.099 (0.6094 - 1.981)	0.7544	0.9878 (0.8164 - 1.195)	0.8992	**0.7859 (0.6422 - 0.9618)**	**0.0192**	1.165 (0.7564 - 1.794)	0.4879

Table 12. Association of 24 Uric Acid associated SNPs to CVD related binary traits.Abbreviations: CAD – Coronary Artery Disease, HTN – Hypertension, HC – Hyprecholestereamia, DB - Diabetes. P values in ***black*** *bold show significantly associated SNPs before Bonferroni Correction, P values in* ***red*** *bold show SNPs that survive after multiple testing corrections*

5.4.6-Genetic Risk Score association to Uric Acid and CVD related traits – Mendelian Randomization

We identified at least nominally significant causal estimates for the effect of uric acid on creatinine (p = 0.0017) (*Table 13*) For other quantitative and binary traits we haven't observed any significant causal relationship (*Table 13*)

The IV estimators for creatinine show a causal effect of higher uric acid levels on the levels of creatinine (β = 0.072 (0.034 – 0.11); SE = 0.023).

				IV Estimator		
GRS Traits	â (95% C.I)	SE	*P-Value*	â (95% C.I)	SE	*P-Value*
CREA	0.0046 (0.0022-0.0071)	0.0012	0.0002	**0.0727 (0.0341-0.1112)**	**0.0232**	**0.0017**
BMI	0.0183 (-0.0453-0.0819)	0.0324	0.572	0.2844 (-0.7038-1.2727)	0.5061	0.574
DBP	0.1114 (-0.0476-0.2705)	0.0811	0.17	1.7309 (-0.7408-4.2028)	1.294	0.181
SBP	0.1424 (-0.1443-0.4292)	0.1462	0.33	2.2122 (-2.243-6.6682)	2.302	0.3365
				IV Estimator		
GRS Phenotypes	OR (95% C.I)	SE	*P-Value*	OR (95% C.I)	SE	*P-Value*
CAD	1.0249 (0.9378-1.131)	0.0483	0.6101	0.961 (0.8794-1.0605)	0.0463	0.6114
DB	1.05336 (0.9823-1.1344)	0.0369	0.159	0.9876 (0.921-1.0637)	0.036	0.1706
HC	0.9958 (0.9667-1.0258)	0.0151	0.781	0.9337 (0.9064-0.9619)	0.0171	0.7812
HTN	1.0284 (0.9997-1.0581)	0.0144	0.0524	**0.9643 (0.9374-0.9922)**	**0.0167**	**0.0654**[a]

Table 13.Mendelian Randomization results on effect of Serum Uric Acid on CVD related

5.4.7 – Discussion

We attempted to find a causal relationship between SUA and CVD/CVD related traits, a relationship that has been previously postulated in many observational studies, by implementing a Mendelian randomization approach and using genetic variants associated with SUA levels as our instrumental variable. Initially, through an extensive literature search, we identified known genetic variants associated with SUA levels. We devised a set of criteria to select 24 genetic variants, from the complete list of researched variants, to efficiently and effectively allow greater power to detect any causal relationship between SUA and CVD. In the past the use of a single instrumental variable in MR analysis was considered appropriate however, due to the fact that the use of genetic variants, which impart very small overall risk, requires multiple instrumental variables to be used to increase statistical association to the intermediate phenotype. This increases the likelihood of finding a true causal association to the outcome of interest, and reduces some of the limitations that inherently companion a MR analysis such as LD. Hence, careful selection of independent SNPs overcomes any unfavorable LD between IVs and also allows inclusion of genetic variants that might be in close vicinity, or within a locus that is already associated with CVD or a CVD related trait.

6. A view towards the future, beginning from the past

6.1 Genomic correlates of atherosclerosis in Ancient Humans

Twenty-first century humans often think that much of the heart disease we suffer is the consequence of our modern lifestyle. We live sedentary, mentally stressful lives, consume large amounts of calories from ill chosen foods, indulge sugar and salt in high quantities, and are exposed to high levels of environmental toxins, including tobacco smoke. To avoid dying of a heart attack, we are urged to engage in regular exercise; maintain an ideal body weight; avoid unsaturated fat, salt, and sugar; avoid smoke; and take medicine to keep our blood pressure, lipid levels, and blood sugar in an ideal range. Perhaps we think nostalgically of our ancestors who lived a purer life and did not have to engage in corrective actions to prevent heart disease.

Our ancestors going back thousands of years show signs of atherosclerosis, as suggested by modern research using CT to detect evidence of calcium deposits associated with atherosclerotic plaques in the arteries of mummies as old as 5,000 years. Even though our human ancestors lived far different lives than we do, their environments and lifestyles were not protecting them against the development of atherosclerosis.

What is similar between now and then is the human genetic material, our genome, including ancient polymorphisms that were uncovered to predispose the carrier to the development of atherosclerotic cardiovascular disease. Seminal studies of genetic material obtained from 5,300-year-old Ötzi showed mutations in the 9p21 chromosomal region identical to SNP identified in contemporary humans that are strong predictors for the development of CAD. Our ancient ancestors were certainly susceptible to many other conditions, such as infectious diseases, nutritional deprivation, and trauma, which often resulted in an early age death, before atherosclerotic heart disease became clinically manifest.

Nevertheless, the study of ancient humans and the investigation of the interaction

between environmental and genetic influences on the development of heart disease may provide unique pathophysiologic insights and lead to more effective prevention and treatment of the most common cause of death in the modern world.

