Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
" BIOCHIMICA, BIOLOGIA MOLECOLARE E BIOTECNOLOGIE"

CICLO XXVIII

COORDINATORE Prof. BARNARDI FRANCESCO

AN INTEGRATED GENOMIC-TRANSCRIPTOMIC APPROACH TO DETECT GENES ASSOCIATED WITH ATHEROSCLEROSIS. THE PROTO-ONCOGENE BCL3 IS A POTENTIAL CANDIDATE.

Settore Scientifico Disciplinare BIO/12

Dottoranda

Dott.ssa Meneghetti Silvia

Tutore

Prof.ssa Marchetti Giovanna

Anni 2013/2015

Table of contents

Acknowledgements

I would like to express my gratitude and to thank people who made this possible and supported me throughout these three years.

Firstly, I would like to express the deepest gratitude to my supervisor: Giovanna Marchetti for her constructive criticism and continuous guidance. I thank her for all the time dedicated to me, without her supervision and constant help this thesis would not have been possible.

I am grateful and would like to thank my phD coordinator Bernardi Francesco for his guidance, instructions and encouragement.

Thanks to Marie-Luce Bochaton Piallat and her research groups (Anita, Antonia, Chiraz) who accepted me in lab at the CMU at Geneva. My presence at her lab improves my scientific and technical background and my English knowledge. Moreover, allowed me to spend a beautiful experience.

I wish to thank my colleagues for their help: Alessio, Irving, Marcello, Mattia, Matteo, Silvia. I am grateful for their support, for cheering me up and the "*sometimes*" needed distractions. Thanks to Paolo for never-ending discussions and the laughs during the coffee break.

I am graciously thankful to Elena my best colleague-friend in life, who stood beside me throughout these years, for her "unconditional love", her encouragement, understanding and trust in me.

My warmest thanks to my dear Riccardo for his continuous support and encouragement.

Most importantly, I wish to thank my parents, Adriana and Luciano, who have been great over the years, for their unconditional love and unlimited support.

Riassunto

L'aterosclerosi è una patologia degenerativa e progressiva delle arterie di grosso e medio calibro, caratterizzata da disfunzione endoteliale, accumulo di lipidi, infiltrazione di linfociti, migrazione e proliferazione di cellule muscolari lisce e deposizione di matrice extracellulare nella parete vascolare, con conseguente formazione di una placca ateromasica.
Il processo aterosclerotico è alla base della patogenesi della malattia cardiovascolare nelle sue diverse manifestazioni cliniche, quali la malattia coronarica (CAD), l'infarto del miocardio (MI), l'arteriopatia ostruttiva periferica e l'ictus. L'aterosclerosi e' responsabile delle prime tre cause di mortalità e morbilità nel mondo.
L'aterosclerosi è una patologia complessa, alla quale concorrono molteplici fattori genetici ed ambientali.

Lo scopo di questa tesi è stato l'identificazione di nuove componenti genetiche e molecolari della patologia aterosclerotica. Lo studio si e' sviluppato a livello di DNA, RNA e proteina, applicando diverse metodologie e integrando dati ottenuti mediante: i) analisi di polimorfismi a singolo nucleotide (SNPs) in soggetti con e senza CAD, ii) analisi dei livelli di espressione genica in cellule muscolari lisce vascolari (VSMCs) e in placche aterosclerotiche umane, tramite microarray e Real time PCR, iii) analisi dei livelli di proteina mediante immunoistochimica di *specimens* di parete arteriosa carotidea.

Partendo dai dati relativi a 91 SNPs analizzati in 510 pazienti con CAD e MI e in 388 soggetti di controllo senza patologia coronarica nella fase di replicazione dello studio di associazione "genome-wide" del Consorzio MiGen, sono stati selezionati 15 SNPs associati nominalmente al CAD ($P<0.1$).

Per studiare la potenziale associazione di questi SNPs con il processo aterosclerotico, i livelli di espressione dei 71 geni prossimali ai 15 SNPs sono stati analizzati sia in VSMCs, che rivestono un ruolo primario nel processo aterosclerotico, sia in porzioni di parete arteriosa derivanti da interventi di endoarteriectomia carotidea. A tal fine sono state condotte due consecutive ed indipendenti analisi del trascrittoma mediante microarray, la prima in culture primarie di VSMCs isolate da placche aterosclerotiche e dalla porzione prossimale virtualmente sana, la seconda in *specimens* di placche carotidee e delle corrispondenti porzioni "sane". L'analisi dei profili di espressione ha permesso di identificare tre geni differenzialmente modulati, *BCL3, PVRL2 e ABCA1*, quest'ultimo già ampiamente studiato in relazione alla patologia aterosclerotica.

Ponendo l'attenzione ai due geni adiacenti sullo stesso cromosoma (19q13) sono stati analizzati 4 SNPS intragenici, due per *BCL3* e due per *PVRL2* ed uno intergenico *BCL3-PVRL2* in 393 soggetti di controllo e 442 pazienti con CAD senza pregresso MI. Quest'ultima coorte è stata selezionata per poter investigare in modo preferenziale l'eventuale associazione con l'aterosclerosi piuttosto che con la sua complicanza trombotica acuta.
L'analisi dei genotipi ha mostrato che i portatori dell'allele G del polimorfismo *BCL3* rs2965169 erano più rappresentati tra la popolazione con CAD e l'associazione con la patologia rimaneva significativa anche dopo correzione per i tradizionali fattori di rischio cardiovascolare. L'allele A del polimorfismo *BCL3*rs8100239 correlava con l'indice di massa corporea, l'ipertensione e il profilo lipidico.
L' analisi della distribuzione dei genotipi non ha identificato associazioni significative con la patologia coronarica o con variabili metaboliche sia per entrambi gli SNPs di *PVRL2* che per la variante intergenica *BCL3-PVRL2.*

L'analisi immunoistochimica di placche aterosclerotiche (n=10) e di porzioni adiacenti virtualmente sane (n=5) ha evidenziato espressione della proteina BCL3 solamente nella porzione aterosclerotica, associata a VSMCs e *foam cells.* Nella parete vascolare aterosclerotica l'espressione del mRNA per BCL3 sembra pertanto correlare con l'espressione della proteina.

Complessivamente l'approccio integrato utilizzato nello sviluppo di questa tesi supporta il coinvolgimento della proteina BCL3 nel processo aterosclerotico.

Abstract

Atherosclerosis is the common ground of several clinical manifestations of cardiovascular diseases (CVD), including coronary artery disease (CAD), myocardial infarction (MI), peripheral artery occlusive disease and stroke. CVD is still one of the major causes of mortality and morbidity in the worldwide.

Atherosclerosis is a complex multifactorial disease of the wall of medium-sized and large arteries, characterized by endothelial cell dysfunction, smooth muscle cell proliferation (VSMCs) and migration, inflammation, lipid and matrix accumulation. Susceptibility to atherosclerosis is in turn influenced by interplay of genetic and environmental factors.

The aim of this thesis was to unravel new potential genetic and molecular signatures of the atherosclerosis, by using an integrated approach which joins information from i) single nucleotide gene polymorphisms (SNPs) analysis in a case-control study of subjects with or without CAD, ii) microarray-based gene expression analysis on human cultured VSMCs and on carotid artery specimens, and iii) immunohistochemical analysis in carotid artery specimens.

Firstly, 15 SNPs nominally associated with CAD (P< 0.1) were selected from 91 SNPs, investigated within replication of a genome-wide association study –MiGen- (510 patients with CAD and MI and 388 subjects with normal coronary arteries CAD-free). The expression levels of 71 genes proximal to the 15 tag-SNPs were evaluated by two subsequent steps of microarray-based RNA profiling, the former in VSMC populations isolated from grossly non-atherosclerotic (NP) and atherosclerotic (DP) human carotid portions, and the latter in whole carotid specimens. *BCL3* and *PVRL2*, located on chromosome 19, and *ABCA1*, extensively investigated before, were found to be differentially expressed. Focusing the attention on *BCL3* and *PVRL2*, the only couple of contiguous genes differentially expressed in the transcriptomic analysis, a total of 5 SNPs, two within *BCL3* gene,

two within PVRL2 gene and one *BCL3-PVRL2* intergenic, were genotyped within CAD-free subjects (n=393) and CAD patients without MI history (n=442). This cohort enabled to preferentially investigate the atherosclerosis pathways, rather than its acute thrombotic complication.

The carriership of the *BCL3* rs2965169 G allele was more represented among CAD patients and remained independently associated with CAD after adjustment for all the traditional cardiovascular risk factors, while the *BCL3* rs8100239 A allele correlated with metabolic abnormalities. No significant associations were found for either *PVRL2* SNPs or *BCL3-PVRL2* intergenic variant.

A BCL3 positive immunostaining was detected in the intima-media of atherosclerotic specimens, but not within non-atherosclerotic ones, thus indicating a correlation between BCL3 mRNA and protein levels.

This thesis, which integrates GWAS data with the downstream changes in the RNA and protein levels in human arterial wall specimens, supports a role for BCL3 in atherosclerosis.

Chapter I: Introduction

1. 1 HISTOLOGY OF THE BLOOD VESSEL

The arterial wall consists of a concentric, well-developed trilaminar structure, called, from the luminal side outward, tunica intima, tunica media and tunica adventitia. The thickness of these three layers varies greatly depending upon the size and type of vessel (large, medium, small arteries and veins, capillaries) [1].

The tunica intima is the luminal or innermost layer in direct contact with the blood. It is composed of a single layer of endothelial cells and a small amount of subendothelial connective tissue. The tunica intima is separated from the media by a dense elastic membrane called the internal elastic lamina. In addition to forming the physical interior wall of the blood vessel, the endothelium acts as a selectively permeable barrier and can secrete chemicals that stimulate dilation or constriction of the vessel. The healthy endothelium has two known key functions [2]. First, it provides a non-thrombogenic surface that prevents blood coagulation, but when tissue around vessel is inflamed, the endothelial cells produce cell-adhesion molecules that induce leukocytes to adhere to the surface. Second, the endothelium acts as a "gate keeper" by selectively controlling the passage of circulating blood cells (i.e. monocytes and lymphocytes) and active molecules (i.e. nitric oxide and L-arginine) into and out of the arterial wall [2].

The middle layer, the tunica media, is separated from the intima by the internal elastic lamina which serves as a border with the overlying intima. The media is composed of only one cell type, the vascular smooth muscle cells (VSMCs) embedded in VSMC-produced extracellular matrix. The VSMCs present a variety of different structural and physiological functions. In fact, the VSMCs are

responsible of the capacity of withstanding the high pressure of the circulating blood compartment and are also involved in arterial repair after injury in adult life [3]. In physiologic conditions, SMCs are in a quiescent state characterized by a low rate of cell replication and death [4]. The abluminal side of the tunica media is characterized by the external elastic lamina, a prominent fenestrated membrane on the border with the adventitial layer [1].

The outermost layer, called adventitia, is a connective tissue layer and mostly consists of collagen fibers and proteoglycans, and a loose network of elastic fibers. It contains the vasa vasorum that provide nutrients and remove waste products from the outer portion of the arterial wall, and nerves. Fibroblasts constitute the most abundant cells of the adventitia [1] (***FIGURE i-1***).

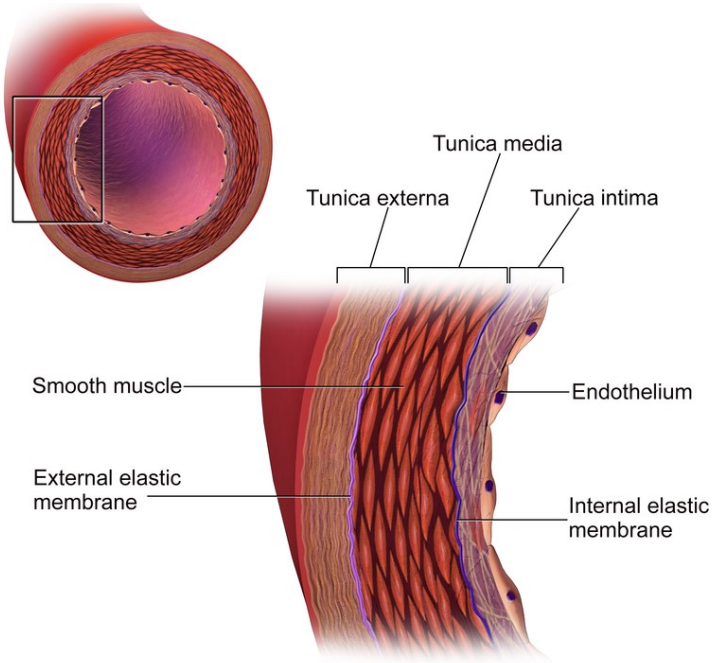


***FIGURE i-1*. Structure of the arterial wall** [from internet]

1.2 ATHEROSCLEROSIS

Atherosclerosis is a complex multifactorial disease of the wall of medium-sized and large arteries. Atherosclerosis is the common ground of several clinical manifestations of cardiovascular disease (CVD), such as coronary artery disease

(CAD), myocardial infarction (MI), peripheral artery occlusive disease and stroke. Therefore, atherosclerosis is the single largest cause of death and disability in the western world. This inflammatory, chronic, progressive disease to become clinically manifest requires the formation of a fibro-lipid plaque within the arterial wall resulting in reduced blood flow.

The lesions occur mainly at the level of the bifurcations in which the flow changes from laminar to turbulent. Two mechanisms have been identified, which could explain the link between disturbed blood flow and atherosclerosis development.

The first, the “mass transport theory”, is based on the evidence that the transport of certain bioactive substances (e.g. LDL) from the circulation to the vessel wall may be promoted at sites of disturbed flow due to prolonged contact between blood and vascular ECs. This differs from the “shear stress theory”, which emphasized the effect of blood flow-induced mechanical forces on vascular physiology. Of note, these theories are not mutually exclusive. Both mass transport and shear stress influence plaque formation, and these factors interact at a functional level, e.g. shear stress alters vessel permeability and in turn molecular transport. [5].

The formation of atherosclerotic lesions involves several processes, i) the accumulation of lipids, mainly cholesterol free and esters of cholesterol in the intima (especially in the subendothelial space); ii) the establishment of an inflammatory infiltration of lymphocytes and macrophages engulfing lipids and becoming foam cells; c) migration and proliferation of VSMCs with production of extracellular matrix. Although the term "atherosclerosis" is derived from the greek atheré, "mush" and sclerosis, "hardening", it is important to emphasize that in the lesions can be a great variability in the formed tissue due to the prevalence of each of the above mentioned processes [6;7].

Atherosclerosis involves a large genetic network which extends to interaction with environmental factors [8]. The main risk factors for developing this disease, largely related to lifestyle, are smoking, hyperlipidaemia and hypertension source of stress on the endothelium, obesity, diabetes and sedentary life. Other important risk factors are gender and age: women are in fact protected by oestrogen until menopause, and the disease typically occurs after the 40-50 years [8].

1.2.1 Pathogenesis of atherosclerosis

Atherosclerosis is a chronic, progressive, disease with a long asymptomatic phase. Its development usually extends over decades [6;9].

The initial event in atherogenesis is the endothelial dysfunction, triggered by exposure to irritative stimuli such as hyperlipidaemia, high shear forces, and proinflammatory cytokines. This leads to acquire new functional and antigenic properties. Simultaneous changes in endothelial permeability and the composition of the extracellular matrix beneath the endothelium promote entry and retention of cholesterol-rich low-density lipoproteins (LDL) in the arterial wall. The dramatic changes in endothelial functional properties are mediated by induced expression of several genes by nuclear factor- Kappa B (NF-κB) commonly a proinflammatory transcription factor [10]. Through the activation of NF-κB the endothelial cells begin to express cell adhesion molecules (e.g., E-selectin, P-selectin, intercellular adhesion molecule-1), which trigger capture and adhesion of leukocytes, many of which are monocytes, thus forming a first non-obstructive lesion called fatty streak. Enzymes expressed by ECs, phagocytic cells, VSMCs and fibroblasts have been proposed as a possible source of oxidizing agents involved in the oxidation of LDL (e.g., NADPH oxidase, NO-synthase, myeloperoxidase and lipoxygenase). The oxidizing species most involved in the formation of LDL oxidized (oxLDL) are oxygen, chlorine, nitrogen and free radicals derived from lipid peroxidation. The oxLDL are involved in many key processes in the development of atherosclerotic lesion: induce the transformation of monocytes into macrophages and these then in foam cells; stimulate ECs and VSMCs to secrete monocyte chemotactic protein-1 and thereby encourage their migration within the wall; promotes the adhesion and chemotaxis of T lymphocytes and macrophages. Under the influence of the macrophage colony stimulating factor, monocytes are differentiated into macrophages that upregulate both toll-like receptors, involved in macrophage activation, and scavenger receptors, that internalize oxLDL, leading to lipid accumulation and foam cell formation [10]. As the inflammatory process advances, a high number of monocytes is attracted by chemotactic cytokines and induced to maturate into macrophages which after uptake of lipid, become foam cells (***FIGURE i-2b***).

The progressive enlargement of the lesion involves the recruitment of VSMCs from the tunica media into the tunica intima (***FIGURE i-2c***). In this layer the VSMCs can switch from a differentiated (also termed contractile) to a dedifferentiated (also termed synthetic) phenotype in response to mediators such as platelet derived growth factor and transforming growth factor-β (TGF-β). The VSMCs with contractile phenotype are characterized by high levels of expression of contractile genes and low rates of proliferation, migration, and extracellular matrix synthesis. Conversely, the synthetic VSMCs have increased rates of proliferation, migration, and production of extracellular matrix, as well as reduced expression of contractile genes [11].

The synthetic VSMCs secrete collagen, elastin and proteoglycans, and form a fibrous cap that covers the plaque which changes the lipid rich fatty streak into more advanced lesion. At this stage, the plaque appears white-yellowish and tends to protrude into the lumen of the artery producing an alteration of the blood flow. The cap increases the size of the plaque but at the same time protects the deeper thrombogenic components of the core from exposure to circulating blood.

As the cap matures, the underlying region undergoes some transformations. The most obvious is the death by apoptosis of the VSMCs and foam cells, which release oxidized lipids that accumulate in the central region of the plaque called lipid core (***FIGURE i-2d***) [1;9].

The advanced lesion may also be calcified by an active process which greatly resembles the active formation of bone [9].

Calcification of a lesion may also contribute to its likeliness to rupture. Virmani et al. [12] have hypothesized that in certain types of lesions calcified nodules may cause physical stress to the fibrous cap and lead to its rupture. On the other hand, Abedin et al. [13] argue that calcification originally destabilizes a plaque, but after the area has become highly calcified, it actually protects the plaque from rupture. Another study indicated that lesions with superficial calcium deposits are more prone to rupture than those with deep deposits [14].

In addition to calcification, neovascularization affects the structure of the plaque. As the plaque size increases, the oxygen from the bloodstream does not reach all areas of the lesion, and the inner section gets hypoxic. Hypoxia has been shown to increase lesion progression in atherosclerosis by promoting lipid accumulation,

inflammation and especially the neovascularization following the hypoxic state has been connected to rupture prone plaques [15].

Plaques generally cause clinical manifestations by producing flow-limiting stenosis that lead to tissue ischaemia, or by provoking thrombi that can interrupt blood flow locally or embolize. Thrombi often arise after physical disruption of the plaque, most commonly a fracture of the fibrous cap that exposes tissue factor (TF) a potent trigger of coagulation cascade. Thrombosis is the last stage in the advancement of the disease and is responsible for the acute clinical events, to be paid by coronary, cerebral and peripheral districts [9;16].