Paleogenetics offers a unique opportunity to study human evolution, population dynamics, and disease evolution in situ. Although histologic and computed x-ray tomographic investigations of ancient mummies have clearly shown that atherosclerosis has been present in humans for more than 5,000 years, limited data are available on the presence of genetic predisposition for cardiovascular disease in ancient human populations. In a previous whole-genome study of the Tyrolean Iceman, a 5,300 year old glacier mummy from the Alps, an increased risk for coronary heart disease was detected. The Iceman's genome revealed several single nucleotide polymorphisms that are linked with cardiovascular disease in genome-wide association studies. Future genetic studies of ancient humans from various geographic origins and time periods have the potential to provide more insights into the presence and possible changes of genetic risk factors in our ancestors.

6.2 1000 years of genetic variability in Brisighella's population: survey on the predisposition to cardiovascular diseases from the XIst century today

Genetic analysis of biological samples of the past allows to explore fascinating new Research horizons through the ability to analyze the genomes of individuals who lived in distant times. This area of research has many possible applications, which are based on the analysis of ancient DNA: evolutionary studies, identification of infectious or genetic diseases, staff identification, definition of family relationships between individuals and determination of characteristics or phenotypic markers related to specific diseases, as well as a reconstruction of the diet and genetic characterization of human groups in the past.

The success of these studies suggests that the application of the latest technologies in the field of paleogenetics can bring significant contributions in the study of

preparation/genetic resistance to diseases with high incidence in the present population, among which cardiovascular ones surely play an important role. In particular, the knowledge of any changes, during time, of the frequencies of the markers correlated with the above pathologies and possible diachronic changes in DNA induced by epigenetic phenomena, could serve as an appropriate understanding of the pathogenesis of these diseases and their prevention.

However, despite some studies based on histological and computed tomographic analysis have shown the presence of atherosclerosis in populations dating back over 5000 year [129,130]there are few data regarding the preparation to cardiovascular diseases in ancient human populations.

Currently, in one individual of the past, the man of Similaun, dating back to about 5300 years, the presence of polymorphisms (ie single nucleotide genetic modifications) related to the risk of coronary diseases[131,132]has been investigated.

Among the polymorphisms related to the risk of cardiovascular diseases, the chromosomal region 9p21, already implicated in the pathogenesis of cancer[133], was the subject of several studies in 2007,independent on the human genome that have reported the association of the same locus on chromosome 9p21 with coronary artery disease and myocardial infarction[134,135].

This discovery has led to a considerable increase of further investigations that have largely confirmed the preliminary data and the identification of other genetic variations closely associated with cardiovascular events [136].Alongside these studies, recent research gives a glimpse of the role that epigenetic DNA modifications can have in transforming environmental stimuli inheritable genetic expressions can influence significantly, even in later generations, the susceptibility to diseases. In particular, as well as in malignancies[137], neurodegenerative[138] and metabolic diseases[139], epimutations of DNA (for example of hyper- or hypomethylation processes) appear to have a0 significant role in the pathogenesis of atherosclerosis and cardiovascular related disease[140].

As for individuals of the past , the human genome of the Iceman has been entirely sequenced in 2012 and the analysis of the data obtained showed that this individual had a high genetic predisposition to cardiovascular inauspicious events, such as highlighted by a study of some genetic polymorphisms associated with cardiovascular risk previously identified in genome-wide association studies ().

Chr.	dbSNP#	1,000 Genomes minor allele frequency	Gene	Coverage Iceman A	C	G	T	N	SNP association
chr1	rs1801133	A = 0.325	*MTHFR*	10	0	1	0	0	Cardiovascular disease
chr1	rs1764391	T = 0.354	*GJA4*	0	4	0	9	0	Atherosclerosis
chr4	rs1870377	A = 0.241	*KDR*	8	0	0	9	0	Coronary heart disease
chr9	rs10757274	G = 0.396	*CDKNBAS*	1	0	8	0	0	Ischemic stroke, sudden cardiac death
chr9	rs2383206	G = 0.459	*CDKNBAS*	0	0	8	0	0	Coronary heart disease
chr13	rs5351	T = 0.436	*EDNRB*	0	14	0	20	1	Atherosclerosis
chr19	rs1613662	G = 0.141	*GP6*	4	0	3	0	0	Myocardial infarction, age

chr, chromosome; dbSNP, single nucleotide polymorphism database; SNP, single nucleotide polymorphisms.

Table 14. SNP related to cardiovascular disease identified in the Iceman genome

However, as far as we know, there are no systematic studies, at level of ancient people, in particular the second at diachronic perspective, both as regards genetic variants, and associated with epimutations in cardiovascular lesions.

With a view to laying the foundations for an integrated project of biology, genetics and medicine inhuman field , the following project has as main objective the analysis of the preparation of genetic cardiovascular diseases in the population of Brisighella (RA) during a span of about 1000 years.

The peculiarity of this research project lies in the ability to conduct a study on genetic markers related to cardiovascular diseases on the territory of Brisighella, with the extraordinary opportunity to expand the chronological horizon of the already established DNA biobank.

In order to study the distribution of polymorphisms related to cardiovascular disease in the people of Brisighella from the ninth century to the present and correlate them with lifestyle and diet, it aims to address the following analysis :

1) verifying genetic continuity between the finding buried in the castle and the current Rontana people of Brisighella , previously sampled and which constitute the current DNA biobank (Brisighella Heart Study Biobank) . The analysis will be carried out through the study of markers uniparental (hypervariable region 1 - HVR1 of mitochondrial DNA polymorphisms Y chromosome) that allow to identify and follow the female parental lines and men's long subsequent generations;

2) analysing the presence of genetic variants related to cardiovascular disease identified in the association studies published in the literature in the ancient and modern population of Brisighella ;

3) reconstruction of the style of life and state of health of ancient populations through anthropological and paleopathological study;

4) reconstruction of the ancient population diet through stable carbon and nitrogen isotope analysis;

5) extrapolation of data on lifestyle and diet for selected modern standards in the biobank ;

The study of ancient humans and a better understanding of the interaction between environmental and genetic influences on the development of heart diseases may lead to a more effective prevention and treatment of the most common cause of death in the modern world.