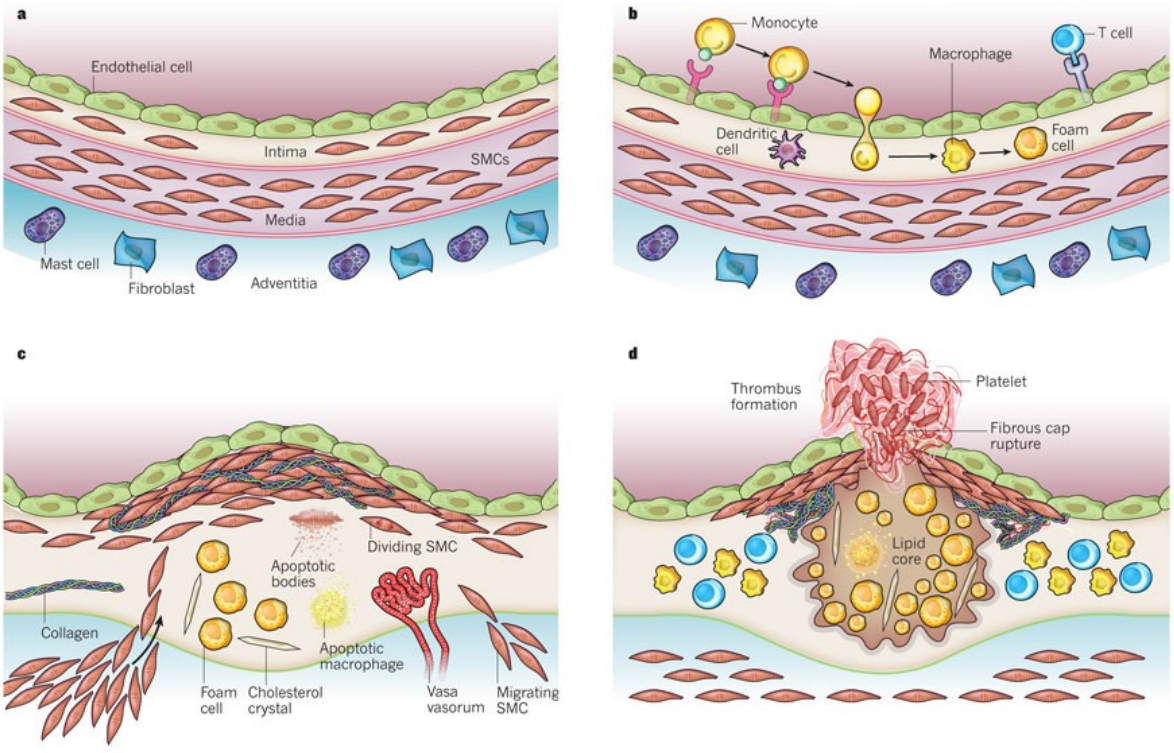


***FIGURE i-2*. Stages in the development of the atherosclerotic lesion** [9].

1.2.2 Classification of the atherosclerotic lesions

The American heart association (AHA) has classified the atherosclerotic process into six main stages, based on histological analysis of plaque at the point of maximum extension [6;7].
The particular characteristics of each stage are described below and are shown graphically in ***FIGURE i-3***

The first three stages are considered an early stage of atherosclerosis, these lesions show no disruption of intimal structure and there is no effect on the media or adventitia and are considered to be reversible changes.
Type I: is classified as the initial lesion. This includes an increase in macrophages and macrophage foam cells due to atherogenic lipoproteins. These changes can be seen in their adaptive intimal thickening considered as physiological response to insults artery blood flow. The thickening is a natural process occurring in arteries in everyone since birth extension [6].
Type II: lesions are characterized by fatty streaks. These consist of macrophage foam cells and lipid-laden smooth muscle cells, arranged in stripes of yellow grossly identifiable on the surface of the intima [6].
Type III: is an intermediate stage between stage II, fatty streaks, and IV which can lead to serious cardiac events. Type III is non-threatening since the endothelial layer of the artery is healthy and lipid deposits are small. These arteries have type II characteristics with additional scattered areas of extracellular lipid collections that accumulate between intimal smooth muscle cells [6].

Type IV-VI lesions are often called advanced lesions. These lesions show intimal disorganization and are often symptomatic. They predispose to sudden lesion progression and are not reversible.
Type IV: these lesions can cause stenosis since the lipid pool determines the degree which the lumen is constricted. These lesions still have smooth muscle cells and do not have extensive fibrotic tissue. If the lipid begins to rapidly grow, this lesion can narrow the lumen and be at high risk for plaque rupture, sending a flood of lipid into the blood stream. This leads to clotting which can lead to coronary occlusion, stroke, or death [7].

Type V: lesion is characterized by thick layers of fibrous tissue. Type V can stem from Type IV lesions with lipid deposits with additional fibrous tissue or multilayered fibrous tissue with multilayered lipid deposits. Type V can also have calcium deposits or no lipid or calcium deposits [7].

These type of lesions can occlude the artery slowly leading to gradual complications.

Type VI: this lesion can cause fissure, hematoma, and thrombus. Type VI additionally can have the lipid characteristics of type IV or the calcium and/or fibrous tissue characteristics of type V. Fissures, which are tears or grooves in the surface of a lesion, can cause hematoma and thrombosis. In extreme cases they can structurally weaken the artery. Fissures vary in size and length and consequently the effect of fissures varies [7].

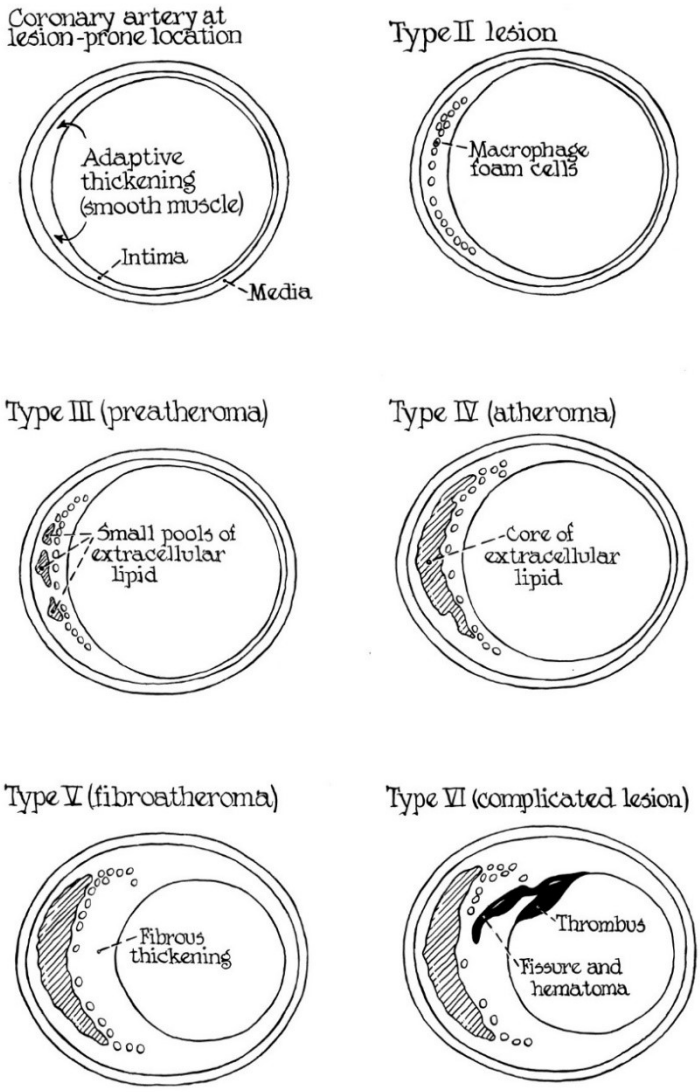


FIGURE i-3. Classification of atherosclerotic lesions according to the American Heart Association [7].

1.2.3 Carotid atherosclerosis

Carotid arteries are paired structures located in both the left and right sides of the neck. The left common carotid artery (CCA) originates at the aortic arch, whereas the right CCA originates at brachiocephalic artery. The CCA bifurcates into the internal carotid artery (ICA) and external carotid artery (ECA) as shown in ***FIGURE i-4***. The ICA supplies oxygenated blood to the brain and the ECA supplies oxygenated blood to the face, scalp, and skull. Atherosclerotic plaques usually occur at sites of non-laminar turbulent flow such as the carotid bifurcation and the proximal ICA.

Hemodynamic forces at the carotid bifurcation play a significant role in the localization of intimal thickening in the vessel wall. The magnitude and rate of change of blood flow at the luminal surface have been closely tied to the pathogenesis of atherosclerotic plaque formation. Laminar flow results in a gradient of fluid velocities as you move from the wall towards the center of a tube. Friction between fluid along the wall and the wall itself creates a tangential force exerted by flowing fluid on the wall of the tube, and this is referred to as “wall shear stress.” Arterial segments with low and oscillatory wall shear stress appear to be atheroprone. In vivo and in vitro research has shown that disturbed flow and low shear conditions predispose to endothelial dysfunction [5].

The carotid arteries are vulnerable to atherosclerosis, and because they supply blood to the brain, vessel narrowing and embolisms from carotid atherosclerosis could cause stroke and subsequent cognitive impairment [17]. Some studies also found that carotid atherosclerosis was independently associated with poorer cognitive function in subjects without a clinical stroke [18;19].

The carotid intima-media thickness (IMT) is a widely used surrogate marker for atherosclerosis worldwide because it can be simply, reproducibly, and noninvasively measured. Many studies have reported that carotid IMT measurements are useful for evaluating the risk and incidence of cardiovascular disease (CVD) [20].

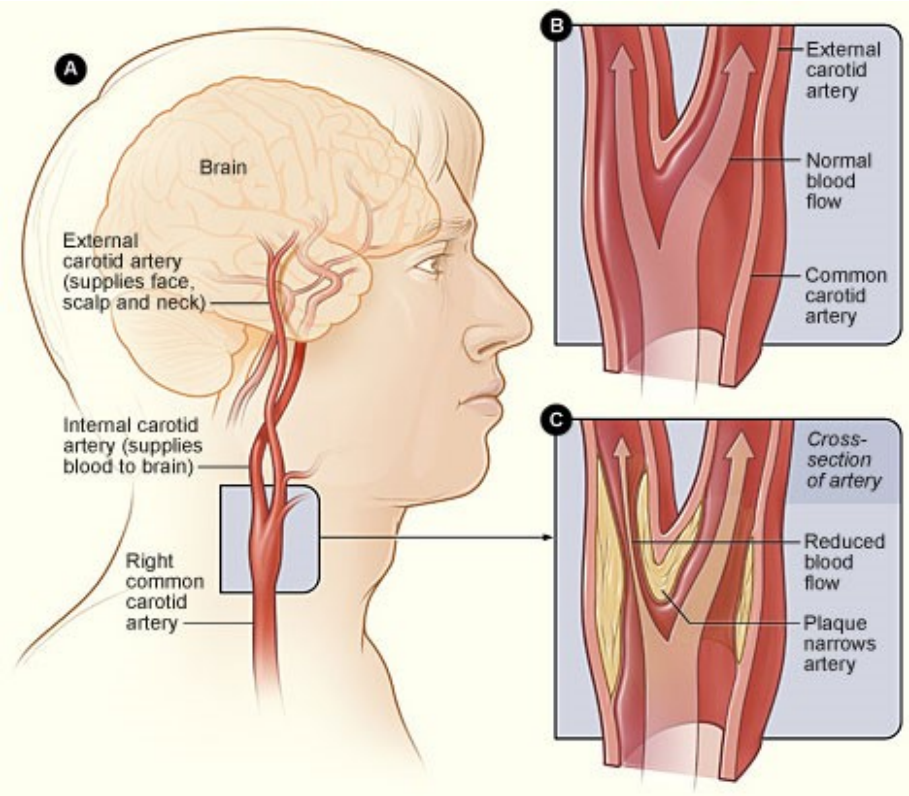


FIGURE i-4. Carotid artery. A) location of the right carotid artery in the head and neck. B) shows the inside of a normal carotid artery that has normal blood flow. C) show the inside of a carotid artery that has plaque build-up and reduced blood flow (images from NIH web site)

1.2.4 VSMCs and atherosclerosis

VSMCs are the predominant cellular elements of the vascular media and are involved in all the physiological functions (e.g. blood vessel tone-diameter, blood pressure, and blood flow distribution) and the pathological changes taking place in the vascular wall. VSMCs show considerable differences depending on their position in the arterial tree, their embryologic origin, and their organ-dependent microenvironment [21].

In healthy adult blood vessels, VSMCs generally display a contractile or differentiated phenotype, characterized by a slow rate of proliferation and the expression of contractile, or smooth muscle cell, markers [22]. These markers include calponin, smooth muscle alpha actin (α-SMA), SM myosin heavy chain (MHC), and SM22 [23].

Contractile VSMCs are arranged in circumferential direction with a well-organized structure and are filled with myofilaments and dense bodies but contain a relatively poorly developed Golgi apparatus and rough endoplasmatic reticulum.

Conversely, in the context of vascular injury VSMCs often switch from a contractile to synthetic, or undifferentiated phenotype, which is characterized by a decrease in the expression of contractile markers. Moreover, synthetic VSMCs display increased rates of VSMC proliferation, migration and extracellular matrix (ECM) remodelling. The VSMCs in this synthetic phenotype are myofilament-poor but are relatively rich in organelles involved with protein synthesis, such us rough endoplasmic reticulum and Golgi apparatus [7].

Phenotypic changes of VSMCs are associated with development and progression of the atherosclerotic plaque [24,25].
In atherosclerosis VSMCs appear to play a dual role: contribution to plaque development and to plaque stability. Their migration, proliferation and transformation into foam cells drive to atherogenesis and plaque development. Conversely, abundant VSMCs and matrix form a thick fibrous cap that contributes to the plaque stability [9,25].

VSMCs have been extensively used as a model to characterize molecular signatures of atherosclerosis, as well as to functional validation of candidates for atherosclerosis/CAD [26-28].

1.3 NUCLEAR FACTOR-κB PATHWAY IN ATHEROGENESIS

NF-κB is the general name for a family of transcription factors, which consists of five members, p50, p52, p65 (RelA), c-Rel, and RelB (***FIGURE i-5***). These five NF-κB subunits assemble combinatorially into functioning homo- and heterodimers, to produce 15 possible NF-κB transcription factor complexes that bind a set of specific DNA elements, known as κB sites, in the enhancers/promoters of target genes

[10;29]. The prototypical form of NF-κB is a heterodimer of p50 and p65. A distinctive characteristic of all NF-κB family members is the presence of the Rel homology domain (RHD) in the N-terminus. RHD is responsible for dimerization, interaction with NF-κB inhibitors, nuclear localization, and DNA binding p65, RelB and c-Rel subunits also contain a transactivation domain in their C-terminus [10;29].The p50 and p52 proteins lack the transactivating domains found in the C-terminal regions of the Rel proteins and thus, DNA-bound p50 and p52 homodimers inhibit gene transcription [10;29] (***FIGURE i-5***).

There are seven inhibitors (IκB): IκBα, IκBβ, IκBε, IκBγ, BLC-3 and the two precursors, p100 and 105 [10;29].
The IκB members can be broadly divided into three categories: cytoplasmic IκBs (IκBα, -β and –ε); precursor IκBs which are cleaved to form Rel subunits p50 and p52 (p100/IκBδ and p105/IκBγ) and nuclear IκBs (IκBζ, BCL3 and IκBNS) [30].

In resting cells, NF-κB dimers are sequestered in the cytoplasm through interaction with an IκB protein, which masks nuclear localization sequences (NLS) [29]. Upon stimulation (e.g. cytokines, oxidized LDL, endotoxins, AGE, and angiotensin II), various signalling pathways leads to activation of the IκB kinase (IKK) complex, which in turn leads to serine phosphorylation of IκB by IKKβ, poly-ubiquitination and proteasome-mediated degradation of IκB. Degradation of IκB frees the NF-κB dimer to translocate to the nucleus. In the nucleus, NF-κB binds to κB elements in promoter regions of its target genes, and induces target gene expression. Target genes of NF-κB include cytokines, cell adhesion molecules, stress response genes, growth factors and other transcription factors (e.g. BCL3, c-REL, rel-B, Snail, Sox9 and YY1) [10;29].

Overall, the nuclear factor κB (NF-κB) family proteins regulate the transcription of a vast collection of inducible effector genes, which influence a broad range of biological processes, including innate and adaptive immunity, inflammation, stress responses, B-cell development, and lymphoid organogenesis.

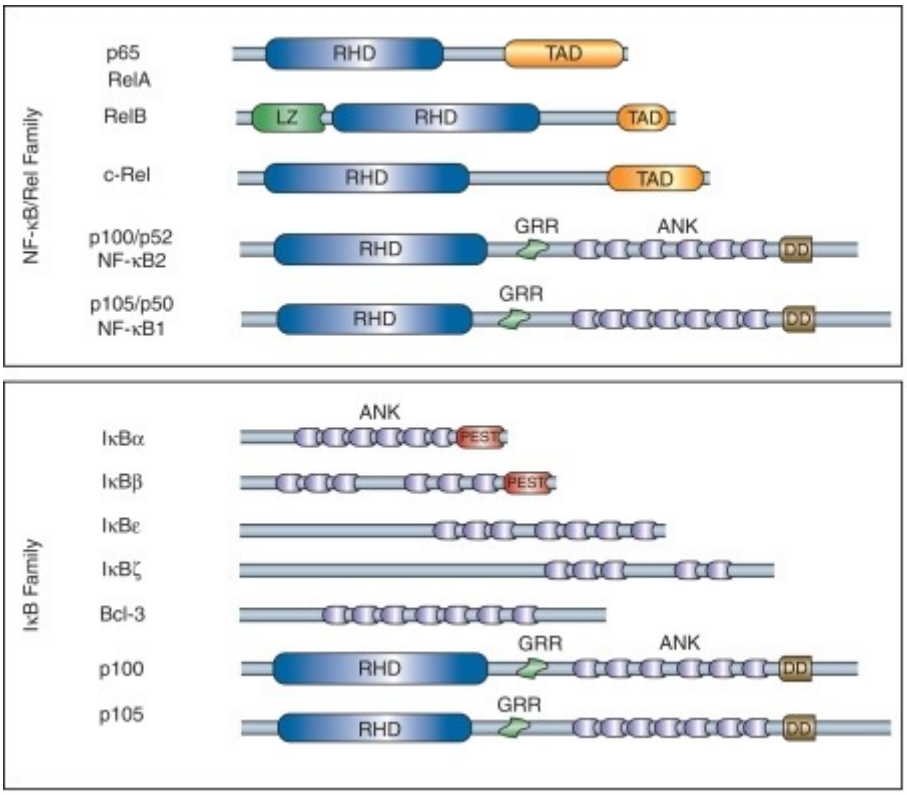


FIGURE i-5. **Members of the NF-κB and IκB protein families.** The domains that typify each protein are indicated schematically. CC, coiled-coil; DD, death domain; GRR, glycine-rich region; HLH, helix-loop-helix; IKK, IκB kinase; LZ, leucine-zipper; NBD, NEMO binding domain; PEST, proline-, glutamic acid-, serine-, and threonine-rich region; TAD, transactivation domain; ZF, zinc finger [29].

Evidence suggests that NF-κB signalling pathway is relevant, if not central, in atherosclerosis. In animal and human arteries, NF-κB activation and expression of NF-κB target genes has been observed in atherosclerotic lesions [10;31;32].

Mice that are deficient in NF-κB activating receptors, such as IL6 receptor, RAGE, TLR4, TLR2 and an upstream adapter molecule, MyD88, showed reduced atherosclerotic lesion development [33], as well as mice with p50 deficiency in hematopoietic cells [34]. Many risk factors for atherosclerosis, such as oxidative stress and oxidized LDL are also activators of the NF-κB signalling pathway [35]. In addition, many target genes of NF-κB, such as adhesion molecules VCAM-1, ICAM-1, E-selectin, and proinflammatory cytokines and chemokines like monocyte chemoattractant protein-1, are essential for recruitment of leukocytes to the intima in the early stages of atherosclerosis and progression of atherosclerotic lesion [10].

Positive feedback of NF-κB activation by cytokines and chemokines that are induced by NF-κB signalling can result in amplification of cell signalling, which may help perpetuate atherosclerosis lesion formation and growth. Apoptosis resulting from aberrant or unbalanced activation of NF-κB can potentially contribute to formation of the necrotic core in advanced atherosclerotic lesion [32].