Bibliography

[1]**Paul, J. R.** " Epidemiology " in Green, David E.,and Knox, W. Eugene (Eds.), Research in MedicalScience. New York: Macmillan, 1950. p. 53.2. Maxcy, Kenneth F. (Ed.), Papers of Wade Hampton Frost, M.D. New York: Commonwealth Fund,1941. p.1

[2]**Gaddi A**., A. F. G. Cicero, F. O. Wani, A. Dormi, V. Pasquarelli, and S. D'Addato, "The realization of a project aimed at reducing the plasmatic lipid level in a large Italian population improves the mean calcium daily intake: the Brisighella study," *European Journal of Clinical Nutrition*, vol. 55, no. 2, pp. 97–106, 2001.

[3]**A F. G. Cicero**, A. Dormi, S. D'Addato, A. V. Gaddi, and C. Borghi, "Long-term effect of a dietary education program on postmenopausal cardiovascular risk and metabolic syndrome: the brisighella heart study," *Journal of Women's Health*, vol. 19, no. 1, pp. 133–137, 2010.

[4]Framinghan Heart Study Database

[5]**Holle R**[1], Happich M, Löwel H, Wichmann HE; MONICA/KORA Study Group. KORA--a research platform for population based health research.*Gesundheitswesen*. 2005 Aug;67 Suppl 1:S19-25.

[6]**Pattaro C**[1], Gögele M[2], Mascalzoni D[2] et al. The Cooperative Health Research in South Tyrol (CHRIS) study: rationale, objectives, and preliminary results.*J Transl Med.* 2015 Nov 5;13:348. doi: 10.1186/s12967

[7]**Makarow M**., Hoejgaard L., Vanommen, G. J. &Skorpen, F. European Science Foundation (2008) "Population Surveys and Biobanking. Science Policy

Briefing".BBMRI_Master_Galley.indd 81 2013/06/05 19:35 82 European Science Foundation 32: 1-10

[8]**Carey MC et al.,** Lipid digestion and absorption. *Annu Rev Physiol* 45, 651-677, 1983

[9]**Bisgaier CL et al.,** Intestinal synthesis, secretion and transport of lipoproteins *AnnuRev Physiol* 45, 625-536, 1983.

[10]**Kane JP**, Apo-lipoprotein B: structural and metabolic heterogeneity. *Annu Rev Physiol*45, 637-650, 1983

[11]**Breckenridge WC et al.,** Hypertriglyceridemia associated with deficiency of apolipoprotein CII. *N Engl J Med* 298, 1265-1273, 1978.

[12]**Brunzell JD**, Familial lipoprotein lipase deficiency and other causes of chylomicronemia syndrome. In: Scriver CR Et al., eds. the metabolic basis of inherited disease. 6th ed. New York: McGraw-Hill, 1165-1180, 1989.

[13]**Demel RA and De Kruyff B**. The function of sterols in membranes. *BiochimBiophysActa*457, 109-132, 1976.

[14]**Grundy SM.,**Cholesterol metabolism in man. *West J Med* 128, 13-25, 1978

[15]**Mahley RW.,**Lipoprotein receptor and cholesterol homeostasis. *BiochimBiophysActa*737, 197-222, 1983.

[16]**Nakanishi M. et al.,** Multivalent control of 3-hydroxy-3-methylglutaryl coenzyme Areductase. *J BiolChem* 263, 8929-8937, 1988.

[17]**Beg ZH. et al.,** Modulation of enzymatic activity of 3-hydroxy-3-methylglutaryl

coenzyme A reductase by multiple kinase systems involving reversible phosphorylation: a review. *Metabolism* 36, 900-917, 1987.

[18]**Bjorkhem I.M**., Mechanism of bile acid biosynthesis in mammalian liver (review).In: Danielsson H, Sjovall J, eds. Comprehensive Biochemistry. Amsterdam: Elsevier, 231-278, 1985.

*[19]***Goldstein JL et al.,** Familial hypercholesterolemia. In: Scriver CR, Beaudet AL, SlyWS, et al., eds. The metabolic basis of inherited disease. 6th ed. New York: McGraw-Hil, 1215-1250, 1989.

[20]**Kovanen PT et al.,** Regulation of low density lipoprotein receptors by adrenocorticotropin in the adrenal gland of the mice and rats in vivo. *J BiolChem* 255,5591-5598, 1980.

[21]**Von Eckardstein A. et al.,** Physiological role and clinical relevance of high-density lipoprotein subclasses. *CurrOpin Lipidol*.5, 404–416, 1994.

[22]**Brown MS. And Goldstein JL.,** Lipoprotein metabolism in the macrophage: implications for cholesterol deposition in atherosclerosis. *Annu Rev Biochem* 52, 223-261, 1983.

[23]**Oram JF. et al.,** Regulation of high density lipoprotein receptor activity in culturedhuman skin fibroblast and human arterial smooth muscle cells. *J Clin Invest* 72, 1611-1621, 1983.

[24]**Genest J. et al.,** High density lipoproteins in health and disease. *J Invest Med*. 47, 31–42, 1999.

[25]**Von Eckardstein A. and Assmann G.,** Prevention of coronary heart disease by raising of high density lipoprotein cholesterol *CurrOpin Lipidol*.11, 627–637, 2000.

[26]**Dietschy JM. et al.,** Role of liver in the maintenance of cholesterol and low density lipoprotein homeostasis in different animal species including humans. *J Lipid Res.* 34,1637–1639, 1993.

[27]**Phillips MC. et al.,** Mechanism and consequence of cellular cholesterol exchange and transfer. *BiochimBiophys Acta*.906, 223–276, 1987.

[28]**Krieger M.,** Charting the fate of the "good cholesterol": identification and characterization of the high-density lipoprotein receptor SR-BI. *Annu Rev Biochem*.68,523–558, 1999.

[29]**Williams DL. et al.,** Scavenger receptor BI and cholesterol trafficking. *CurrOpin Lipidol.*10, 329–339, 1999.

[30]**Thuahnai ST. et al.,** Scavenger receptor class B, type-I-mediated uptake of various lipids into cells. *J Biol Chem*. 276, 43801–43808, 2001.

[31]**Rigotti A. et al., .**The class B scavenger receptors SR-BI and CD36 are receptors for anionic phospholipids. *J Biol Chem*. 270, 16221–16224, 1995.

[32]**Bortnick AE, et al.,** The correlation of ABC1 mRNA levels with cholesterol efflux from various cell lines. *J Biol Chem*. 275, 28634–28640, 2000.

[33]**Remaley AT, et al.,** Apo-lipoprotein specificity for lipid efflux by the human ABCA Itransporter. *BiochemBiophys Res Commun*.280, 818–823, 2001.

[34]**Steinberg D., et al.,** Beyond cholesterol: modifications of low-density lipoprotein that increase its atherogenicity. *N Engl J Med*. 320, 915-924, 1989.

[35]**Brown MS. And Goldstein JL.,** Drug used in the treatment of hyperlipidaemias. In:Goodman LS, Gilman A, Rall TW, Nies AS, Tsylor P, eds. Pharmacological bases of therapeutics. New York: Pergamon Press, 874-896, 1990.

[36]**Grundy SM, et al.,** Summary of the second report of the national cholesterol education programme (NCEP)expert panel of detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). *JAMA* 269, 3015-3023, 1993.

[37]**Faggiotto et al.,** Studies of hypercholesterolemia in non-humans primate: I. Changes that lead to fatty streak formation. *Arteriosclerosis* 4:323, 1984.

[38]**Brown B.G., et al.,** Familial atherosclerosis treatment study: A randomized trial demonstrating coronary disease regression and clinical benefit from lipid-altering therapy among men with high apo-lipoprotein-B. *N. Engl. J. Med*.

[39]**Miki, H. et al.,** Analysis of the kinesin superfamily: insights into structure and function. *Trends Cell Biol*. 15, 467–476, 2005.

[40]**Vale, R.D.** The molecular motor toolbox for intracellular transport. *Cell* 112, 467–480,2003.

[41]**Bananis, E. et al.,** Microtubule and motor-dependent endocytic vesicle sorting in vitro. *J. Cell Biol*. 151, 179–186, 2000.

[42]**Hameed A, et al.** No Evidence for Cardiac Dysfunction in Kif6 Mutant Mice. PLoSONE 8(1): e54636, 2013.

[43]**Levy R.I.,** Drugs used in the treatment of hyperlipoproteinaemias. In Goodman, A.S.,Gilman, A.S., and Gilman, A., (eds.): The Pharmacologic bases of therapeutics. NewYork, Macmillan, 834-877, 1980.

[44]**Sitori C.R. and Chiero G.,** Effects of lipid lowering agents and other treatment regimens on serum lipoprotein. *Current Opinion in Lipidology* 1:262, 1990.

[45]**Chatzizisis Y. S, et al.,** Risk factors and drug interactions predisposing to statin-induced myopathy: implications for risk assessment, prevention and treatment. *DrugSaf*. 33(3),171–187, 2010.

[46]**Giorgi MA,** Pharmacogenomics and adverse drug reactions: the case of statins. *Expert OpinPharmacother* 12(10):1499–509, 2011.

[47]**Kalow W and Gunn DR.,** Some statistical data on atypical cholinesterase of human serum. Ann Hum Genet 23, 239 –250, 1959; Evans DA et al., Genetic control of isoniazid metabolism in man.*Br Med J* 2, 485-491, 1960.

[48]**Meyer UA.,**Pharmacogenetics and adverse drug reactions. Lancet 356, 1667–1671,2000.

[49]**Evans WE and Relling MV,**Moving towards individualized medicine with pharmacogenomics. *Nature* 429, 464 – 468, 2004.

[50]**Qiang Ma and Anthony Y. H. Lu.,**Pharmacogenetics, Pharmacogenomics, and Individualized Medicine, *Pharmacol Rev* 63:437–459, 2011.

[51]**Ravi Sachidanandam, et al.,** A map of human genome sequence variation containing1.42 million single nucleotide polymorphisms. *Nature*vol 409, 2001.125

[52]**Vesell, E. S.,** Pharmacogenetics perspectives gained from twin and family studies. *Pharmacol. Ther*.41, 535–552, 1989.

[53]**Guengerich, F. P. et al.,** Twenty years of biochemistry of human P450s: purification,expression, mechanism, and relevance to drugs. *Drug Metab.Dispos*. 26, 1175–1178,1998.

[54]**Meyer U. A. and Zanger, U. M.,** Molecular mechanisms of genetic polymorphisms of drug metabolism *Annu. Rev. Pharmacol. Toxicol*.37, 269–296, 1997.

[55]**Gonzalez, F. J. et al.,** Characterization of the common genetic defect in humans deficient in debrisoquine metabolism. *Nature* 331, 442–446, 1988.

[56]**Collins FS et al.,** A DNA polymorphism discovery resource for research on humangenetic variation. *Genome Res* 8(12), 1229-1231, 1998.

[57]**Sim Sc. and Ingelman-Sundberg M.,** Pharmacogenomics biomarkers: new tools incurrent and future drug therapy. *Trends PharmacolSci*, 2010.