1.4 BCL3, B-CELL LYMPHOMA 3

B-Cell Lymphoma 3 (BCL3) was originally discovered by its translocation into the immunoglobulin alpha-locus in B cell t (14;19) (q32.3; ql3.2), found in a subset of patients with B-cell chronic lymphocytic leukaemia. The translocation leads to the juxtaposition of BCL3 to the immunoglobulin heavy chain gene locus, resulting in high-level expression of the BCL3 transcript [36;37].

The BCL3 gene maps to chromosome 19q13, and encodes a phosphoprotein of 454 amino acids exhibiting an apparent molecular weight between 47 and 60 kDa [37].

BCL3 transcript shows a broad expression pattern in multiple cell types: it is highly expressed in spleen, liver and kidney, with no apparent expression in brain. Transcription of BCL3 is regulated through several signalling pathways, included the NF-kB [29].

The overexpression of BCL3 is proposed to cause dysregulation of genes normally regulated by NF-kB transcription factors implicated in cell proliferation, apoptosis and differentiation [37;38].

Although the precise role of BCL3 in the NF-kB signalling is not fully understood, it is proposed that BCL3 regulates transcriptional activity and subcellular localization of the transcription factor NF-κB. BCL3 associates specifically with the inhibitory p50 and p52 homodimers in the nucleus, and in contrast to the classical IκB proteins, this association with BCL3 leads to an increased transcriptional activity.

These complexes can either activate or repress transcription of target genes [10;39;40].

The paradoxicall dual role of BCL3, promotion or inhibition of NF-kB target gene expression, could depend on the type of cell, the type of activating stimulus, and type of the NF-kB target gene involved [37].
The potential of BCL3 to either promote or impede transcription has been shown to depend to some extent on different post-translational modifications [39;41]. BCL3 protein is modified by phosphorylation and polyubiquitination. The extent of BCL3 phosphorylation has been shown to affect its interaction with both NF-kB p50 and p52 [39].
In addition, polyubiquitination has also been shown to regulate BCL3 entry into the nucleus. Deubiquitination of BCL3 was found to prevent its nuclear translocation and, as a result, inhibited p50:BCL3- or p52:BCL3-dependent transcription [41].

1.5 GENES AND ATHEROSCLEROSIS

For decades epidemiological and familial studies have claimed that 30-60% of the interindividual variation in CAD risk is accounted by heritable factors [42].
Since the beginning of this millennium a large effort has been devoted to understanding the genetic bases of atherosclerosis/cardiovascular disease by using genome-wide association study (GWAS) [43].

A GWAS approach uses methods of standard design: DNA from cases (people who carry the disease) are compared to controls (non-carriers) using micro-array chips which can examine up to a million SNPs (occurring at a frequency greater than 1%) per sample [44;45]. After obtaining candidate variants, an important next step is to evaluate and select variants by utilizing appropriate significance tests.

To minimize the false negative rate due to multiple test correction, conventional GWAS analysis requires stringent threshold of P values of less than 5x10-8. However, this stringent requirement has excluded many genuinely associated

SNPs that have moderate or weak association signals. Once identified, the putative disease loci are then subject to further study by a replication stage. Replication studies should be conducted in an independent dataset drawn from the same population as the GWAS, in an attempt to confirm the effect in the GWAS target population. Once an effect is confirmed in the target population, the disease loci are then subject to further study to identify causal variants through sequencing of the region and further functional studies to identify causal mechanisms [44;45].

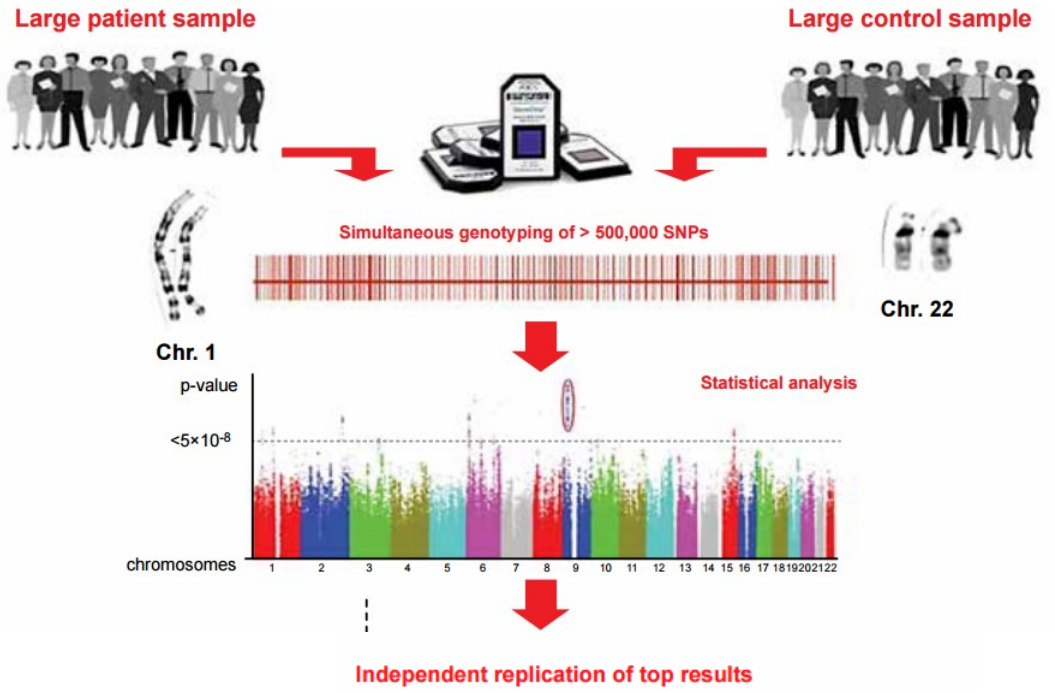


FIGURE i-6. Principle of genome-wide association studies [modify from 45].

The first GWAS studies in CAD were independently and simultaneously published in 2007 from two groups, led by McPherson and colleagues [46] and Helgadottir and colleagues [47].

Either these studies determined only a few polymorphisms linked to CAD, all located in the 9p21.3 genome region. The CAD association with 9p21 represents

one of the most consistent and robust SNP-disease association- in the GWAS era, having been replicated in several independent samples in numerous ethnicities including European, Korean, and Japanese populations [48,49,50].

Despite the strength and consistency of CAD associations with 9p21 variant the biologic role of this locus and how it may confer increased CAD risk is not completely understood [51]. At the molecular level, disentangling pathways linking such variant to CAD risk has been made difficult by the fact that the 9p21 variant does not reside within any known protein-coding genes. The nearest protein-coding genes, which are many thousand base pairs away, include CDKN2A, CDKN2B, and MTAP, which encode cyclin proteins involved in cell cycle regulation. However, 9p21 variant localizes within ANRIL, a large antisense noncoding RNA, the biologic function of which remains obscure. The contribution of 9p21 variant on CDKN2A and CDKN2B expression in VSMCs [26] and in knout-out mice for CDKN2A [reviewed in 42] might represent an important mechanism for the association between this genetic locus and susceptibility to atherosclerosis and CAD.

The MI Genetics (MIGen) [52] Consortium, the largest GWAS consortium to date for MI, confirmed prior findings and also identified three novel loci: 21q22 in an intergenic region, 6p24 within an intron of PHACTR1, and 2q33 within an intron of WDR12 [52]. The previous associations with 9p21, 1p13, 10q11, and 1q41 were replicated by the MIGen Consortium, further establishing these as robust associations.

In addition, a number of GWAS have also been completed to examine the genetic determinants of a number of traditional and non-traditional risk factors, including serum LDL cholesterol, high-density lipoprotein (HDL) cholesterol, and triglycerides [53].

These lipid GWAS have identified ~30 robust genetic loci that influence serum lipids, with several of them not previously implicated in the traditional pathways of CVD. Common variants near LDLR (19p13) and PCSK9 (1p32), which are known to be involved in cholesterol metabolism, were also strongly associated with CAD [52].

In addition, GWAS have also identified dozens of previously uncharacterized loci associated with type 2 diabetes mellitus, systolic blood pressure/hypertension, obesity, cigarette smoking, and C-reactive protein (CRP) [54,55].

Several variants associated with cardiovascular disease in candidate-gene association studies have not been confirmed in GWAS studies, e.g. C677T SNP in the 5,10-methylenetetrahydrofolate reductase (MTHFR) gene, and Q192R SNP in the paraoxonase 1 (PON1) gene [53].

In just over 6 years a total of 50 genetic variants predisponing to CAD of genome wide significance with confirmation in independent population [45] have been discovered (**TABLE i-1**). Only 15 of the 50 genetic risk variants are associated with conventional risk factors for CAD: seven associated with low density lipoprotein-cholesterol (LDL -C); one with high density lipoprotein (HDL); two with triglycerides; four with hypertension; and one with coronary thrombosis. The remaining 35 risk variants operate through mechanisms yet to be determined [42].
The genetic risk variants for CAD have many feature that are similar to genetic variant for the other polygenic disorder: i) the genetic variant risk for CAD occurring on average in 50% of the population with a frequency from 2% to 9%; ii) the relative increased risk of each genetic variant is small, averaging 18% with an odds ratio varying from 2% to 90%; iii) for CAD as well as other common polygenic disorders, multiple genetic risk variants are inherited by everyone. Those at high genetic risk for CAD have a greater genetic risk burden due to inheritance of a greater number of common risk variants, as opposed to inheriting one or more genetic variant of high risk; IV) most of the genetic risk variants for CAD are located in DNA sequences that do not code for protein. This means the risk variant mediates its increased risk for CAD directly or indirectly through regulation of DNA sequences that do code for protein [42].

TABLE i-1. List of 50 genetic variants (Genome-Wide Significant) associated with CAD or MI [42]**.**

Chromosomal Location	SNP	Nearby Genes (Allele)	Risk Allele Frequency	Odds Ratio (95% CI)
6q25.3	rs3798220	*LPA*	0.02 (C)	1.92 (1.48–2.49)
2p24.1	rs515135	*APOB*	0.83 (G)	1.03
1p13.3	rs599839	*SORT1*	0.78 (A)	1.29 (1.18–1.40)
19p13.2	rs1122608	*LDLR*	0.77 (G)	1.14 (1.09–1.19)
19q13.32	rs2075650	*APOE*	0.14 (G)	1.14 (1.09–1.19)
2p21	rs6544713	*ABCG5-ABCG8*	0.29 (G)	1.07 (1.04–1.11)
1p32.3	rs11206510	*PCSK9*	0.82 (T)	1.15 (1.10–1.21)
6p21.31	rs12205331	*ANKS1A*	0.81 (C)	1.04
Risk variants associated with triglycerides				
8q24.13	rs10808546	*TRIB1*	0.65 (A)	1.08 (1.04–1.12)
11q23.3	rs964184	*ZNF259, APOA5-A4-C3-A1*	0.13 (G)	1.13 (1.10–1.16)
12q24.12	rs3184504	*SH2B3*	0.44 (T)	1.13 (1.08–1.18)
10q24.32	rs12413409	*CYP17A1, CNNM2, NT5C2*	0.89 (G)	1.12 (1.08–1.16)
4q31.1	rs7692387	*GUCYA3*	0.81 (G)	1.13
15q26.1	rs17514846	*FURIN-FES*	0.44 (A)	1.04
9q34.2‡	rs579459	*ABO*	0.21 (C)	1.10 (1.07–1.13)
Risk variants (mechanism of risk unknown)				
9p21.3	rs4977574	*CDKN2A,CDKN2B*	0.46 (G)	1.25 (1.18–1.31) to
1q41	rs17465637	*MIA3*	0.74 (C)	1.20 (1.12–1.30)
10q11.21	rs1746048	*CXCL12*	0.87 (C)	1.33 (1.20–1.48)
2q33.1	rs6725887	*WDR12*	0.15 (C)	1.16 (1.10–1.22)

Chromosomal Location	SNP	Nearby Genes (Allele)	Risk Allele Frequency	Odds Ratio (95% CI)
6p24.1	rs12526453	*PHACTR1*	0.67 (C)	1.13 (1.09–1.17)
21q22.11	rs9982601	*MRPS6*	0.15 (T)	1.19 (1.13–1.27)
3q22.3	rs2306374	*MRAS*	0.18 (C)	1.15 (1.11–1.19)
10p11.23	rs2505083	*KIAA1462*	0.42 (C)	1.07 (1.04–1.09)
1p32.2	rs17114036	*PPAP2B*	0.91 (A)	1.17 (1.13–1.22)
5q31.1	rs2706399	*IL5*	0.48 (A)	1.02 (1.01–1.03)
6q23.2	rs12190287	*TCF21*	0.62 (C)	1.08 (1.06–1.10)
7q22.3	rs10953541	*BCAP29*	0.75 (C)	1.08 (1.05–1.11)
7q32.2	rs11556924	*ZC3HC1*	0.62 (C)	1.09 (1.07–1.12)
10q23.31	rs1412444	*LIPA*	0.34 (T)	1.09 (1.07–1.12)
11q22.3	rs974819	*PDGF*	0.29 (T)	1.07 (1.04–1.09)
13q34	rs4773144	*COL4A1, COL4A2*	0.44 (G)	1.07 (1.05–1.09)
14q32.2	rs2895811	*HHIPL1*	0.43 (C)	1.07 (1.05–1.10)
15q25.1	rs3825807	*ADAMTS7*	0.57 (A)	1.08 (1.06–1.10)
17p13.3	rs216172	*SMG6, SRR*	0.37 (C)	1.07 (1.05–1.09)
17p11.2	rs12936587	*RASD1, SMCR3, PEMT*	0.56 (G)	1.07 (1.05–1.09)
17q21.32	rs46522	*UBE2Z, GIP, ATP5G1, SNF8*	0.53 (T)	1.06 (1.04–1.08)
5p13.3*	rs11748327	*IRX1, ADAMTS16*	0.76 (C)	1.25 (1.18–1.33)
6p22.1*	rs6929846	*BTN2A1*	0.06 (T)	1.51 (1.28–1.77)
6p24.1†	rs6903956	*C6orf105*	0.07 (A)	1.65 (1.44–1.90)
6p21.3	rs3869109	*HCG27 and HLA-C*	0.60 (C)	1.15
1q21	rs4845625	*IL6R*	0.47 (T)	1.09

Chromosomal Location	SNP	Nearby Genes (Allele)	Risk Allele Frequency	Odds Ratio (95% CI)
Chr4	rs1878406	*EDNRA*	0.15 (T)	1.09
7p21.1	rs2023938	*HDAC9*	0.10 (G)	1.13
2p11.2	rs1561198	*VAMP5-VAMP8*	0.45 (A)	1.07
Chr2	rs2252641	*ZEB2-AC074093.1*	0.45 (A)	1
Chr5	rs273909	*SLC22A4-SLC22A5*	0.14 (C)	1.11
6p21	rs10947789	*KCNK5*	0.76 (T)	1.01
6q26	rs4252120	*PLG*	0.73 (T)	1.07
8p22	rs264	*LPL*	0.86 (G)	1.06
13q12	rs9319428	*FLT1*	0.32 (A)	1.1

A indicates adenine; C, cytosine; CI, confidence interval; G, guanine; HDL, high-density lipoprotein; LDL, low-density lipoprotein; OR, odds ratio; SNP, single-nucleotide polymorphism; and T, thymine.
* Variant identified only in Japanese.
† Variant identified only in Han Chinese.
‡ The risk variant at 9q34.2 is associated with myocardial infarction but not with coronary atherosclerosis

1.6 TRANSCRIPTOMIC ANALYSIS IN ATHEROSCLEROSIS

Unlike DNA sequence variation, which is normally fixed within an individual, there is tremendous variability in gene expression in different tissues and in response to stimuli. Transcriptomic is the quantitative study of all genes expressed in a given biological state, and it can provide important insights into pathological processes [56]. Recently, robust high throughput techniques such as gene expression microarray or RNA sequencing have become mainstream methods to quantify the abundance of all transcripts expressed in a tissue of interest under a given biological state [57].

A typical modern microarray consists of patches of DNA probes (11–50 bp) complementary to the transcripts whose presence is to be investigated, and immobilized on a solid substrate. Transcripts are extracted from samples (cells or tissues to be investigated), labelled with fluorescent dyes (either one colour or two), hybridized to the chip and scanned with a laser. Probes that correspond to transcribed RNA hybridize to their complementary target. Since transcripts are labelled with fluorescent dyes, light intensity can be used as a measure of gene expression [57,58].

Microarray analysis of RNA from animal models for dyslipidemia, atherosclerosis, and vascular disease, as well as cell culture microarray studies, have been particularly fruitful in identifying key genes and pathways involved in lipoprotein metabolism, atherosclerosis, and vascular disease [58].

The most commonly animals used for gene expression profiling studies in atherosclerosis, are mice, rats and pigs [43]. *In vivo* atherosclerosis experiments are conducted predominantly with mouse models because they are small, breed quickly, and are relatively inexpensive to maintain. However, mice have a different lipoprotein profile compared to humans and do not normally develop atherosclerosis. Some strains, when given a high-fat diet, can develop plaques, but most atherosclerotic mouse models are genetically engineered.

A common transgenic model is the apolipoprotein E knockout (Apoe-/-) mouse. Apoe is an apolipoprotein responsible for the lipoprotein uptake from the circulation by the LDL receptor. Without apoE, mice develop hypercholesterolemia and atherosclerotic lesions [59].

The best way to identify gene involved in the development of human atherosclerosis is to perform a transcriptomic analysis on entire vessel segments. However, a limitation of this approach is the lack of insight into the underlying reason for observed differences in gene expression levels. These differences may reflect the changed composition of the vessel wall during atherosclerosis (eg, thinning of the medial smooth muscle cell layer) and thus the presence of different cell types (eg, infiltration of T-lymphocytes), or reflect a change in the gene expression caused by a pro-atherosclerotic environment (eg, differentiation of macrophages to foam cells) [43].

To prevent the problem of analyzing tissues made up of multiple cell types, 3 different approaches have been used: macroodissection, laser capture and

isolation of pure cell line form the entire vessel wall. The macrodissection was used to separate the VSMC-rich fibrous cap, media, and non-atherosclerotic intima. Laser capture microdissection was used to dissect SMCs or macrophages from whole mount specimens for subsequent cell-specific RNA isolation [43].

In a recent study, laser capture microdissection technique has been used for expression profiling of neovascular region in human carotid plaque [60].

A drawback of both macrodissection and laser capture microdissection is the low tissue and RNA yield. To avoid this prolem is possibile pooling samples and/or use (several rounds of) amplification techniques. Pooling of samples limits the possibilities for downstream statistical analysis. Pros and cons of amplification have been discussed in several articles and include issues such as reproducibility and effects on magnitude of differential expression caused by amplification [43].

An alternative approach to obtain relatively pure cell populations to culture cells after isolation from entire vessel wall samples. This approach was recently used by Li et al., to identify miRNAs differentially expressed in cultured SMCs isolated from human aortic artery [61].

An advantage of this approach is that cell type-specific transcripts are amplified in culture omitting the need for pooling or amplification procedures. However, a disadvantage is that the in vitro culture may cause a shift in the transcriptome. As a result, the expression profiles of cultured cells may not be entirely representative of the in vivo situation [43].

Chapter II: Aim of the study

The principal aim of this study was to identify genetic and molecular signatures in atherosclerosis. The experimental design used an integrated approach, joining information from DNA, RNA and protein analyses.

The study at DNA level started from GWAS data and expanded upon the analysis of single nucleotide polymorphisms (SNPs) in a case-control study represented by subjects with or without CAD. At the RNA level, the transcriptome profile was analysed by microarray and qPCR methodologies, both in cells relevant in the atherosclerotic process (VSMCs), and in atherosclerotic and grossly non-atherosclerotic carotid tissues. The expression at protein level was evaluated by immunohistochemical analysis in carotid artery specimens.