[58]**Taylor F**, Ward K, Moore TH, Burke M, Davey Smith G, Casas JP et al. Statins for the primary prevention of cardiovascular disease. Cochrane Database Syst Rev 2011; : CD004816.

[59]**Phan BAP**, Bittner V. Lipid-Lowering Therapy in Patients 75 Years and Older: Clinical Priority or Superfluous Therapy? Prog.Cardiovasc.Dis. 2014. doi:10.1016/j.pcad.2014.03.005.

[60]WHO; WORLD HEART FEDERATION; Global atlas on cardiovascular disease prevention and control. In: Global atlas on cardiovascular disease prevention and control, Policies, strategies and interventions. Geneva, 2011, p 164.

[61]**Do R**, Willer CJ, Schmidt EM, Sengupta S, Gao C, Peloso GM et al. Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet* 2013; 45: 1345–52.

[62]**Mitchell AP**, Simpson RJ. Statin cost effectiveness in primary prevention: A systematic review of the recent cost-effectiveness literature in the United States. BMC Res. Notes.2012; 5: 373.

[63]**Hopewell JC**, Parish S, Offer A, Link E, Clarke R, Lathrop M et al. Impact of common genetic variation on response to simvastatin therapy among 18 705 participants in the Heart Protection Study. *Eur Heart J* 2013; 34: 982–992.

[64]**Chasman DI**, Giulianini F, MacFadyen J, Barratt BJ, Nyberg F, Ridker PM. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: The justification for the use of statins in prevention: An intervention trial evaluating rosuvastatin (JUPITER) trial. *Circ Cardiovasc Genet* 2012; 5: 257–264.

[65]**Poduri A**, Khullar M, Bahl A, Sehrawat BS, Sharma Y, Talwar KK. Common variants of HMGCR, CETP, APOAI, ABCB1, CYP3A4, and CYP7A1 genes as predictors of lipid-lowering response to atorvastatin therapy.*DNA Cell Biol* 2010; 29: 629–637.

[66]**Polisecki E**, Peter I, Simon JS, Hegele RA, Robertson M, Ford I et al. Genetic variation at the NPC1L1 gene locus, plasma lipoproteins, and heart disease risk in the elderly. *J Lipid Res* 2010; 51: 1201–1207.

[67]**Bare LA**, Morrison AC, Rowland CM, Shiffman D, Luke MM, Iakoubova OA et al. Five common gene variants identify elevated genetic risk for coronary heart disease. *Genet Med* 2007; 9: 682–689.

[68]**Iakoubova OA**, Tong CH, Rowland CM, Kirchgessner TG, Young BA, Arellano AR et al. Association of the Trp719Arg Polymorphism in Kinesin-Like Protein 6 With Myocardial Infarction and Coronary Heart Disease in 2 Prospective Trials. The CARE and WOSCOPS Trials. *J Am CollCardiol* 2008; 51: 435–443.

[69]**Iakoubova OA**, Robertson M, Tong CH, Rowland CM, Catanese JJ, Blauw GJ et al. KIF6 Trp719Arg polymorphism and the effect of statin therapy in elderly patients: results from the PROSPER study. *Eur J CardiovascPrevRehabil* 2010; 17: 455–461.

[70]**Li Y**, Iakoubova OA, Shiffman D, Devlin JJ, Forrester JS, Superko HR. KIF6 polymorphism as a predictor of risk of coronary events and of clinical event reduction by statin therapy. *Am J Cardiol* 2010; 106: 994–8.

[71]**Hopewell JC**, Parish S, Clarke R, Armitage J, Bowman L, Hager J et al. No impact of KIF6 Genotype on vascular risk and statin response among 18,348 randomized patients in the heart protection study. *J Am CollCardiol* 2011; 57: 2000–2007.

[72]**Ridker PM**, MacFadyen JG, Glynn RJ, Chasman DI. Kinesin-Like Protein 6 (KIF6) Polymorphism and the Efficacy of Rosuvastatin in Primary Prevention. *Circ Cardiovasc Genet* 2011; 4: 312–317.

[73]**Assimes TL**, Hlm H, Kathiresan S, Reilly MP, Thorleifsson G, Voight BF et al. Lack of Association between the Trp719Arg Polymorphism in Kinesin-Like Protein-6 and Coronary Artery Disease in 19 Case-Control Studies. *J Am CollCardiol* 2010; 56: 1552–1563.

[74]**Krauss RM**, Mangravite LM, Smith JD, Medina MW, Wang D, Guo X et al. Variation in the 3-hydroxyl-3-methylglutaryl coenzyme A reductase gene is associated with racial differences in low-density lipoprotein cholesterol response to simvastatin treatment. *Circulation* 2008; 117: 1537–1544.

[75]**Cicero AFG**, Dormi A, D'Addato S, Borghi C, Staff BHS.From risk factor assessment to cardiovascular disease risk and mortality modification: the first 40 years of the Brisighella Heart Study. *ClinLipidol* 2011; 6: 269–276.

[76]**Cicero AFG**, Rosticci M, Baronio C, Morbini M, Parini A, Grandi E et al. Serum LDL cholesterol levels and new onset of arterial hypertension: an 8-year follow-up. *Eur J Clin Invest* 2014; 44: 926–32.

[77]**Rydén L**, Grant PJ, Anker SD, Berne C, Cosentino F, Danchin N et al. ESC guidelines on diabetes, pre-diabetes, and cardiovascular diseases developed in collaboration with the EASD - summary. *DiabVasc Dis Res* 2014; 11: 133–73.

[78]**Thygesen K**, Alpert JS, Jaffe AS, Simoons ML, Chaitman BR, White HD et al. Third universal definition of myocardial infarction. *Eur Heart J* 2012; 33: 2551–2567.

[79]**Mancia G**, Fagard R, Narkiewicz K, Redon J, Zanchetti A, Böhm M et al. 2013 ESH/ESC Practice Guidelines for the Management of Arterial Hypertension. *Blood Press* 2014; 23: 3–16.

[80]**Willer CJ**, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S et al. Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; 45: 1274–83.

[81]The International Consortium for Blood Pressure Genome-Wide Association Studies. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011; 478: 103–9.

[82]R Core Team. R: A Language and Environment for Statistical Computing. 2014.http://www.r-project.org.

[83]Purcell S. PLINK v1.07 (10-Oct-2009). http://pngu.mgh.harvard.edu/purcell/plink/.

[84]**Purcell S**, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81: 559–575.

[85]**Marchini J**, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; 39: 906–913.

[86]**Mägi R**, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010; 11: 288.

[87]**O'Reilly PF**, Hoggart CJ, Pomyen Y, Calboli FCF, Elliott P, Jarvelin M-R et al. MultiPhen: joint model of multiple phenotypes can increase discovery in GWAS. *PLoS One* 2012; 7: e34861.

[88]**Teslovich TM**, Musunuru K, Smith A V, Edmondson AC, Stylianou IM, Koseki M et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 2010; 466: 707–713.

[90]**Ballantyne CM**, Raichlen JS, Cain VA. Statin therapy alters the relationship between apolipoprotein B and low-density lipoprotein cholesterol and non-high-density lipoprotein cholesterol targets in high-risk patients: the MERCURY II (Measuring Effective Reductions in Cholesterol Using Rosuvas. *J Am CollCardiol* 2008; 52: 626–32.

[91]**Sniderman A**, Furberg C, Keech A, van Lennep JR, Frohlich J, Jungner I et al. Apolipoproteins versus lipids as indices of coronary risk and as targets for statin treatment. *Lancet* 2003; 361: 777–780.

[92]**Barter PJ**, Ballantyne CM, Carmena R, Cabezas MC, Chapman MJ, Couture P et al. Apo B versus cholesterol in estimating cardiovascular risk and in guiding therapy: Report of the thirty-person/ten-country panel. *J. Intern. Med*. 2006; 259: 247–258.

[93]**Charlton-Menys V**, Betteridge DJ, Colhoun H, Fuller J, France M, Hitman GA et al. Targets of statin therapy: LDL cholesterol, non-HDL cholesterol, and apolipoprotein B in type 2 diabetes in the collaborative atorvastatin diabetes study (CARDS). *ClinChem* 2009; 55: 473–480.

[94]**Sniderman AD**, Scantlebury T, Cianflone K. Hypertriglyceridemic hyperapoB: The unappreciated atherogenic dyslipoproteinemia in type 2 diabetes mellitus. *Ann. Intern. Med*. 2001; 135: 447–459.

[95]**Robinson JG**, Wang S, Jacobson TA. Meta-analysis of comparison of effectiveness of lowering apolipoprotein B versus low-density lipoprotein cholesterol and non high-density lipoprotein cholesterol for cardiovascular risk reduction in randomized trials. *Am J Cardiol* 2012; 110: 1468–1476.

[96]**Iakoubova OA**, Tong CH, Rowland CM, Kirchgessner TG, Young BA, Arellano AR *et al.* Association of the Trp719Arg Polymorphism in Kinesin-Like Protein 6 With Myocardial Infarction and Coronary Heart Disease in 2 Prospective Trials. The CARE and WOSCOPS Trials. *J Am Coll Cardiol* 2008; **51**: 435–443.

[97]**Iakoubova OA**, Robertson M, Tong CH, Rowland CM, Catanese JJ, Blauw GJ *et al*. KIF6 Trp719Arg polymorphism and the effect of statin therapy in elderly patients: results from the PROSPER study. *Eur J Cardiovasc Prev Rehabil* 2010; **17**: 455–461.

[98]**Ference BA**, Yoo W, Flack JM, Clarke M. A common KIF6 polymorphism increases vulnerability to low-density lipoprotein cholesterol: Two meta-analyses and a meta-regression analysis. *PLoS One* 2011; 6.doi:10.1371/journal.pone.0028834

[99]**Hopewell JC**, Parish S, Clarke R, Armitage J, Bowman L, Hager J *et al.* No impact of KIF6 Genotype on vascular risk and statin response among 18,348 randomized patients in the heart protection study. *J Am Coll Cardiol* 2011; **57**: 2000–2007.

[100]**Ridker PM**, MacFadyen JG, Glynn RJ, Chasman DI. Kinesin-Like Protein 6 (KIF6) Polymorphism and the Efficacy of Rosuvastatin in Primary Prevention. *Circ Cardiovasc Genet* 2011; **4**: 312–317.

[101]**Assimes TL**, Hlm H, Kathiresan S, Reilly MP, Thorleifsson G, Voight BF *et al.* Lack of Association between the Trp719Arg Polymorphism in Kinesin-Like Protein-6 and Coronary Artery Disease in 19 Case-Control Studies. *J Am Coll Cardiol* 2010; **56**: 1552–1563.

[102]**Willer CJ**, Schmidt EM, Sengupta S, Peloso GM, Gustafsson S, Kanoni S *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* 2013; **45**: 1274–83.

[103]**Teslovich TM**, Musunuru K, Smith A V, Edmondson AC, Stylianou IM, Koseki M *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* 2010; **466**: 707–713.

[104]**Shiffman D**, Sabatine MS, Louie JZ, Kirchgessner TG, Iakoubova OA, Campos H et al. Effect of Pravastatin Therapy on Coronary Events in Carriers of the KIF6 719Arg Allele from the Cholesterol and Recurrent Events Trial. *Am J Cardiol* 2010; 105: 1300–1305.