The feasibility and the work of this study have been largely supported by well-established collaborations with i) the group of the University of Verona, headed by Professor Oliviero Olivieri, who provided us with data of a replication stage of GWAS –MIGen–, and the complete clinical characterization of CAD and CAD-free subjects; ii) the Unit of Vascular and Endovascular Surgery of Sant'Anna University-Hospital, of Ferrara, headed by Dott. Francesco Mascoli, who provided us with samples of atherosclerotic and non-atherosclerotic arterial walls from patients undergone to carotid endarterectomy and iii) the group of Marie Luce Bochaton Piallat of the Department of Pathology and Immunology (CMU) at the University of Geneva, for its expertise on immunohistochemical analysis in vascular samples.

Chapter III: Materials and Methods

3.1 STUDY POPULATION

3.1.1 Recruitment of patients and their characterization

This part of the study was performed within the Verona Heart Study (VHS), a regional survey designed for identification of new risk factors for CAD in subjects with objective angiographic documentation of their coronary vessels [62,63].
The VHS has been concepted and is still conducted by the research group of the University of Verona, headed by Professor Olivieri. Through the Verona Heart Study (VHS), this research group participated to the first GWAS on CAD and MI risk, a study coordinated by the MIGEN Consortium at BROAD institute (MIT/Harvad, Boston) [52].

Subjects with proven CAD had at least one of the main epicardial coronary arteries affected (left anterior descending, circumflex, or right) with ≥ 1 significant stenosis (≥50%). CAD patients were classified into MI and non-MI groups on the basis of a thorough review of medical records including history, electrocardiogram, enzyme changes, and/or the typical sequelae of MI on ventricular angiography. CAD-free subjects had completely normal coronary arteries, being submitted to coronary angiography for reasons other than CAD, mainly valvular heart disease. These controls were also required to have neither history nor clinical or instrumental evidence of atherosclerosis in vascular districts beyond the coronary bed. All participants came from the same geographical area (North-East Italy). At the time of blood sampling, a complete clinical history was collected, including the assessment of cardiovascular risk factors such as obesity, smoking, hypertension and diabetes.

The first study population was represented by 510 patients with a history of MI before the age of 65 years and 388 CAD-free subjects who were included in the MIGen Consortium as replication population of a GWAS [52].

The second study population analysis, addressing specifically the atherosclerotic (more than thrombotic) phenotype, included 442 CAD patients without MI history and 393 CAD-free subjects ***TABLE m-1***.

The study was approved by the Ethic Committee of Azienda Ospedaliera Universitaria Integrata, Verona, Italy. A written informed consent was obtained from all the participants after a full explanation of the study.

***TABLE m-1*. General characteristics of the study populations, with or without coronary artery disease (CAD).**

CHARACTERISTICS	CAD-free (n=393)	CAD (n=442)	P
Age, years	58.9 ± 12.1	63.1 ± 9.1	<0.001*
Male sex, %	62.8	72.9	0.002†
BMI, kg/m²	25.4 ± 3.5	26.8 ± 3.4	<0.001*
Hypertension, %	41.2	72.6	<0.001†
Smoking, %	41.9	61.3	<0.001†
Diabetes, %	7.1	20.4	<0.001†
Glucose, mmol/L	5.47 (5.37-5.58)	5.70 (5.58-5.81)	0.007 *
Creatinine, mmol/L	89.1 (86.5-90.9)	90.0 (88.2-91.8)	0.587 *
Lipid Lowering Therapies, %	5.9	31.6	<0.001†
Total cholesterol, mmol/L	5.44 ± 1.09	5.45 ± 1.15	0.904*
LDL-cholesterol, mmol/L	3.50 ± 0.93	3.62 ± 0.98	0.098
HDL-cholesterol, mmol/L	1.43 ± 0.42	1.22 ± 0.32	<0.001*
Triglycerides, mmol/L	1.35 (1.28-1.40)	1.68 (1.62-1.77)	<0.001*
high sensitivity- C Reactive Protein, mg/l	1.97 (1.72-2.27)	3.22 (2.83-3.67)	<0.001*

CAD patients were requested to have no history of previous myocardial infarction.
* by t-test.
† by chi-square test.

3.1.2 Biochemical analyses on blood samples

The biochemical analyses on blood samples have been performed by the group of professor Olivero Oliveri of University of Verona.
Samples of venous blood were withdrawn from each VHS subject, after an overnight fast. Plasma glucose, creatinine, triglycerides and total and high-density lipoprotein (HDL) cholesterol were determined by using a Technicon DAX 96 automated analyzer (Technicon Instruments, Tarrytown, NY). Low-density lipoprotein (LDL) cholesterol levels were calculated using the Friedewald formula. High-sensitivity C-reactive protein (hs-CRP) was measured by particle-enhanced nephelometric immunoassay (Siemens Healthcare Diagnostics, USA).

3.1.3 DNA analysis

Genomic DNA was prepared from whole blood samples by phenol-chloroform extraction.

The 91 SNPs investigated as replication stage of MIGen study, including intergenic rs10402271 polymorphism (*BCL3/PVRL2* locus), were genotyped using the iPLEX MassARRAY platform (Mass Array, Sequenom). PCR primers were designed by Sequenom Mass-Array-Assay-Design program, as described in [52].

The *BCL3* polymorphisms, rs2965169 (5' UTR, NM_005178.4:c.-892T>G) and rs8100239 (intron 1, NM_005178.4:c.256+801T>A) and the PVRL2 polymorphisms, rs3810143 (5' UTR, NM_001042724.1:c.-381T>C or NM_002856.2:c.-381T>C) and rs1871047 (intron 1, NM_001042724.1:c.88+1876A>G or NM_002856.2:c.88+1876A>G) were genotyped by allele-specific real-time PCR (TaqMan SNP Genotyping Assays, Applied Biosystems, Foster City, CA).
Each predesigned TaqMan SNP Genotyping Assay included two allele-specific TaqMan probes containing distinct fluorescent dyes (VIC 5' dye-labeled probe, FAM 5' dye-labeled probe) and a PCR primer pair to detect specific SNP targets. These TaqMan probe and primer sets uniquely align with the genome to provide unmatched specificity for the allele of interest. TaqMan probes incorporate MGB technology at the 3' end to deliver superior allelic discrimination. The MGB molecule binds to the DNA helix minor groove, improving hybridization based assays by stabilizing the MGB probe-template complex. If only the first probe's fluorophore

wavelength is detected during the assay, then the individual is homozygous to the wild type. If only the second probe's wavelength is detected, then the individual is homozygous to the mutant allele. Finally, if both wavelengths are detected, then both molecular beacons must be hybridizing to their complements and thus the individual must contain both alleles and be heterozygous.

One hundred ng of gDNA were amplified on CFX96 Real-Time PCR Detection System (BioRad, Hercules, CA) using Taqman Genotyping Master Mix (Applied Biosystems, Foster City, CA) in the following condiction: 95°C for 10 minutes, then 40 cycles of 15 seconds at 92°C and 1 minute at 60°C.

3.2 COLLECTION OF CAROTID SPECIMENS

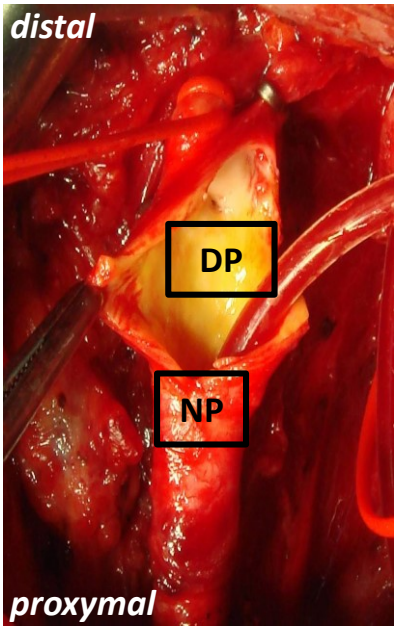
Carotid artery specimens were obtained from patients who underwent carotid endarterectomy (CEA) for extracranial high-grade (>70%) internal carotid artery stenosis [64]. Restenotic lesions were excluded. CEA were performed by the same surgeon at the Unit of Vascular and Endovascular Surgery of S. Anna University-Hospital (Ferrara, Italy) The study was approved by the Ethic Committee of the University-Hospital and written informed consent was obtained from all patients.

Arteriotomy was performed on the common carotid artery and extended to the internal carotid artery (***FIGURE m-2***). The CEA specimen was removed as a single piece and included the atheromatous area and two small adjacent areas, proximal and cranial respectively, without clear evidence of atherosclerotic lesions. CEA specimen was cut transversally at the bifurcation and the portion towards the aortic arch was used. This segment, consisting of the common carotid artery, was further cut into: a proximal small portion without evidence of atherosclerotic lesions (grossly non-atherosclerotic portion, NP), and the distal one (cephalad) characterized by diffuse atherosclerosis (diseased portion, DP).

Although the NP showed some histological alterations, like a thin thickened intima, this portion was the available and suitable endogenous control NP and DP portions were processed for i) histology-immunohistochemistry analysis, ii) total RNA extraction, iii) VSMCs cultures. The portions for total RNA isolation were

immediately placed into RNAlater (Ambion Inc., Austin, TX) and stored at 4°C overnight to allow the solution to thoroughly penetrate the tissue. The tissues were retrieved from RNAlater with sterile forceps, quickly dried from excess RNAlater with sterile gauze, and then moved in a new sterilize RNase-free eppendorf tubes and immediately stored at -80°C.

FIGURE m-2: **Macroscopic anatomical features of grossly non-atherosclerotic (NP) and diseased (DP) carotid portions.**



3.2.1 Histological and immunohistochemical analysis of carotid specimens

For histology, 3 µm thick sections of formaldehyde-fixed paraffin-embedded NP and DP were cutted and stained with Masson's trichrome.

Immunostaining for α-SMA, BCL3, and CD68 was performed on adjacent sections with: IgG2a recognizing α-smooth muscle actin (1: 100, α-SMA, clone 1A4) [65], IgG1 recognizing CD68 (1:200, clone KP1; Dako, Glostrup, Denmark) and IgG2a recognizing BCL3 (clone1E8, Abcam, Cambridge, UK). Before using the first antibody, immunoreactivity was intensified by pressure cooker (3 minutes) in citrate

buffer (10 mM, pH 6.0) for BCL3 or microwave treatment in citrate buffer (750 W, 5 minutes) for α-SMA and CD68. Goat anti-mouse or anti-rabbit biotinylated antibodies (Dako) were used as secondary antibodies. The presence of the specific proteins was revealed by means of the streptavidin-biotin peroxidase complex, and staining was visualized using either the 3,3′-diaminobenzidine (DAB).

Paraffin-embedded human tonsil sections were used as positive controls for BCL3 immunohistochemistry data. The optimal dilution for anti-BCL3 antibody (1:25) was obtained by testing a dilution curve set. All immunohistochemistry experiments included a negative control which received identical treatment with the exception of addition of the primary antibody. This ruled out the possibility of non-specific binding causing the generation of false positive results.

3.2.2 VSMC cultures

Primary VSMC cultures were obtained from NP (n=10) and DP (n=11) carotid specimens collected at surgery. The specimens retrieved at surgery were conserved in cold sterile Roswell Park Memorial Institute (RPMI) supplemented with 2% HEPES, 100 U/ml penicillin, 100 µg/ml streptomycin and 2 mM L-glutamine (Gibco-Invitrogen, Carlsbad, CA). Specimens were carried to the laboratory within 1 hour from surgical intervention. Specimens were washed several times with warm (37°C) phosphate buffered saline (PBS) to remove blood in excess.

After luminal gentle scraping to remove endothelial tissue, the NP and the DP were cut into 3x3mm pieces and plated on separate dishes. The abluminal side of the explants was carefully placed in contact with the culture dish (5-10 tissue pieces for the NP portion and 15- 20 tissue pieces for the DP portion per 60-mm dish). After 20 minutes, during which specimens dried, RPMI 1640 medium supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA), 100 U/ml penicillin, 100 µg/ml streptomycin, 400mM l-glutamine was added to petri dish. Tissue explanation were maintained at 37 °C in a humidified atmosphere of 95% air and 5% CO_2. Explanted tissues were removed 7 to 10 days after the first VSMCs appeared. VSMC populations from both NP and DP were maintained in identical baseline culture conditions and studied at the third passage.

Cultured VSMCs lineage was confirmed by immunofluorescence staining using a mouse IgG2a recognizing α-smooth muscle actin (α-SMA, clone 1A4) [65] and rabbit polyclonal IgGs recognizing both SMMHC types 1 and 2 (BT-562, Biomedical Technologies Inc, Stoughton, MA).

3.3 RNA EXPRESSION STUDY

3.3.1 RNA extraction

To isolate intact, high-quality RNA, it was essential that RNases were not introduced into RNA preparations. For this reason, sterile and RNases-free tips, tubes and solutions were used. Table surface, gloves and micropipettes were decontaminated with RNaseZap solution (Ambion Inc., Austin, TX).

Total RNA was isolated from either cultured VSMCs or tissue specimens.

Confluent cells were directly lysed in a culture dish by adding 1 ml of TRIzol Reagent (Invitrogen Carlsbad, CA) to a 3.5 cm diameter dish.

A more complex process was used for the lysis of the tissue samples. In fact, NP and DP specimens were homogenised directly in TRIzol Reagent (1 mL for 50-100 mg of tissue) 2-3 times 20 seconds using a disperser homogenizer (Ultra Turrax, IKA Labortechnik, Staufen, Germany) at maximum speed. The homogenization was conducted while keeping sample on ice and with intervals of 40 seconds (on ice), to avoid RNA degradation due to increase of temperature. To prevent cross contamination between samples, the dispersing element was serial rinsed with ethanol 70%, NaOH 1M, ethanol 70%, followed by 3 washes in DEPC-treated water (Diethylpyrocarbonate) and in the last decontaminated with RNaseZap solution (Ambion Inc., Austin, TX). Following homogenization, the insoluble material was removed from the homogenate by centrifugation at 12,000 g for 10 minutes at 4°C. The resulting pellet contains extracellular membranes, polysaccharides, and high molecular weight DNA, while the supernatant contains RNA. In some DP samples, an excess of fat collects as a top layer which should be removed. After we transferred the cleared homogenate solution in to a fresh tube.

From this step the procedure was the same for cells and tissues. 200 µL of chloroform for ml of TRIzol Reagent used for homogenization was added and the mixture was shaken vigorously by vortex for 15 s. The mixture was incubated for 5 min at room temperature and then centrifuged at 12,000 g for 15 minutes at 4°C to separate the 3 layers. The upper aqueous phase contains mostly RNA, and the interphase contains mostly DNA. The aqueous phase was collected in a fresh tube and a 0.5 mL of 100% isopropanol (per 1 mL of TRIzol Reagent used for homogenization) equal volume of cold isopropanol was added. To increase yield, this mixture was stored overnight at -80°C and then centrifuged at 12,000 g for 10 min at 4°C to precipitate the RNA. After discarded supernatant, the pellet was washed with 1ml of 70% cold ethanol RNase-free at 4°C and then dried under hood to ensure that no contaminating ethanol remained. DEPC-treated water was added (20-40 µL, depending on the size of the RNA pellet) to dissolving RNA.

3.3.2 RNA quality control

High-quality RNA is required for the downstream applications. Microarray experiments require RNA samples with specific purity and integrity values, while qPCR-analyses may accept samples with lower quality scores because the amplicons are small.

To evaluate the quality of the total RNA extracted from VSMCs and tissues a three steps analyses of quality control were used. The first step to determine the quality and the integrity of the isolated RNA was performed after each extraction via gross examination of the 18S and 28S ribosomal RNA bands on 1.5% (w/v) TAE (40 mM Tris (pH 7.6), 20 mM acetic acid, 1 mM EDTA) agarose gels containing 1% bleach [66]. If the 28S and 18S rRNA exhibited an expected near 2:1 ratio on ethidium bromide staining, this indicated that no gross degradation of RNA occurred.

The purity was evaluated by spectrophotometric analysis, by measuring the absorbance (A260/280 and A260/230 ratios) contributed by the nucleic acid to the absorbance of the contaminants. The accepted range for A260/A280 ratios is 1.8–2.2, which indicates the absence of protein, phenol or other contaminants that absorb at 280 nm. The 260/230 ratio was used as a secondary measure of nucleic

acid purity. The 260/230 values for “pure” nucleic acid is 1.8 or higher. Samples with 260/280 and 260/230 ratios out of the recommended ranges were excluded, because a significant presence of contaminants (phenol, ethanol, proteins and solvents, used in RNA extraction) could interfere with the downstream analyses, lowering their efficiency.

The last step to evaluate the integrity was by Agilent Bioanalyzer. The Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). This analysis was performed by loading samples onto a chip that contains micro channels that are filled with fluorescent dye and a gel- like matrix. By an electrophoretic separation, the sample move through the gel. The Agilent Bioanalyzer evaluates the ratio between the 18S and the 28S rRNA, as well as the presence of degraded short fragments. This is used to calculate a RIN value (RNA Integrity value), which is an evaluation of the intactness of the RNA sample. An intact total RNA sample would have a RIN value of 10. The RNA samples with low-quality RIN value (RIN < 7) were excluded from further analyses.

3.3.3 Microarray-based expression profiling of cultured VSMCs and of whole carotid artery specimens

The microarray profiling of cultured VSMCs and whole carotid artery were performed by the Microarray Facility at the Laboratory for Technologies of Advances Therapies (LTTA) of the University of Ferrara.

Labeled cRNA was synthesized from 0.5 µg of total RNA isolated from cultured VSMCs or from carotid tissues, using the Low RNA Input Linear Amplification Kit (Agilent Technologies, Palo Alto, CA) in the presence of cyanine 3-CTP (Perkin-Elmer Life Sciences, Boston, MA).

RNA samples from VSMCs were hybridized on Agilent whole human genome oligo microarray (Cat.No. G4112F, Agilent Technologies, Palo Alto, CA). This microarray consists of 60-mer DNA probes synthesized in situ, which represent 41,000 unique human transcripts. One-colour gene expression was performed according to the manufacturer’s procedure. Hybridizations were performed at 65°C for 17 hours in a rotating oven. Images at 5 um resolution were generated by Agilent scanner and

the Feature Extraction 9.5 software (Agilent Technologies) was used to obtain the microarray raw-data.

RNA samples from eight carotid specimens (DP=5 and NP =3) were hybridized on Agilent whole human genome oligo microarray (Cat.No. G4851A, Agilent Technologies) which represents 60,000 unique human transcripts. RNA labelling and hybridisation were performed in accordance to manufacturer's indications. Feature Extraction software v.10.7 (Agilent Technologies) was used to obtain the microarray raw-data.

Microarray results were analyzed using the GeneSpring GX software 7.3.1 (Agilent Technologies). Data files were pre-processed using the GeneSpring plug-in for Agilent Feature Extraction software results. Data transformation was applied to set all the negative raw values at 5.0, followed by on-chip median normalization. All microarray data have been registered in an official web-site of the University of Ferrara.

3.3.4 cDNA preparation and real-time quantitative polymerase chain reaction (qPCR*).*

cDNA was obtained from 1µg (cultured cells) or 0.5 µg (carotid tissues) of total RNA by reverse transcription using SuperScript VILO cDNA Synthesis Kit (Invitrogen, Carlsbad, CA) according to the manufacturer's recommendations.

Aliquots of diluted (1/10) first-strand cDNA were amplified on CFX96 Real-Time PCR Detection System (BioRad, Hercules, CA) using SsoFast EvaGreen Supermix (BioRad, Hercules, CA). PCR protocol was: 95°C for 30 seconds, then 40 cycles of 5 seconds at 95°C and 10 seconds at 60°C. Forward and reverse primers were designed by OligoCalc software [67] and are reported in **TABLE m-2**. Each reaction was performed in triplicate. The relative levels of mRNAs were calculated by the comparative CT method [68,69], with 18S rRNA as endogenous control for RNA from cultured cells or ACTB and B2M for RNA from carotid tissue. The values were expressed as mean fold change in DP- compared to NP- derived VSMCs (or DP compared to NP tissue) ± standard error of the mean.