[105]**Shen X**, Wang W, Wang L, Houde C, Wu W, Tudor M et al. Identification of genes affecting apolipoprotein B secretion following siRNA-mediated gene knockdown in primary human hepatocytes. *Atherosclerosis* 2012; 222: 154–157.

[106]**Mueller JC**, Lõhmussaar E, Mägi R, Remm M, Bettecken T, Lichtner P et al. Linkage disequilibrium patterns and tagSNP transferability among European populations. *Am J Hum Genet* 2005; 76: 387–398.

[107]**Cicero AFG**, Salvi P, D'Addato S, Rosticci M, Borghi C. Association between serum uric acid, hypertension, vascular stiffness and subclinical atherosclerosis: data from the Brisighella Heart Study. *J Hypertens* 2014; 32: 57–64.

[108]**Ruilope, L**. M., &Pontremoli, R. (2013).Serum uric acid and cardio-renal diseases. *Current Medical Research and Opinion*, *29 Suppl 3*, 25–31. doi:10.1185/03007995.2013.790803

[109]**Becker, B**. (1993). Towards the Physiological Function of Uric Acid. *Free Radical Biology and Medicine*, *14*, 615–631.

[110]**Oda, M**., Satta, Y., Takenaka, O., &Takahata, N. (2002).Loss of urate oxidase activity in hominoids and its evolutionary implications. *Molecular Biology and Evolution*, *19*(5), 640–53.

[111]**Heinig, M**., & Johnson, R. (2006).Role of uric acid in hypertension, renal disease, and metabolic syndrome. *Cleveland Clinic Journal of Medicine*, 1059–1064.

[112]**Gagliardi, A. C. M**., Miname, M. H., & Santos, R. D. (2009).Uric acid: A marker of increased cardiovascular risk. *Atherosclerosis*, *202*(1), 11–7.doi:10.1016/j.atherosclerosis.2008.05.022

[113]**Khosla, U. M**., Zharikov, S., Finch, J. L., Nakagawa, T., Roncal, C., Mu, W. et al. (2005). Hyperuricemia induces endothelial dysfunction. *Kidney International*, *67*(5), 1739–42.doi:10.1111/j.1523-1755.2005.00273.x

[114]**Vekic, J**., Jelic-Ivanovic, Z., Spasojevic-Kalimanovska, V., Memon, L., Zeljkovic, A., Bogavac-Stanojevic, N., & Spasic, S. (2009). High serum uric acid and low-grade inflammation are associated with smaller LDL and HDL particles. *Atherosclerosis*, *203*(1), 236–42.doi:10.1016/j.atherosclerosis.2008.05.047

[115]**Feig, D. I**., Soletsky, B., & Johnson, R. J. (2014).Effect of Allopurinol on Blood Pressure Essential Hypertension, *300*(8).

[116]**Köttgen, A.,** Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C.,Gieger, C. (2013). Genome-wide association analyses identify 18 new loci associated with serum urate concentrations. *Nature Genetics*, *45*(2), 145–54.doi:10.1038/ng.2500

[117]**Yang, Q**., Guo, C.-Y., Cupples, L. A., Levy, D., Wilson, P. W. F., & Fox, C. S. (2005). Genome-wide search for genes affecting serum uric acid levels: the Framingham Heart Study. *Metabolism: Clinical and Experimental*, *54*(11), 1435–41. doi:10.1016/j.metabol.2005.05.007

[118]**Kamatani, Y.,** Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., Kamatani, N. (2010). Genome-wide association study of hematological and biochemical traits in a Japanese population. *Nature Genetics*, *42*(3), 210–5.doi:10.1038/ng.531

[119]**Kolz, M**., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M.,Gieger, C. (2009). Meta-analysis of 28,141 individuals identifies common variants within five

new loci that influence uric acid concentrations. *PLoS Genetics*, *5*(6), e1000504. doi:10.1371/journal.pgen.1000504

[120]**Yang, B**., Mo, Z., Wu, C., Yang, H., Yang, X., He, Y.,He, M. (2014). A genome-wide association study identifies common variants influencing serum uric acid concentrations in a Chinese population. *BMC Medical Genomics*, *7*(1), 10.doi:10.1186/1755-8794-7-10

[121]**Vitart, V.,**Rudan, I., Hayward, C., Gray, N. K., Floyd, J., Palmer, C. N. a, … Wright, A. F. (2008). SLC2A9 is a newly identified urate transporter influencing serum urate concentration, urate excretion and gout. *Nature Genetics*, *40*(4), 437–42.doi:10.1038/ng.106

[122]**Woodward, O. M**., Köttgen, A., Coresh, J., Boerwinkle, E., Guggino, W. B., & Köttgen, M. (2009).Identification of a urate transporter, ABCG2, with a common functional polymorphism causing gout. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(25), 10338–42.doi:10.1073/pnas.0901249106

[123]**Taylor, A. E**., Davies, N. M., Ware, J. J., Vanderweele, T., Smith, G. D., & Munafò, M. R. (2014). Mendelian randomization in health research: Using appropriate genetic variants and avoiding biased estimates. *Economics and Human Biology*, *13*, 99–106. doi:10.1016/j.ehb.2013.12.002

[124]**Smith, G. D**., & Ebrahim, S. (2004). Mendelian randomization: prospects, potentials, and limitations. *International Journal of Epidemiology*, *33*(1), 30–42. doi:10.1093/ije/dyh132

[125]**Davey Smith, G.,**& Ebrahim, S. (2003). "Mendelian randomization": can genetic epidemiology contribute to understanding environmental determinants of disease? *International Journal of Epidemiology*, *32*(1), 1–22. doi:10.1093/ije/dyg070

[126]**Price, A. L**., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. a, & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–9. doi:10.1038/ng1847

[127]**Palmer, T. M**., Lawlor, D. a, Harbord, R. M., Sheehan, N. a, Tobias, J. H., Timpson, N. J., et al. (2012). Using multiple genetic variants as instrumental variables for modifiable risk factors. *Statistical Methods in Medical Research*, *21*(3), 223–42.doi:10.1177/0962280210394459

[128] **Dormi A**., D'Addato, S., & Borghi, C. (2011). From Risk Factor Assessment to Cardiovascular Disease Risk and Mortality Modification. *Clinical Lipidology*, 1–10.