TABLE m-2. Forward and Reverse Primers used for qPCR

GENE	SEQUENCE
ALS2CR3	5'- GCTCCTGTTCCAAAGAGTGCACTTAT-3' 5'-CGCGTGCTGCTACTTCTCTGTACA-3'
ACTB*	5'- CATCGAGCACGGCATCGTCA-3' 5'- TAGCACAGCCTGGATAGCAAC-3'
B2M*	*5'*- TTTCATCCATCCGACATTGA-3' 5'- TGTAAGCAGCATCATGGAGG-3'
BCL3	5'-TTTCCTCTGGTGAACCTGCCTACA-3' 5'-TACCCTGCACCACAGCAATATGGA -3'
GATAD2A	5'-CACCCAAACTGCAGAACTCAGCCTC-3' 5'-CTGACAAACGCAAGGGTCCCGC-3'
HAX1	5'-TGAGGGCCGGACAGAGACTACAGTAA-3' 5'-TTAATGGTGGGCAATGGGTGAGAGGTGG-3'
PVRL2	5'-TGAGCAGGAAATGCCTCGATACCA-3' 5'TGTCCAGATACTCTTCCTCCTCCT-3'
SARS	5'-CCGCCAGGACTGCAAGAACTGATC-3' 5'-AAATGAAAGCCTGGCAAATAGGGAGG-3'
18SrRNA*	5'- GTAACCCGTTGAACCCCATT-3' 5'-CCATCCAATCGGTAGTAGCG-3'

* endogenous controls.

3.4 STATISTICAL ANALYSIS

Differences in mRNA expression levels between VSMC populations by microarray-based transcriptome analysis were evaluated by using the log-transformed (log_2) ratio values (DP/NP) and considering that the log ratio should be 0 when there is no difference between conditions [70]. A P-value <0.05 was considered significant without correction for multiple testing. Relationships between the fold-change (log_2 DP/NP ratio) and the P-value (negative log_{10}) were reported in a volcano plot (see results ***FIGURE r-3***).

In tissue microarray a filter on low gene expression was used to keep only the probes expressed in at least one sample (flagged as Marginal or Present). Then, samples were grouped in accordance to their disease status (DP and NP) and compared. Differentially expressed genes were selected as having a 1.5-fold expression difference between their geometrical mean in two or more groups of interest and a statistically significant p-value (<0.05) by ANOVA (analysis of variance) statistic, followed by the application of the Benjamini and Hoechberg correction for false positives reduction. Differentially expressed genes were employed for Cluster Analysis of samples, using the Manhattan correlation as a measure of similarity.

Gene expression levels between VSMC populations or carotid portions in q-PCR analysis were compared by means of paired or unpaired t-test.

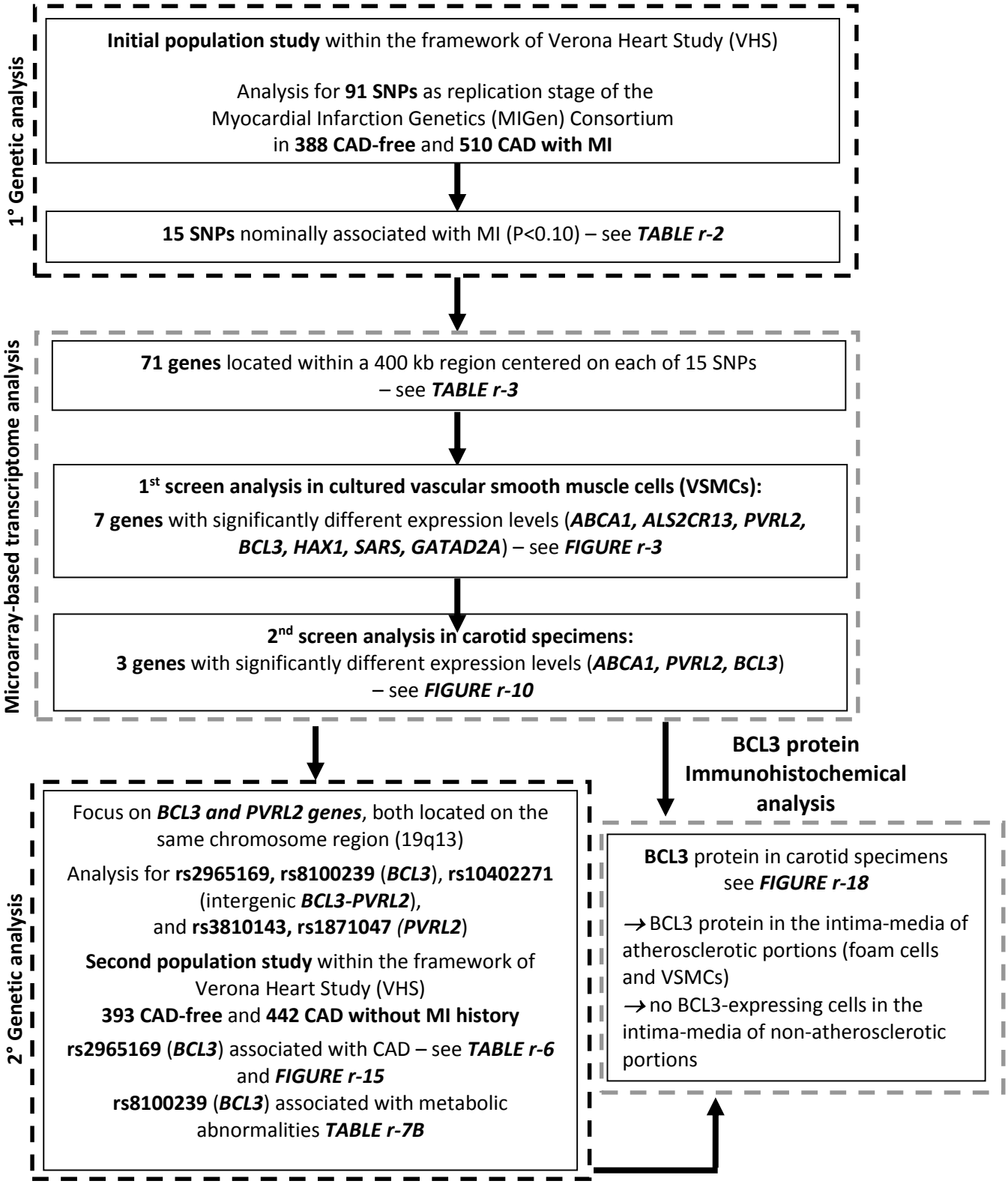
As regards population data, distributions of continuous variables in groups were expressed as means ± standard deviation. Statistical analysis on skewed variables, like glucose, creatinine, triglyceride, and high sensitivity - C reactive protein (hs-CRP), was computed on the corresponding log-transformed values. Thus, results are reported as geometric mean with 95% confidence interval (95% CI). Quantitative data were assessed using the Student's t-test or analysis of variance (ANOVA), with polynomial contrast for linear trend when indicated. Qualitative data were analyzed with either χ^2-test or χ^2 for linear trend analysis when indicated. Within each group examined, the frequencies of the genotypes associated with each of the polymorphisms were compared by the χ^2-test with the values predicted on the basis of the Hardy-Weinberg equilibrium (by Cubex analysis program online available).

In the first case-control population the strength of association with CAD/MI was assessed by means of a sex- and age-adjusted model. In the second case-control population the strength of association with CAD was estimated calculating the odds ratios (ORs) with 95% confidence intervals (CIs) by multiple logistic regression after sex- and age-adjustment and then after adjustment for all the traditional cardiovascular risk factors (i.e. sex, age, body mass index, smoking, hypertension, diabetes, LDL- and HDL-cholesterol, triglycerides, creatinine, and hs-CRP). A value of $P<0.05$ was considered statistically significant.

Chapter IV:

Results

An overall schematic representation of the steps of the present study is shown below, ***FIGURE r-1***. The results are reported in accordance with this flowchart.



4.1 FIRST GENETIC ANALYSIS

The initial population study was conducted in the VHS population within replication stages of MIGen Consortium GWAS. Genotyping was conducted in 510 CAD patients with MI history and in 388 controls without CAD. A total of 91 SNPs, intergenic, intronic and exonic, were investigated by Verona research group. The SNPs are reported in ***TABLE r-1*** (at the end of Results).

Starting from this list of SNPs, 15 SNPs, showing nominally an association with MI by means of a sex- and age-adjusted model with P values arbitrarily set at <0.10, were selected (***TABLE r-2***).

***TABLE r-2*. Fifteen SNPs nominally associated with CAD (P<0.1)**

SNP	Chr	SNP type	*P* *
rs646776	1p13	intergenic	0.001
rs1122608	19p13	intronic	0.005
rs3786722	19p13	intronic	0.005
rs3764261	16q13	intergenic	0.014
rs1333049	9p21	intergenic	0.016
rs8192284	1q21	coding	0.018
rs4977574	9p21	intergenic	0.019
rs4607103	9p14	intergenic	0.023
rs17482753	8p22	intergenic	0.041
rs16996148	19p13	intergenic	0.042
rs3890182	9q31	intronic	0.044
rs328	8p21	coding	0.065
rs10402271	19q13	intergenic	0.068
rs16988929	20q12-13	intronic	0.078
rs6725887	2q33	intronic	0.084

As more than half of the selected SNPs (8/15) localize within intergenic regions, the genomic regions tagged by the 15 SNPs were examined in the NCBI database for the presence of validated or putative coding gene sequences within a region of 400 kb centred on each SNP. For example, the region tagged by the intergenic rs10402271 on chromosome 19q13 includes 14 genes (***FIGURE r-2***).

The search for each of the 15 SNPs identified 77 neighbouring genes, listed in the ***TABLE r-3***.

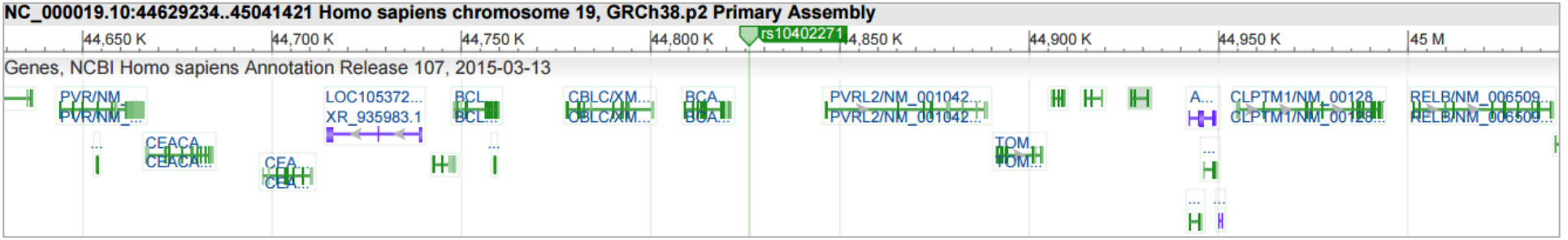


***FIGURE r-2*. Genes present in the ~400 kb region on chromosome 19q13 tagged by the intergenic rs10402271 (NCBI, SNPs database)**

***TABLE r-3*. List of genes located within a 400 kb-region centred on each selected SNP (n=15) associated with MI (P<0.10)**

SNP	Chr	Closest Genes (- upstream, +downstream) *
rs646776	**1p13**	*SYPL2*†(+200kb), *PSMA5*(+126 kb), *SORT1* (+33 kb) *MYBPHL* † (+16 kb), *PSRC1* (+0.6 kb), *CELSR2* (-0.2 kb), *SARS* (-38 kb), *KIAA1324* (-73 kb), *C1orf194* (-162 kb),*TMEM 167* † (-179 kb), *TAF13* (-200kb)
rs1122608 ‡	**19p13**	*DOCK6* (+147 kb), *ANKRD25* (+110 kb), *SPBC24* (+93 kb)
rs3786722 ‡		*LDLR* (+27 kb), *SMARCA4* (0 kb), *YIPF2* (-120 kb), *CARM1*† (-130 kb)
rs3764261	**16q13**	*CPNE2* (+133 kb), *NLRC5* (+61 Kb), *CETP* (+2 kb) *HERPUD1* (- 16 kb), *SLC12A3* (- 46 kb), *NUP93*(-115kb)
rs1333049§	**9p21**	*CDKN2B* (-116 kb), *CDKN2A* (-156 kb)
rs4977574§	**9p21**	*CDKN2B* (-89 kb), *CDKN2A* (-130 kb)
rs8192284	**1q21**	*ADAR* (+127 kb), *CHRNB2* (+113 kb), *TDRD10* (+48 kb) *SHE* (+25 kb), *IL6R* (0 kb), *ATP8B2* (-129 kb) *AQP10* (-129 kb), *HAX1* (-179 kb), *UBAP2L* (-184 kb)
rs4607103	**9p14**	*ADAMTS9* (-38 kb)
rs17482753	**8p22**	*SLC18A1* (+178 kb), *LPL* (-8 kb), *INTS10* (-123 kb)
rs328	**8p22**	*LPL* (0 kb)
rs16996148	**19p13**	*ZNF14* (+160 kb), *ZNF101* (+119 kb), *ATP13A1* (+95 kb), *GMIP* (+ 82 kb), *LPA2* (+76 kb), *PBX4* (+16 kb), *CILP2* (- 2 kb), *GATAD2A* (- 41 kb), *KIAA0892* (-191 kb)
rs3890182	**9q31**	*ABCA1*(0 kb), *NIPSNAP3B* (-111 kb), *NIPSNAP3A*(-112 kb) *OR13D1* (-190 kb)
rs10402271	**19q13**	*RELB* (+175 kb), *CLPTM1* (+129 kb), *APOC2* (+120 kb), *APOC4* (+116 kb), *APOC1* (+88 kb), *APOE* (+80 kb), *TOMM40* (+65 kb), *PVRL2* (+20 kb), *BCAM* (- 4 kb), *CBLC* (-26 kb), *BCL3* (-66 kb), *CEACAM16*†(-116 kb) *CEACAM19* (-141 kb), *PVR* (-162 kb)

SNP	Chr	Closest Genes (- upstream, +downstream) *
rs16988929	**20q12-13**	*HNF4A* (+80 kb), *R3HDML* (+61 kb), *C20orf142* (+31 kb) *GDAP1L1* (0 kb), *JPH2* (-88 kb), *TOX2* (-206 kb)
rs6725887	**2q33**	*NBEAL1*†(+134 kb), *ALS2CR8* (+30 kb), *WDR12* (0 kb), *ICA1L* (-41 kb), *ALS2CR13* (-112 kb)

* Approximate distance (kb) of genes from single SNPs.
‡ The rs1122608 and the rs3786722 are both intragenic (SMARCA4).
§CDKN2A and CDKN2B are in the 400 kb window of both rs4977574 and rs1333049.
‖ Closest genes to the intragenic rs328 are the same as for the rs17482753.
† genes for which probes were not present in the Agilent platform used for the microarray-based transcriptome analysis

4.2 MICROARRAY-BASED TRANSCRIPTOME ANALYSES

4.2.1 First transcriptome analysis in cultured vascular smooth muscle cell

To move towards the investigation of a potential association of the 77 genes with the atherosclerotic process, we evaluated the expression of these genes in VSMCs, which play a pivotal role in both the development and progression of atherosclerosis. For this purpose, we analyzed data stemming from a transcriptomic analysis, previously conducted in our laboratory, on VSMCs populations cultured from NP:DP couples of carotid wall specimens, collected from three patients who underwent to thromboendarterectomy. This microarray-based transcriptome analysis represented a preliminary step to identify, among the 77 genes, candidates for further evaluation.

The Agilent whole human genome microarray (44k), used for this expression profile analysis, included probes for 71 of the 77 neighbouring genes.

From the microarray normalized raw-data the mean expression values in DP and in NP were calculated for each gene. The log-transformed (log_2) ratio values (mean DP/mean NP) were used to calculate differences in mRNA expression levels, assuming the log ratio equal to 0 when there is no difference of expression [70]. A positive log_2 DP/NP ratio thus indicates increased expression levels in DP- than NP-derived VSMCs, while a negative log_2 DP/NP ratio indicates lower expression levels in DP- than NP-derived VSMCs. Expression levels of the 77 genes, are reported in the ***TABLE r-4*** and in a volcano plot that shows the relationships between the fold-change and the P value (***FIGURE r-3***).

From this analysis 7 out the 77 genes (*ABCA1, ALS2CR13, PVRL2, BCL3, HAX1, SARS, GATAD2A*, rombs over the dotted line of volcano plot and bold type in ***TABLE r-4***), showed significant difference ($P<0.05$) in expression levels between DP- and NP-derived VSMCs.

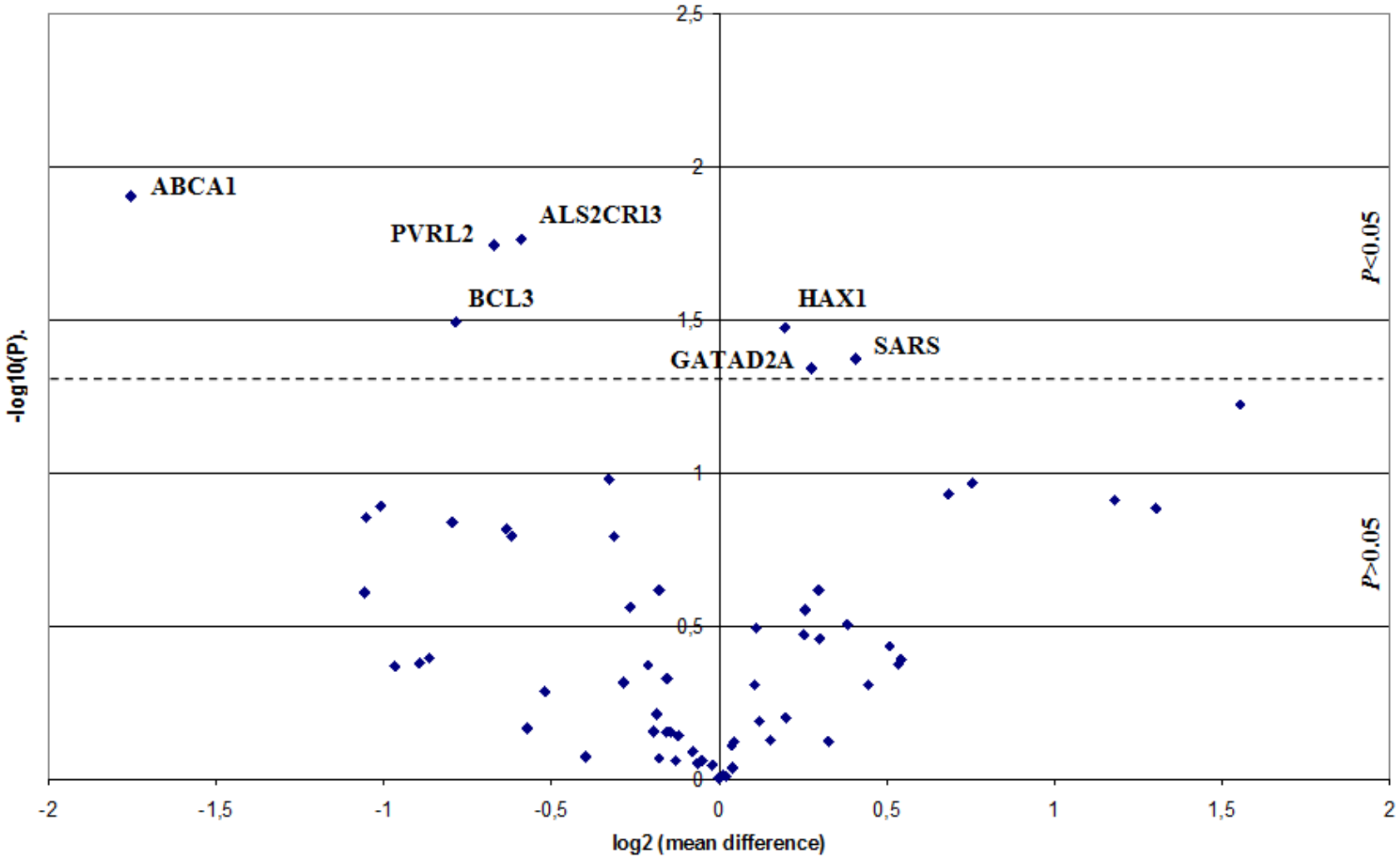


***FIGURE r-3*. Volcano plot of microarray data, reporting differences in expression at RNA level for the 71 considered genes in VSMC populations cultured from carotid artery diseased portions (DP) versus grossly non-atherosclerotic portions (NP).**
The graph shows the relationships between the fold-change (expressed as $\log_2$ of DP/NP mean ratio) on X-axis and the P-values (expressed as negative $\log_{10}$) on Y-axis for each gene (rhombs). Genes showing significant difference are reported and are represented over the dotted line (P<0.05). A positive log2 DP/NP ratio indicates increased expression levels in DP- than NP-derived VSMCs, while a negative log2 DP/NP ratio indicates lower expression levels in DP- than NP-derived VSMCs

***TABLE r-4*. Expression levels of the 71 considered genes in VSMC populations from DP and NP carotid portions**

Genes	Expression Analysis log_2 DP/NP ratio			*P**
	mean ratio	95% CI		
PSMA5	0	-0.167	0.167	0.999
SORT1	-0.020	-0.615	0.576	0.900
PSRC1	-0.001	-1.429	1.426	0.997
CELSR2	-0.187	-1.544	1.169	0.613
SARS	0.407	0.034	0.779	**0.042**
KIAA1324	-0.634	-1.844	0.574	0.152
C1orf194	0.535	-1.755	2.825	0.421
TAF13	-0.130	-3.183	2.924	0.872
DOCK6	0.014	-1.870	1.899	0.977
ANKRD25	0.106	-0.440	0.651	0.493
SPBC24	0.300	-0.761	1.360	0.348
LDLR	-0.266	-1.033	0.502	0.275
SMARCA 4	-0.212	-1.129	0.705	0.424
YIPF2	-0.144	-1.544	1.255	0.701
CPNE2	0.685	-0.426	1.795	0.118
NLRC5	-0.065	-1.847	1.717	0.890
CETP	-0.866	-4.404	2.673	0.403
HERPUD1	-0.314	-0.934	0.306	0.161
SLC12A3	0.038	-0.466	0.541	0.778
NUP93	0.295	-0.477	1.068	0.242
CDKN2B	-0.573	-5.783	4.637	0.683
CDKN2A	-1.011	-2.741	0.720	0.129
ADAR	0.111	-0.255	0.477	0.321
CHRNB2	0.152	-1.604	1.909	0.745
TDRD10	-0.399	-8.148	7.351	0.845
SHE	0.043	-0.478	0.565	0.755
IL6R	0.039	-1.417	1.495	0.918
ATP8B2	0.199	-1.321	1.719	0.630

Genes	Expression Analysis log_2 DP/NP ratio			*P**
	mean ratio	95% CI		
AQP10	-1.058	-3.864	1.748	0.246
HAX1	0.197	0.038	0.356	**0.034**
UBAP2L	0.256	-0.497	1.010	0.281
ADAMTS9	-0.797	-2.267	0.674	0.145
SLC18A1	0.038	-0.466	0.541	0.778
LPL	0.038	-0.466	0.541	0.778
INTS10	0.002	-0.663	0.667	0.991
ZNF14	-0.052	-1.290	1.185	0.872
ZNF101	-0.618	-1.838	0.601	0.161
ATP13A1	0.383	-0.850	1.618	0.313
GMIP	-0.285	-1.726	1.155	0.484
LPA2	0.021	-2.818	2.859	0.978
PBX4	-0.196	-2.086	1.693	0.699
CILP2	1.303	-0.953	3.559	0.131
GATAD2A	0.275	0.013	0.538	**0.046**
KIAA0892	-0.122	-1.396	1.152	0.721
ABCA1	-1.755	-2.608	-0.902	**0.013**
NIPSNAP3B	-0.328	-0.827	0.169	0.105
NIPSNAP3A	-0.157	-1.682	1.367	0.701
OR13D1	0.038	-0.466	0.541	0.778
RELB	-0.179	-3.959	3.600	0.857
CLPTM1	0.543	-1.699	2.785	0.407
APOC2	-0.968	-5.206	3.270	0.429
APOC4	-0.894	-4.697	2.910	0.418
APOC1	1.554	-0.161	3.269	0.060
APOE	0.755	-0.410	1.920	0.108
TOMM40	0.510	-1.395	2.414	0.368
PVRL2 †	-0.671	-1.065	-0.277	**0.018**
BCAM	0.252	-0.618	1.123	0.338
CBLC	-0.520	-3.403	2.364	0.519

Genes	Expression Analysis log_2 DP/NP ratio			*P**
	mean ratio	95% CI		
BCL3	-0.786	-1.410	-0.163	**0.032**
CEACAM19	0.120	-0.841	1.080	0.645
PVR	0.011	-1.071	1.093	0.969
HNF4A	0.038	-0.466	0.541	0.778
R3HDML	0.038	-0.466	0.541	0.778
C20orf142	-0.180	-0.650	0.290	0.242
GDAP1L1	0.444	-1.848	2.735	0.492
JPH2	0.325	-3.571	4.222	0.754
TOX2	-1.054	-2.958	0.850	0.140
ALS2CR8	-0.080	-1.367	1.207	0.814
WDR12	-0.155	-0.909	0.599	0.470
ICA1L	1.180	-0.789	3.148	0.123
ALS2CR13	-0.590	-0.928	-0.252	**0.017**

Expression levels in VSMC populations from diseased (DP) and grossly non-atherosclerotic (NP) carotid portions are reported as log_2DP/NP mean ratio with 95% CI. Genes with significant differences (P<0.05) in expression levels are reported in bold type. A positive log_2 DP/NP ratio indicates increased expression levels in DP- than NP-derived VSMCs, while a negative log_2 DP/NP ratio indicates decreased expression levels in DP- than NP-derived VSMCs. * by t-test † Data refer to the oligonucleotide probe located in exon 9.

Differences in expression levels for the 7 genes detected by microarray analysis were then evaluated by qPCR in a total of 20 VSMCs populations, both couples and non-coupled. These populations were independent from those analysed by microarray. As general approach in qPCR, for each gene one of the two primers was designed in the same exon recognized by the probe used in the microarray, in order to analyse approximately the same region tested by the microarray.

The results of the qPCRs from couples or in independent DP and NP populations are reported in ***FIGURE r-5*** and in ***FIGURE r-6***.

For both *ABCA1* and *BCL3* significant differences in expression levels between DP and NP populations were observed either in couples or in independent populations.

ABCA1 mRNA levels were higher in DP- than in NP – derived VSMCs populations both in 3 couples (mean fold-change, 2.73 ±0.35, P= 0.0397) and in 15 independent populations (mean fold-change, 3.70 ±0.66, P= 0.001).

BCL3 mRNA levels were lower in cell populations obtained from DP than in those from NP, by comparison of 5 couples (mean fold-change, 0.58 ± 0.09, P=0.009) as well as by comparison of 15 independent DP and NP populations (mean fold-change, 0.60 ± 0.08, P=0.010).

HAX1, *SARS* and *GATAD2A* were up regulated in DP than in NP populations both in 3 couples (*HAX1* FC 2.39± 0.78, P=0.22; *SARS* FC 2.35±0.66 P=0.18; *GATAD2A* FC 1.16±0.49, P= 0.78) and in 15 independent populations (*HAX1* FC 1.87 ± 0.38, P=0.057; *SARS* FC 2.03 ± 0.42, P=0.042; *GATAD2A* FC 1.15 ± 0.18, P=0.595)

For *ALS2CR13* mRNA levels were significantly lower in independent cell populations obtained from DP than in those from NP (mean fold-change 0.54 ± 0.10, P= 0.047). However, this difference was not significantly reproduced in 3 couples of VSMCs (mean fold-change 0.94 ± 0.38, P= 0.88).

Only for *PVRL2,* the DP/NP expression levels were discordant between the couples (mean fold-change 1.43± 0.27, P= 0.33) and the independent populations (mean fold-change 0.73± 0.16, P= 0.18) and significant differences were not detected.

Overall the qPCR analysis showed that the DP/NP ratios for *BCL3*, *HAX1*, *SARS*, *ABCA1* and *ALS2CR13* substantially confirmed microarray data. Differently, the slightly higher expression of *GATAD2A* mRNA in DP populations in microarray analysis was not significantly reproduced. Concerning *PVRL2*, expression levels evaluated by microarray were not confirmed by qPCR either in VSMC populations from DP:NP couples or in independent DP and NP populations (***TABLE r-5***).

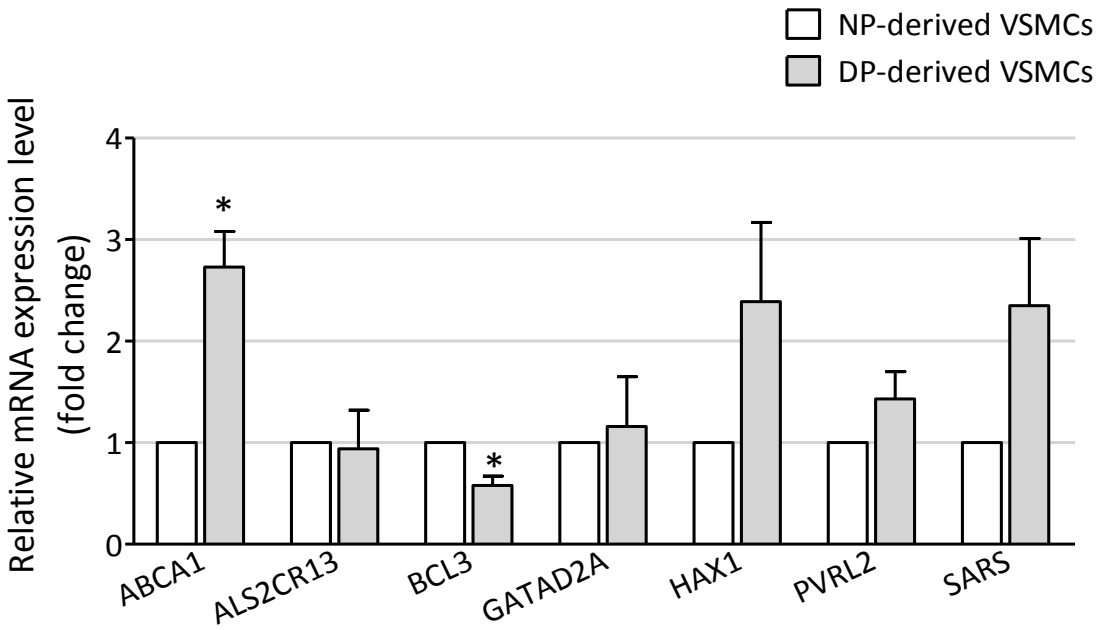


***Figure r-5*. Expression levels by qPCR analysis in couples of NP- and DP- derived VSMCs (5 patients).** 18S was used as reference gene. mRNA expression levels in NP-derived VSMCs cells are set to one as control. Data are reported as mean fold change ±SEM.

* P < 0.05 by paired t-test

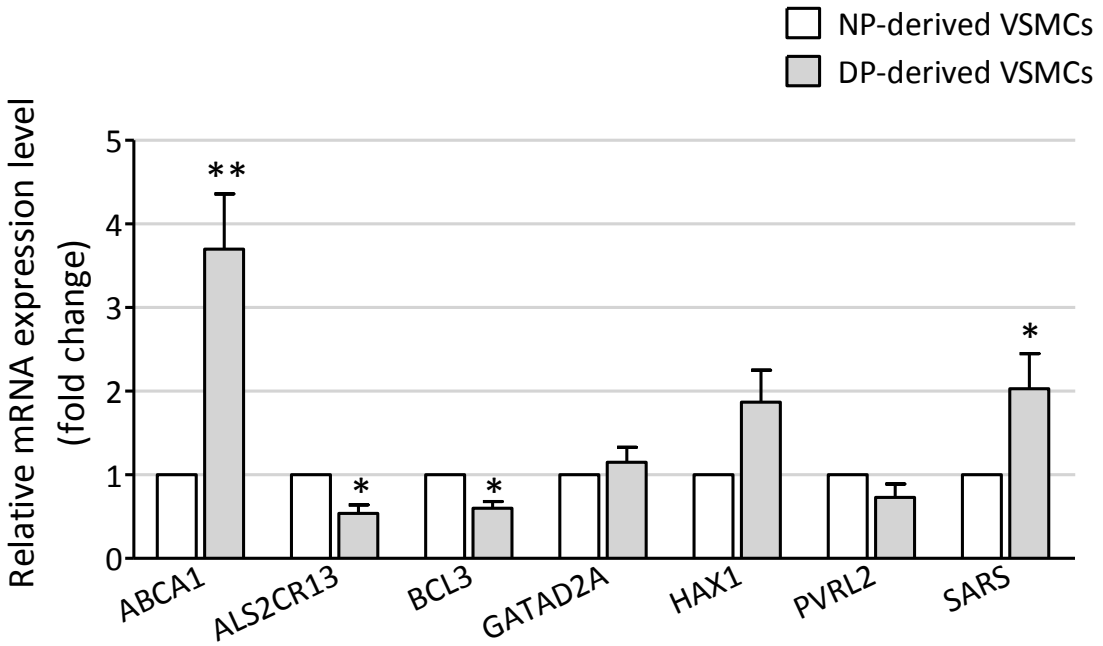


***Figure r-6.* Expression levels by qPCR analysis in independent DP and NP populations (15 patients).** 18S was used as reference gene. mRNA expression levels in NP-derived VSMCs cells are set to one as control.
Data are reported as mean fold change ±SEM.

* P < 0.05 by unpaired t-test ** P < 0.001 by unpaired t-test

TABLE r-5. **Expression levels of the 7 considered genes in VSMC populations from DP and NP carotid portions in microarray and in qPCR analyses.**

Genes	Official Full Name	Microarray Expression Analysis log2 DP/NP ratio		qPCR Expression Analysis			
				coupled DP:NP VSMCs		independent DP and NP VSMCs	
		mean ratio	P*	FC ± SEM	P**	FC ± SEM	P*
SARS	seryl-tRNA synthetase	0.28264	**0.042**	2.35±0.66	0.18	2.03±0.42	**0.042**
HAX1	HCLS1 associated protein X-1	0.13681	**0.034**	2.39±0.78	0.22	1.87±0.38	0.057
GATAD2A	GATA zinc finger domain containing 2A	0.19097	**0.046**	1.16±0.49	0.78	1.15±0.18	0.595
ABCA1	ATP binding cassette subfamily A member 1	-1.755	**0.013**	2.73±0.35	**0.0397**	3.70±0.66	**0.001**
PVRL2 †	poliovirus receptor-related 2	-0.671	**0.018**	1.43±0,27	0.33	0.73±0.16	0.18
BCL3	B-cell CLL/lymphoma 3	-0.786	**0.032**	0.58±0.09	**0.009**	0.60±0.08	**0.01**
ALS2CR13	family with sequence similarity 117 member B	-0.590	**0.017**	0.94±0.38	0.88	0.54± 0.10	**0.047**

Microarray expression levels are reported as log_2DP/NP mean. A positive log_2DP/NP ratio indicates increased expression levels in DP- than NP-derived VSMCs, while a negative log_2DP/NP ratio indicates decreased expression levels in DP- than NP-derived VSMCs.
† Data refer to the oligonucleotide probe located in exon 9.

qPCR values are expressed as fold change in DP- compared to NP-derived VSMCs ± standard error. A FC value > 1 indicated increased expression levels in DP- than NP-derived VSMCs, while a FC < 1 indicates decreased expression levels in DP- than NP-derived VSMCs.
Genes with significant differences (P<0.05) in expression levels are reported in bold type
* by unpaired t-test; ** by paired t-test

4.2.2 Second transcriptome analysis in whole carotid specimens

To explore the expression levels of the 7 selected genes (*ABCA1*, *ALS2CR13*, *PVRL2*, *BCL3*, *HAX1*, *SARS*, *GATAD2A*) in the carotid artery wall, from which VSMC populations were isolated, a second microarray based-transcriptome analysis was performed on whole atherosclerotic (DP) and grossly non-atherosclerotic (NP) carotid specimens.

Total RNA was extracted from DP and NP specimens (N= 72) retrieved at surgery from 36 patients and the RNA quality was evaluated by bleach agarose gel electrophoresis (***FIGURE r-7***), and by the OD 260/280 nm and OD 260/A230nm ratios (NanoDrop Spectrophotometer). The RNA integrity was investigated by assessing the RIN score, using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA).

RNA from 42 out of 72 tissue samples failed the quality control either due to poor integrity or insufficient quantity. The 30 RNA samples that satisfied the purity criteria (A260/A280 ratio 1.8–2.2; A260/A230 ratio >1.8) were analysed by microarray or qPCR.

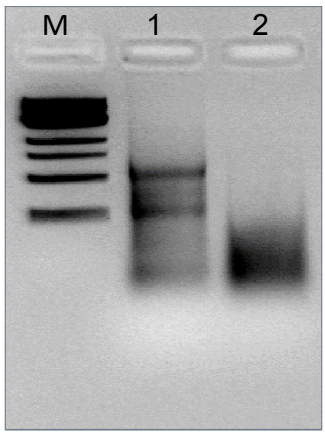


FIGURE r-7. **A representative image of two sample of total tissue RNA analyzed on bleach agarose gel electrophoresis.** Line 1, the bands corresponding to 28 S and 18S are clearly detectable. Line 2, completely degraded RNA. M= MARKER 1 Kb Biolabs

***FIGURE r-8*. Bioanalyzer Agilent report of 8 RNA sample (extracted from 5 DP and from 3 NP carotid specimens) evaluated in microarray analysis**

Eight total RNA samples, from 3 DP:NP couples and 2 independent DP (***FIGURE r-8***), were hybridized on a whole human genome microarray (Agilent Technologies, Palo Alto, CA, USA) which represents 60k unique transcripts.

To evaluate similarities or differences among each group (NP and DP) principal component analysis (PCA) was performed on the raw normalized data. All the three NP samples clustered in a separate area of the PCA plot (***FIGURE r-9***), while two out of 5 DP samples (red circles in ***FIGURE r-9***) grouped with the NP samples. These two DP samples were excluded from the analysis of microarray data, which therefore includes data from 2 paired DP:NP and 2 independent DP and NP samples (DP= 3 and NP=3).

A filter on fold change (FC >1.5) and a moderate t test were used to identify significantly modulated genes with P<0.05 and false discovery rate (FDR)= 10%.

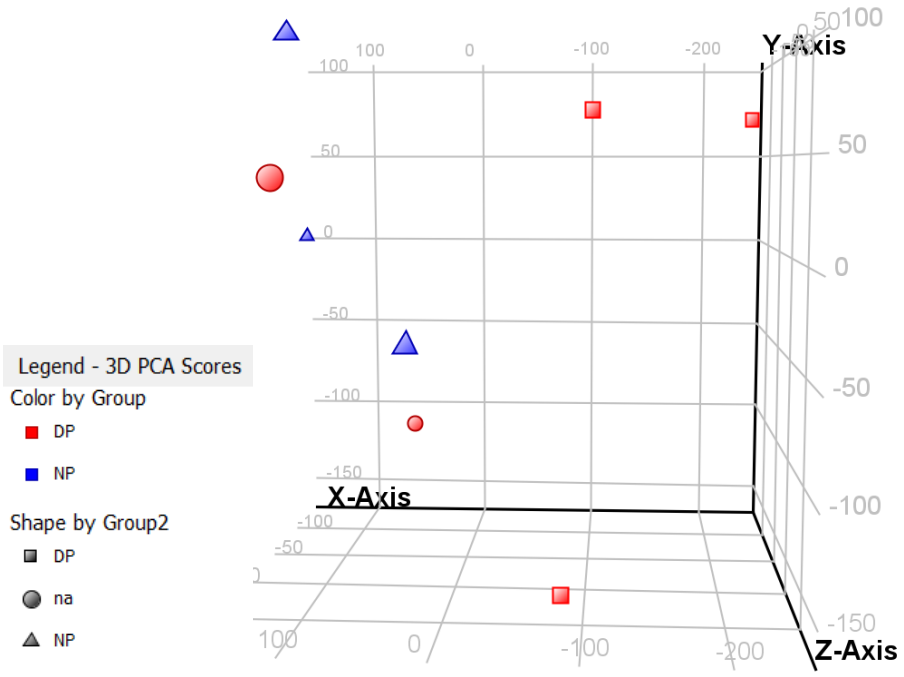


***FIGURE r-9*. Principal component analysis (PCA) between NP and DP samples of microarray.** The PCA plot shows a 3D visualization of the relationships between the samples, which is based on the expression levels of the probe sets. Different sample types are color-coded: samples labelled red are DP and blue are NP. Two out of 5 DP samples (red circles) overlap with the NP samples.