[129]**AdelH**. Allam, MD; Randall C. Thompson, MD; L. Samuel Wann, MD; Michael I. Miyamoto, MD, MS; Gregory S. Thomas, MD, MPH. Computed Tomographic Assessment of Atherosclerosis in Ancient Egyptian Mummies
JAMA. 2009;302(19):2091-2094. doi:10.1001/jama.2009.1641.

[130]**Thompson RC**, Allam AH, Wann LS, Finch CE, Thomas GS. Atherosclerosis in ancient populations - Authors' reply *Lancet*. 2013 Jul 13;382(9887):123-4. doi: 10.1016/S0140-6736(13)61555-X.

[131]**Keller A**[1], Graefen A, Ball M, Matzas M, Boisguerin V, Maixner F, Leidinger P, Backes C, Khairat R, Forster M, Stade B, Franke A, Mayer J, Spangler J, McLaughlin S, Shah M, Lee C, Harkins TT, Sartori A, Moreno-Estrada A, Henn B, Sikora M, Semino O, Chiaroni J, Rootsi S, Myres NM, Cabrera VM, Underhill

PA, Bustamante CD, Vigl EE, Samadelli M, Cipollini G, Haas J, Katus H, O'Connor BD, Carlson MR, Meder B, Blin N, Meese E, Pusch CM, Zink A. New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing. *NatCommun.* 2012 Feb 28;3:698. doi: 10.1038/ncomms1701.

[132]**Zink A**[1], Wann LS[2], Thompson RC[3], Keller A[4], Maixner F[5], Allam AH[6], Finch CE[7], Frohlich B[8], Kaplan H[9], Lombardi GP[10], Sutherland ML[11], Sutherland JD[12],Watson L[13], Cox SL[14], Miyamoto MI[15], Narula J[16], Stewart AF[17], Thomas GS[18], Krause J[19]. Genomic correlates of atherosclerosis in ancient humans. *Glob Heart*. 2014 Jun;9(2):203-9. doi: 10.1016/j.gheart.2014.03.2453.

[133]**Gu F**[1], Pfeiffer RM, Bhattacharjee S, Han SS, Taylor PR, Berndt S, Yang H, Sigurdson AJ, Toro J, Mirabello L, Greene MH, Freedman ND, Abnet CC, Dawsey SM, Hu N, Qiao YL, Ding T, Brenner AV, Garcia-Closas M, Hayes R, Brinton LA, Lissowska J, Wentzensen N, Kratz C, Moore LE, Ziegler RG, Chow WH, Savage SA, Burdette L, Yeager M, Chanock SJ, Chatterjee N, Tucker MA, Goldstein AM, Yang XR. Common genetic variants in the 9p21 region and their associations with multiple tumours.*Br J Cancer*. 2013 Apr 2;108(6):1378-86. doi: 10.1038/bjc.2013.7. Epub 2013 Jan 29.

[134]**Helgadottir A**[1], Thorleifsson G et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction.*Science.* 2007 Jun 8;316(5830):1491-3. Epub 2007 May 3.

[135]**Samani NJ**[1], Raitakari OT, Sipilä K, Tobin MD, Schunkert H, Juonala M, Braund PS, Erdmann J, Viikari J, Moilanen L, Taittonen L, Jula A, Jokinen E, Laitinen T,Hutri-Kähönen N, Nieminen MS, Kesäniemi YA, Hall AS, Hulkkonen J, Kähönen

M, Lehtimäki T Coronary artery disease-associated locus on chromosome 9p21 and early markers of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2008 Sep;28(9):1679-83. doi: 10.1161/ATVBAHA.108.170332. Epub 2008 Jul 3.

[136]**Peloso GM**, Auer PL, Bis JC, et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks.*Am J Hum Genet*. 2014 Feb 6;94(2):223-32. doi: 10.1016/j.ajhg.2014.01.009.

[137]**Sarkar D**[1], Leung EY, Baguley BC, Finlay GJ, Askarian-Amiri ME. Epigenetic regulation in human melanoma: past and future. *Epigenetics*. 2015;10(2):103-21. doi: 10.1080/15592294.2014.1003746.

[138]**Jakovcevski M**[1], Akbarian S. Epigenetic mechanisms in neurological disease. *Nat Med.* 2012 Aug;18(8):1194-204. doi: 10.1038/nm.2828.

[139]**Brasacchio D**[1], Okabe J, Tikellis C et al. Hyperglycemia induces a dynamic cooperativity of histone methylase and demethylase enzymes associated with gene-activating epigenetic marks that coexist on the lysine tail. *Diabetes*. 2009 May;58(5):1229-36. doi: 10.2337/db08-1666. Epub 2009 Feb 10.

[140]**Aavik E**[1], Lumivuori H[1], Leppänen O et al. Global DNA methylation analysis of human atherosclerotic plaques reveals extensive genomic hypomethylation and reactivation at imprinted locus 14q32 involving induction of a miRNA cluster. *Eur Heart J*. 2015 Apr 21;36(16):993-1000. doi: 10.1093/eurheartj/ehu437. Epub 2014 Nov 16.