From this analysis a large number of genes (marked by 6076 probes) was identified as differentially expressed between DP and NP tissues. Among the seven genes (*ABCA1, ALS2CR13, PVRL2, BCL3, HAX1, SARS, GATAD2A*) identified in the first transcriptome screen as differentially expressed between cultured VSMC populations, *ABCA1, BCL3* and *PVRL2* were significantly up regulated in DP tissues (***FIGURE r-10***).

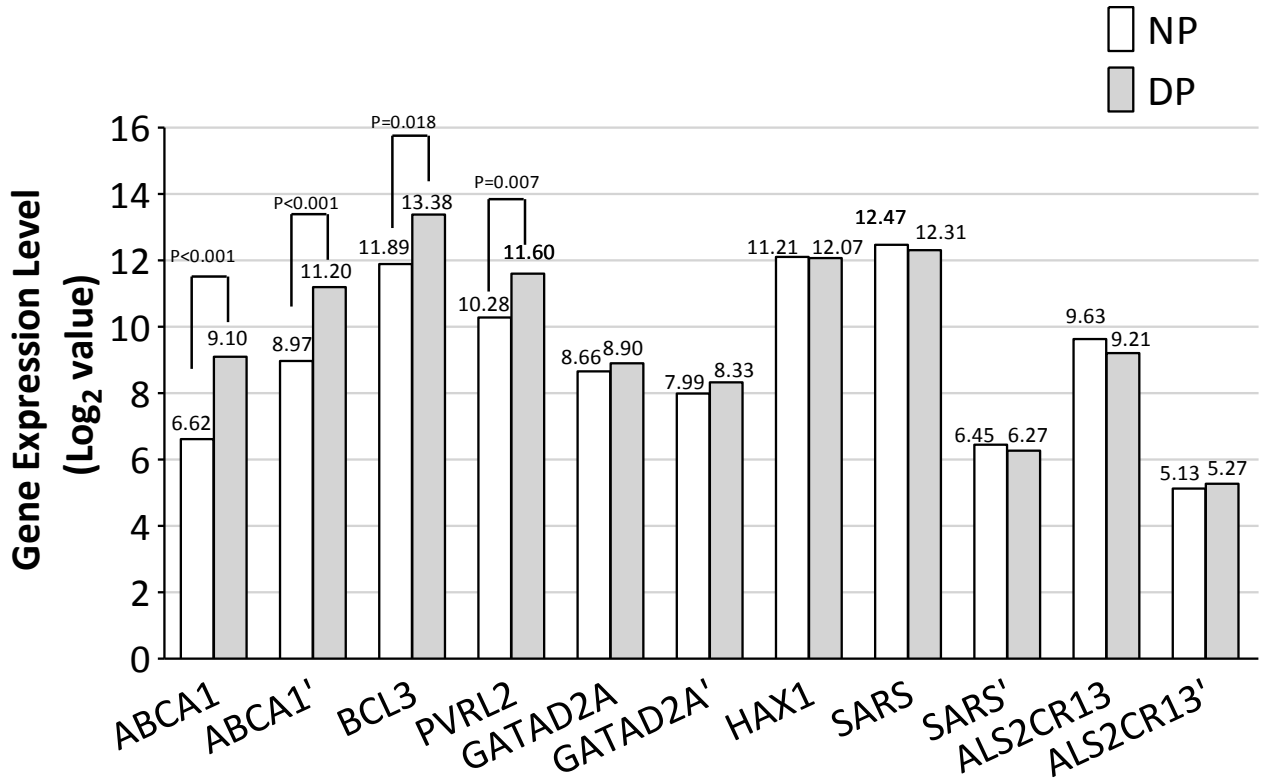


***FIGURE r-10*. Tissue mRNA expression levels by microarray profiling of atherosclerotic diseased (DP) and non-atherosclerotic (NP) portions of carotid artery.** The expression levels of the seven candidate genes, stemming from the preliminary screen in VSMC populations, are reported as $\log_2$ value.
Data obtained with two probes, present in the Agilent microarray, are reported for *ABCA1*, *SARS*, *GATAD2A* and *ALS2CR13*.

Expression data stemming from this second microarray were also investigated for the 77 neighbouring genes (***TABLE r-3***) tagged by the 15 SNPs nominally associated with CAD-MI. Twenty-three genes (27 probes) were found differentially expressed between DP and NP tissues (***FIGURE r-11***). Among them, *ABCA1*, *PVRL2* and *BCL3* were also modulated in VSMCs, as detected by the first microarray analysis.

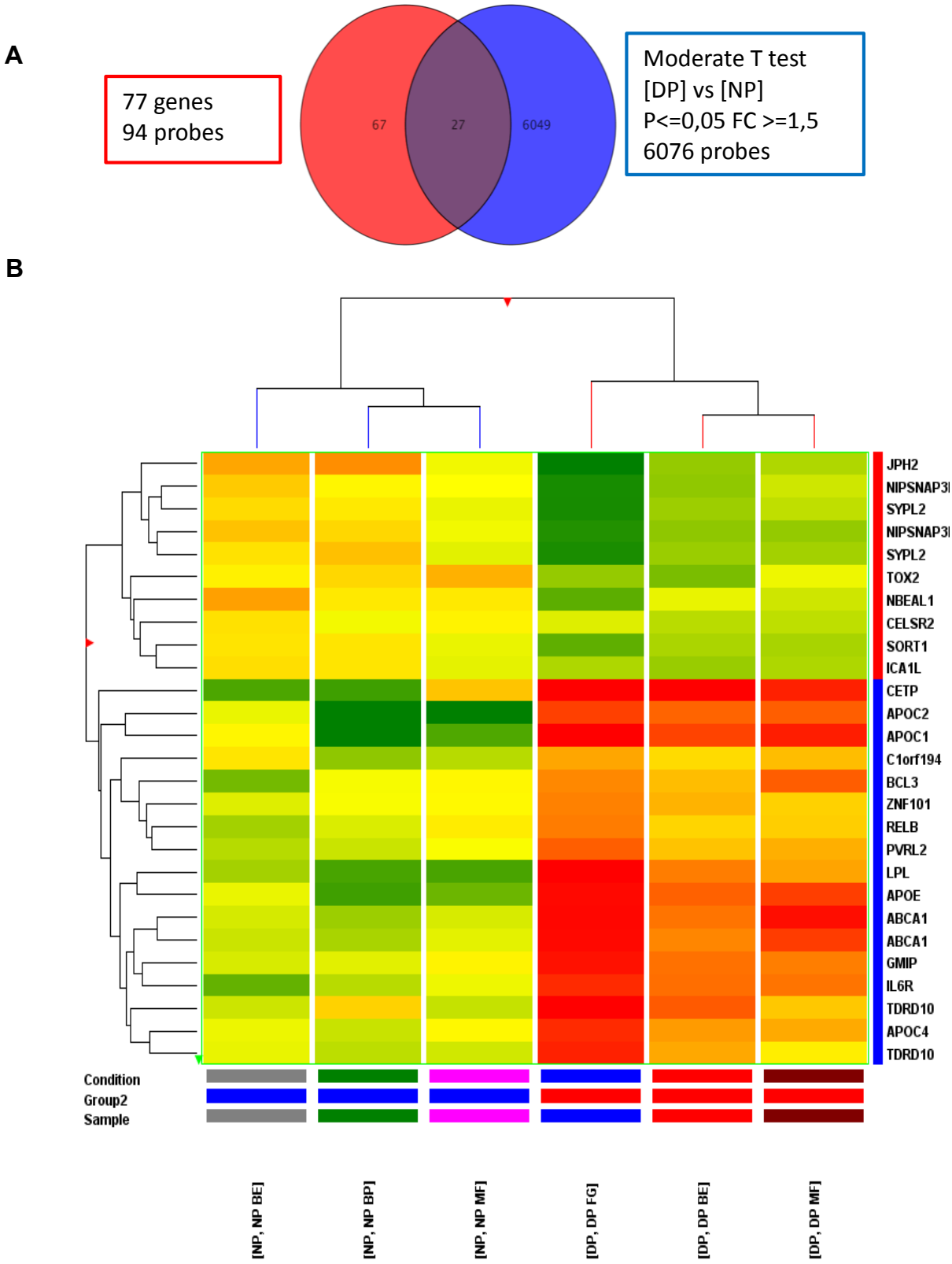


***FIGURE r-11*. A) Diagram showing the relationship between 77 neighbouring genes (94 probes, red circle) and total genes (6076 probes, blue circle) significantly modulated in carotid specimens. B) Heat map representation of the 23 differentially expressed genes (27 probes), intersection in A.** Each column represents one sample (DP and NP) and each row represents one gene (probe), as indicated. Colours represent the expression level fold change: higher-red, lower- green and no difference-yellow.

ATP-binding membrane cassette transporter A1 (*ABCA1*), is a regulator of lipid metabolism whose role in atherosclerosis pathways is extensively studied and well recognised. For this reason, the attention was then addressed to the BCL3 and PVRL2 genes, contiguously located on the chromosome region (19q13) marked by the intergenic rs10402271 (***FIGURE r-2***). This intergenic SNP was included in the list of 15 SNPs nominally associated with CAD-MI.

With the aim to validate findings from the microarray profiling of specimens, qPCR analyses were performed on carotid tissue samples, not previously analysed by microarray. Validation analyses were conducted in 19 tissue samples, both DP:NP couples (5 patients) and independent DP and NP tissue (14 patients). We used two housekeeping genes *ACTB* (beta actin) and *B2M* (beta-2-microglobulin) as internal controls for normalisation (***FIGURE r-12***).

BCL3 mRNA levels were significantly higher in DP than in NP tissue by comparison of couples (mean fold change 10.07 ± 2.87, P=0.034) as well as non-coupled specimens (mean fold change 6.81 ± 3.16, P=0.013) with the *ACTB* as reference gene. Differences were also confirmed with *B2M*, both in couples (mean fold change 3.59 ± 0.91, P=0.11) and in non-coupled specimens (mean fold change 6.13 ± 1.58, P=0.005).

PVRL2 was significantly up regulated in DP, by comparison of DP:NP couples (mean fold change 2.60 ± 0.46, P=0.026) and, as a trend, in non-coupled DP and NP (mean fold change 2.52 ± 0.80; P=0.060) with *ACTB* as reference gene. Coherent data were obtained with normalized *B2M* gene, in couples (mean fold change 1.34 ± 0.32, P=0.35) and in non-paired tissues (mean fold change 1.39 ± 0.36, P=0.38).

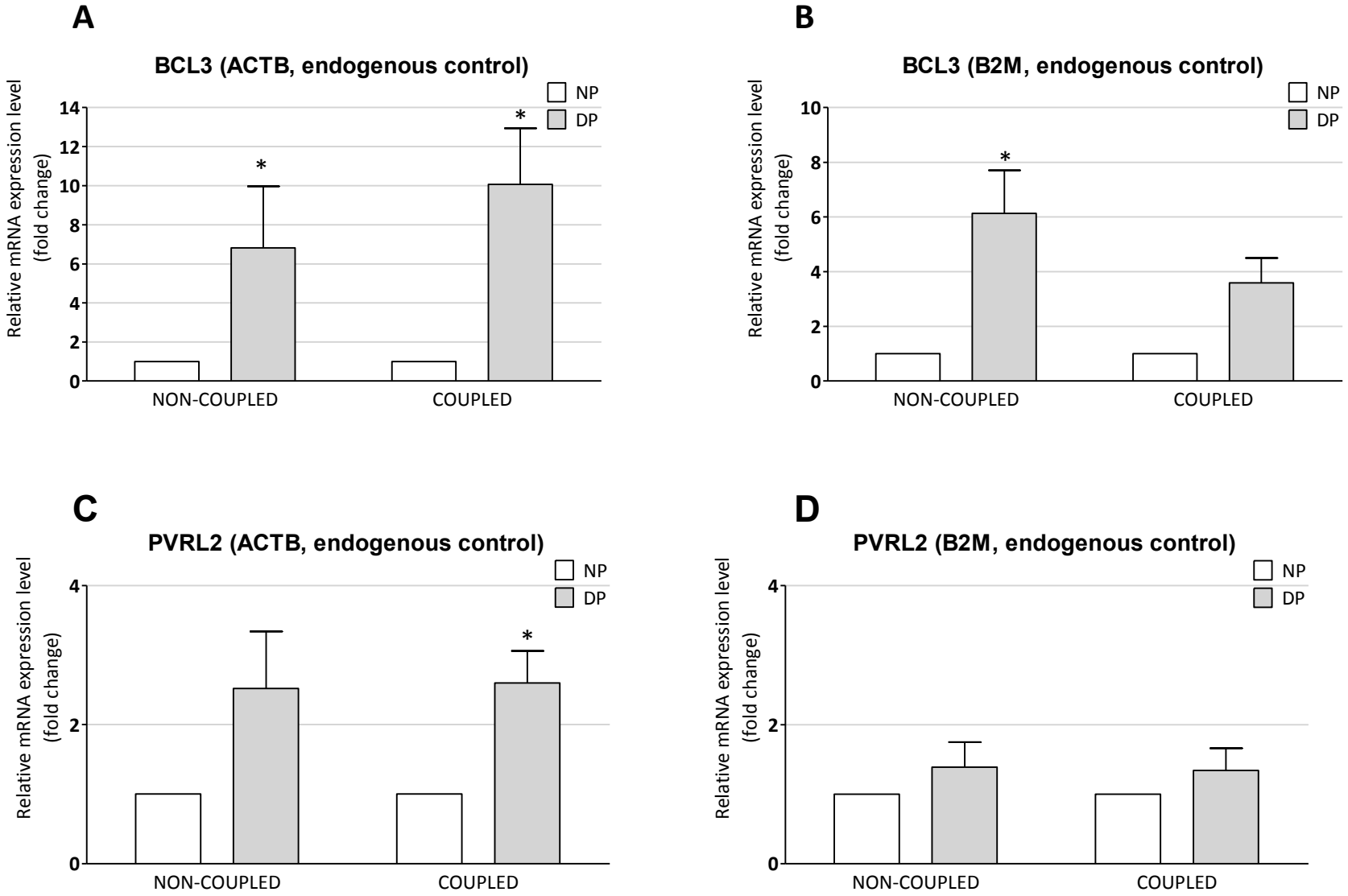


***FIGURE r-12*. Tissue mRNA expression levels of BCL3 and PVRL2 evaluated both in DP:NP couples (patients) and in non-coupled DP and NP tissues (14 patients; NP =7, DP =7) normalized with ACTB and B2M.** mRNA expression levels in NP are set to one as control.
Data are reported as mean fold change ±SEM.
* P < 0.05 by unpaired t-test

4.3 SECOND GENETIC ANALYSIS

4.3.1 *BCL3* SNPs genotyping for association with CAD

Prompted by transcriptomic data suggesting *BCL3* and *PVRL2* as potential candidates for CAD, we moved towards their genotyping in a second population study. This second population study was represented by CAD-free controls (n=393) and CAD patients without previous MI (n=442). This cohort enabled us to preferentially address the atherosclerotic pathway rather than the thrombotic process, and to avoid the potential biases due to selection/survival related to MI history. The clinical and laboratory characteristics of the two groups are summarized in ***TABLE m-1***.

As the originally investigated rs10402271 maps in the intergenic 19q13 region between *BCL3* and *PVRL2*, a search for intragenic SNPs for *BCL3* and *PVRL2* was conducted in the NCBI database. All the exonic SNPs reported for *BCL3* and *PVRL2* had a MAF (Minor Allele Frequency) <0.01 and <0.02 (by NCBI) respectively and thus not eligible for this study.

For this reason, for *BCL3* we selected two SNPs, rs2965169 (MAF=0.49) and rs8100239 (MAF 0.39), which were located in the 5' near gene and in the first intron respectively (***FIGURE r-13 A***). Similarly, for *PVRL2*, the rs3810143 (MAF=0.37, 5' near gene) and the rs1871047 (MAF=0.27, first intron) were selected (***FIGURE r-13 B***). Although there is no reported evidence that *BCL3* SNPs (rs2965169 and rs8100239) and *PVRL2* SNPs (rs3810143 and rs1871047) are functional polymorphisms, the inspection of the data of Encyclopaedia of DNA Elements (ENCODE), for functional elements by H3K27 histone acetylation mark and by chromatin state segmentations analysis (UCSC Genome Browser hg19), provided us with information about potential functionality. Data for cell lines involved in atherosclerosis pathways, other than VSMCs for which data were not available, indicated that both *BCL3* rs2965169 and rs8100239 polymorphisms lie within a potentially active regulatory elements in endothelial-(HUVEC), B-lymphocyte- (GM12878) and fibroblast-(NHLF) cell lines. Also *PVRL2* gene SNPs (rs3810143 and rs1871047) lie within putative functional elements in endothelial cell (HUVEC) and fibroblast(NHLF) (***FIGURE r-13***).

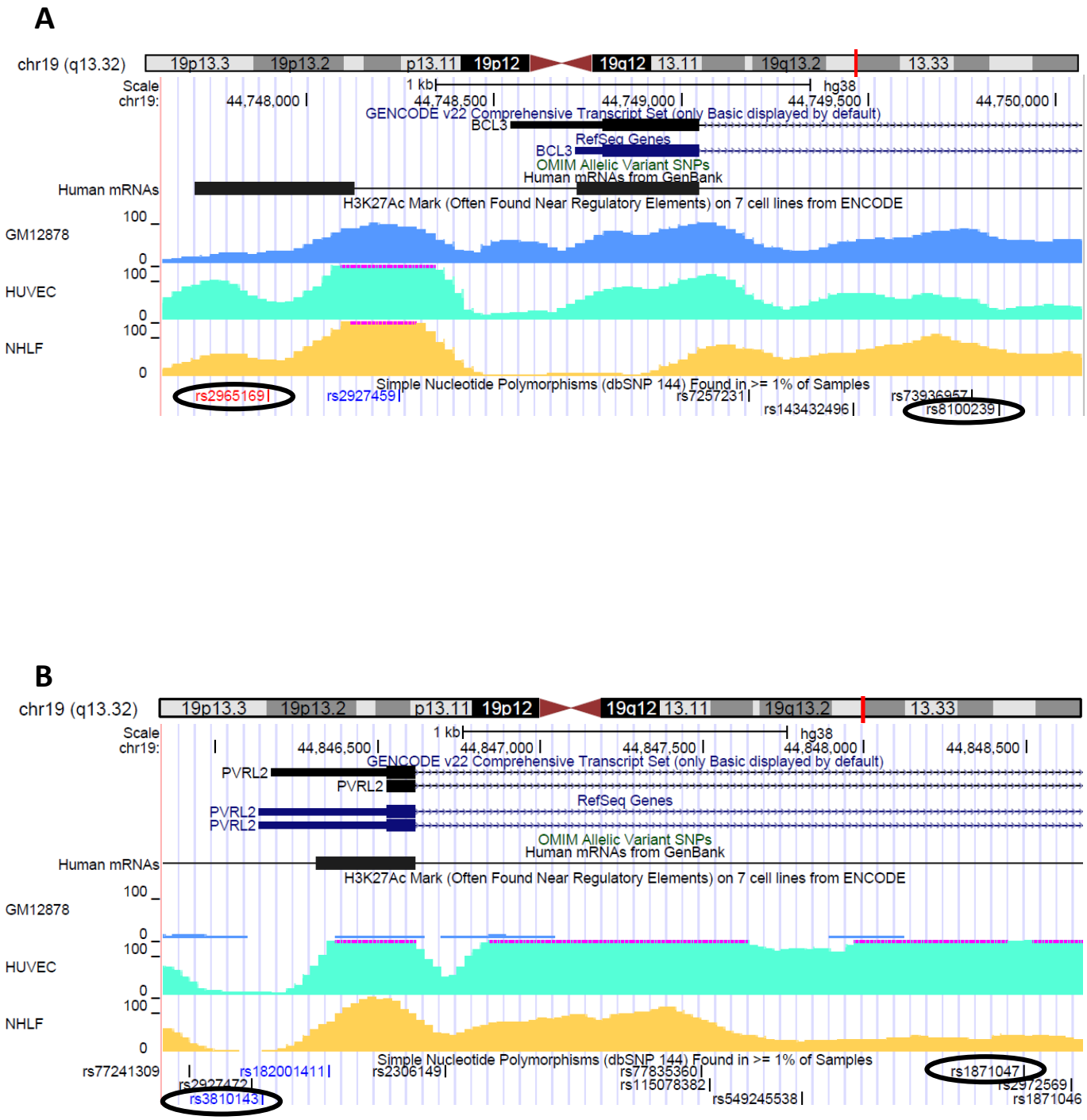


FIGURE r-13. **Position of BCL3 and PVRL2 SNPs in potentially active regulatory elements from the UCSC Genome Browser, hg19, ENCODE data**. A) *BCL3* rs2965169 and rs8100239 SNPs. B) *PVRL2* rs3810143 and rs1871047.

Genotyping for the 5 SNPs in CAD and CAD-free patients is reported in ***TABLE r-6***. In both cases and controls genotypes had a distribution consistent with Hardy-Weinberg equilibrium by Cubex analysis program online available. By using the same analysis program, the linkage disequilibrium between the investigated SNPs was calculated. A moderate degree of linkage disequilibrium was present between the two *BCL3* SNPs as well as between the two *PVRL2* SNPs.

The intergenic rs10402271 showed a low degree of linkage disequilibrium with *BCL3* and a moderate degree with *PVRL2* polymorphisms. The linkage disequilibrium data were in accordance with those reported in Hap Map (***FIGURE r-14***).

The *BCL3* rs2965169 genotype distribution differed significantly between cases and controls (***TABLE r-6***). The carriership of the minor allele G was more frequent in CAD patients (***TABLE r-6***). The estimation of the strength of the association between the carriership of *BCL3* rs2965169 minor allele G and CAD was calculated by means of ORs. The risk associated to the carriership (OR=1.49) remained also after adjustment (OR=1.70) for all the traditional cardiovascular risk factors (sex, age, body max index [BMI], smoking, hypertension, diabetes, LDL- and HDL-cholesterol, triglycerides, creatinine, high sensitive C Reactive Protein [hs-CRP]) independently associated with CAD in different- adjusted logistic regression models (***FIGURE r-15***).

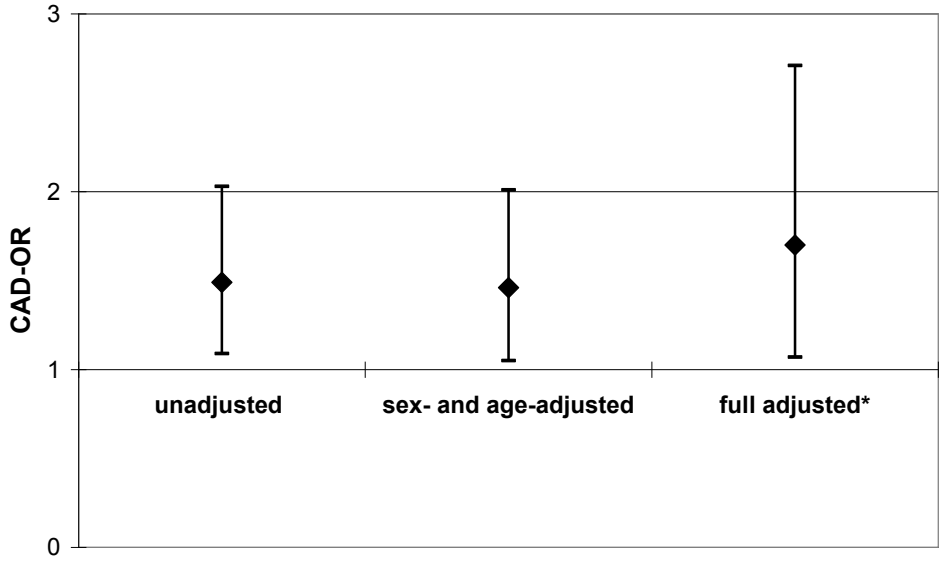


FIGURE r-15. **Association between the carriership of BCL3 rs2965169 minor allele G and CAD**. The strength of the association was estimated by means of ORs with 95% CI at univariate, sex- and age- adjusted, and full adjusted analysis. The homozygotes for the major allele (TT) were considered as reference. * by logistic regression analysis adjusted for sex, age, BMI, smoking, hypertension, diabetes, LDL- and HDL-cholesterol, triglycerides, creatinine and hs-CRP.

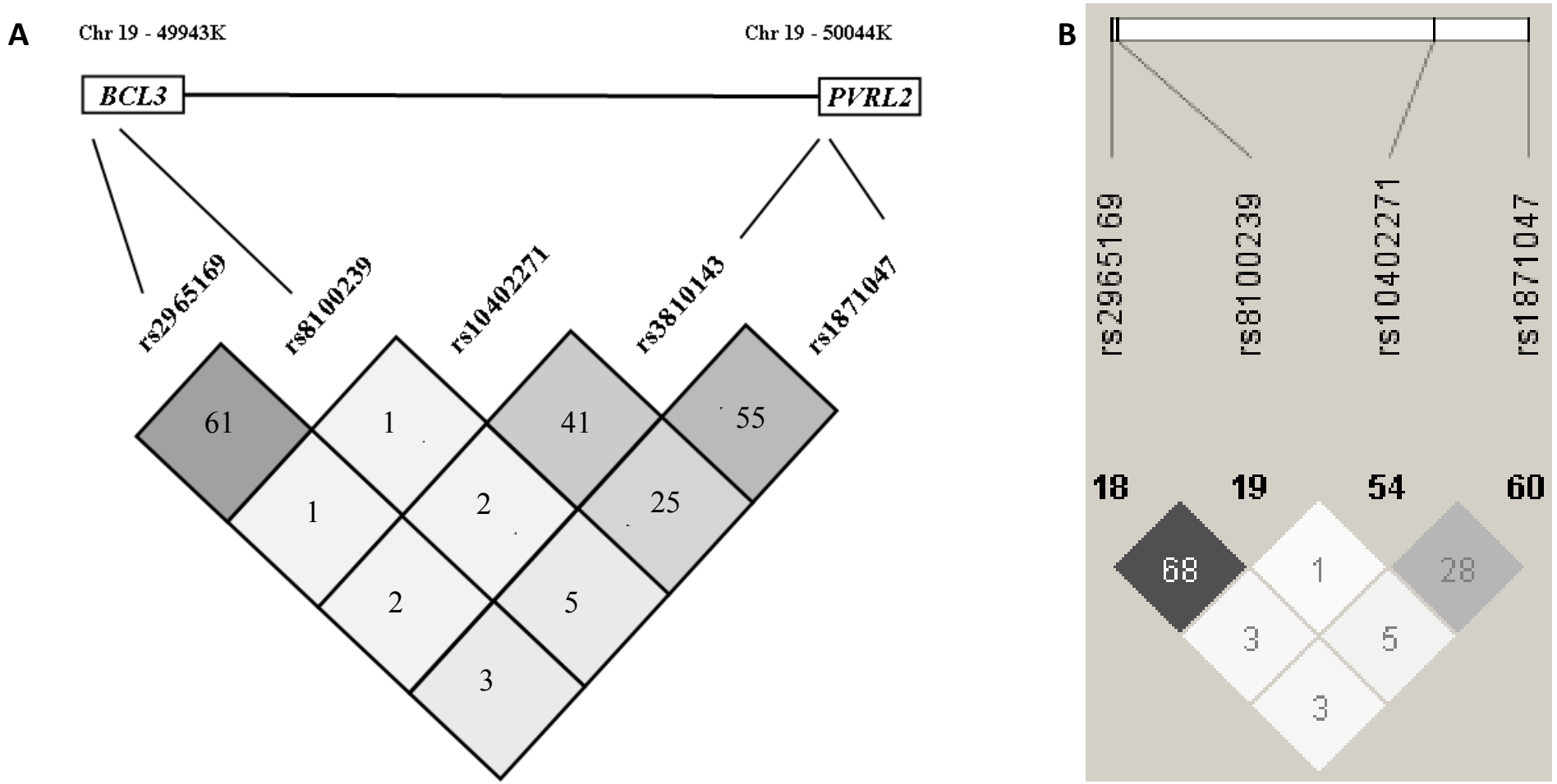


***FIGURE r-14*. Schematic representation of the linkage disequilibrium between SNPs in the genomic region chr19q13.** (A) linkage disequilibrium observed between the 5 investigated SNPs and (B) reported in Hap Map. The r^2 values of the pairwise linkage disequilibrium analysis are displayed. Values for SNPs rs2965169 (*BCL3*, 5'UTR), rs8100239 (*BCL3*, intron 1), rs10402271 (intergenic *BCL3-PVRL2*) and rs1871047 (*PVRL2*, intron 1) are in accordance with Hap Map data. The analysis of linkage disequilibrium for rs3810143 (*PVRL2*, 5'UTR) was not reported in Hap Map. Distances between genes are not in scale. LD plot was constructed by Haploview 4.2 software.

***TABLE r-6*. Distribution of *BCL3* and *PVRL2* genotypes in the study populations, with or without coronary artery disease (CAD)**

GENOTYPES	CAD-free n=393	CAD n=442	*P*
rs2965169 (*BCL3*)	**n (%)**	**n (%)**	
TT	115 (29.3)	96 (21.8)	
TG	194 (49.4)	235 (53.1)	0.020 *
GG	84 (21.4)	111 (25.2)	
rs8100239 (*BCL3*)	**n (%)**	**n (%)**	
TT	130 (33.1)	120 (27.1)	
TA	180 (45.8)	223 (50.5)	0.146 *
AA	83 (21.1)	99 (22.4)	
rs10402271 (intergenic *BCL3-PVRL2*)	**n (%)**	**n (%)**	
TT	197 (50.1)	206 (46.7)	
TG	162 (41.3)	182 (41.1)	0.168 *
GG	34 (8.6)	54 (12.2)	
rs3810143 (*PVRL2*)	**n (%)**	**n (%)**	
TT	128 (32.5)	140 (31.7)	
TC	181 (46.1)	218 (49.4)	0.753 *
CC	84 (21.4)	84 (18.9)	
rs1871047 (*PVRL2*)	**n (%)**	**n (%)**	
AA	144 (36.6)	150 (33.9)	
AG	178 (45.3)	204 (46.2)	0.372 *
GG	71 (18.1)	88 (19.9)	

* by chi-square test for linear trend.

The carriership of the *BCL3* rs8100239 minor allele A appeared to be more represented in CAD patients, but did not reach a statistically significant difference (72.9% versus 66.9%, P=0.062) (***TABLE r-6***). The genotype distribution of rs3810143, rs1871047 and rs10402271 did not significantly differ between CAD and CAD-free.

The analysis of genotypes was also conducted in the entire populations (CAD and CAD-free) in relation to intermediate metabolic phenotypes known to be linked with CAD. The

rs8100239 A allele showed significant associations with higher BMI and lower HDL-cholesterol concentration, as well as with a nominally no-significant trend for a higher prevalence of hypertension (***TABLE r-7 A***). The association with plasma lipids was also confirmed after the exclusion of subjects taking lipid-lowering drugs (data not shown). On the other hand, the BCL3 rs2965169 did not appear associated with lipid profile and metabolic phenotypes (***TABLE r-7 B***).

PVRL2 rs3810143 and rs1871047, and intergenic rs10402271 genotypes were not found associated with lipid profile and metabolic phenotypes.

TABLE *r*-7. Metabolic characteristics in the second study population according to the *BCL3* rs8100239 (A) and rs2965169 genotype (B).

A)

	***BCL3* rs8100239**			***P***
	TT (n=250)	**TA (n=403)**	**AA (n=182)**	
BMI, kg/m²	25.8± 3.4	26.0 ± 3.6	26.7±3.6	0.021 *
Hypertension, %	53.7	58.8	61.8	0.086 †
Diabetes, %	15.6	12.3	16.5	0.942†
Glucose,mmol/L	5.64 (5.53-5.81)	5.58 (5.47-5.70)	5.58 (5.42-5.75)	0.480 *
Total cholesterol, mmol/L	5.50± 1.10	5.41± 1.11	5.47± 1.19	0.718 *
LDL-cholesterol, mmol/L	3.52± 1.00	3.53± 0.92	3.67± 0.96	0.158 *
HDL-cholesterol, mmol/L	1.35± 0.44	1.32± 0.36	1.26± 0.35	0.040 *
Triglycerides, mmol/L	1.51 (1.43-1.60)	1.49 (1.43-1.57)	1.58 (1.48-1.69)	0.389 *

* by ANOVA with polynomial contrast for linear trend.
† by chi-square test for linear trend

B)

	***BCL3* rs2965169**			***P***
	TT (n=211)	**TG (n=428)**	**GG (n=195)**	
BMI, kg/m²	25.9 ± 3.5	26.0 ± 3.6	26.4±3.3	0.181 *
Hypertension, %	54.6	59.8	57.1	0.594†
Diabetes, %	13.4	14.4	14.3	0.790†
Glucose,mmol/L	5.58 (5.47-5.75)	5.64 (5.53-5.75)	5.53 (5.37-5.70)	0.465 *
Total cholesterol, mmol/L	5.50± 1.17	5.43± 1.13	5.42± 1.07	0.442 *
LDL-cholesterol, mmol/L	3.56± 1.05	3.54± 0.94	3.59± 0.86	0.775 *
HDL-cholesterol, mmol/L	1.32± 0.41	1.32± 0.39	1.30± 0.35	0.502 *
Triglycerides, mmol/L	1.52 (1.43-1.62)	1.51 (1.45-1.58)	1.51 (1.42-1.62)	0.898 *

* by ANOVA with polynomial contrast for linear trend.

† by chi-square test for linear trend

4.4 IMMUNOHISTOCHEMICAL ANALYSIS OF CAROTID SPECIMENS

Specimens of NP and of DP were processed for histology characterization by Miller and Masson's trichrome staining. All the specimens included the entire intima and the media comprising the layers up to the external elastic lamina. NP specimens exhibited a media and a thin thickened intima, consistent with American Heart Association (AHA) type III intermediate lesions. The DP exhibited a media underlying an atheromatous plaque, corresponding to AHA type IV-V atheroma-fibroatheroma (***FIGURE r-16).***

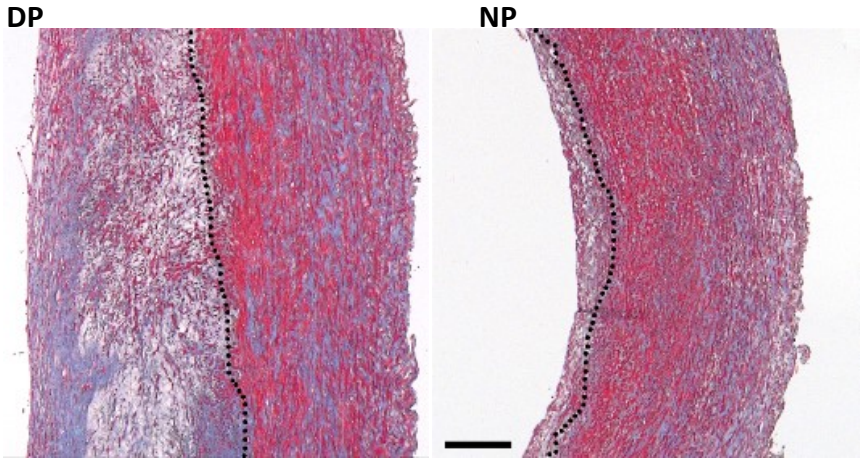


***FIGURE r-16*. Histological features of DP and NP carotid portions**. Dotted lines highlight the internal elastic lamina. The lumen is located on the left side of the pictures. Bar = 200 μm.

The expression of BCL3 protein was investigated by immunostaining of carotid specimens (n=15) with a mouse monoclonal antibody, previously tested on human tonsil (see Methods and ***FIGURE r-17***), a tissue known to overexpress BCL3.

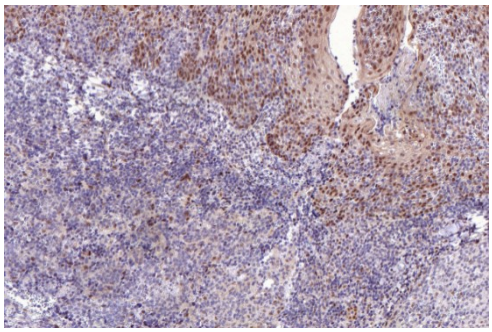


***FIGURE r-17*. Paraffin sections of normal human tonsil stained with anti-BCL3 antibody.**

Immunostaining for a-SMA and CD68 was also performed on adjacent sections.
BCL3 protein was found expressed only in the DP specimens (n=10). A positive staining for BCL3 was found in plaque foam cells (approximately 20-40%) and in rare intimal and medial VSMCs in 9 of 10 DP samples (***FIGURE r-18***).
We did not detect any BCL3-expressing cells in the intima-media of all the NP samples (n=5) (***FIGURE r-18***). BCL3-positive inflammatory cells which, however, were scarcely present in the analyzed atherosclerotic lesions, were not detected.

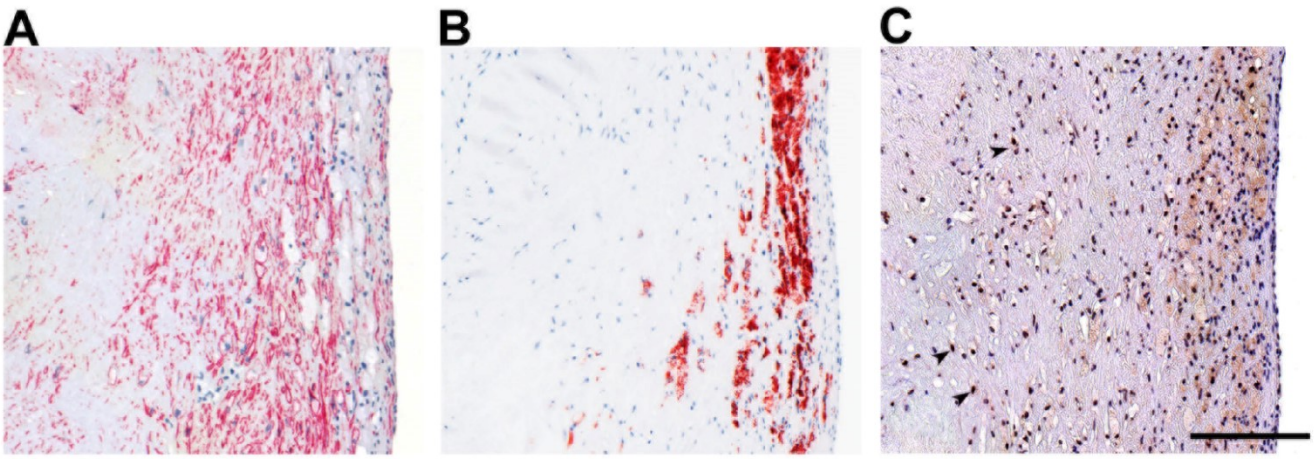


***FIGURE r-18*. BCL3 protein expression in atherosclerotic (DP) carotid specimens.** Representative immunohistochemical staining of the DP plaque for α-SMA (A) CD68 (B) and BCL3 (C). BCL3 is expressed by plaque foam cells and rare VSMCs (arrows). The lumen is located on the right side of the pictures. Bar= 200 μm.

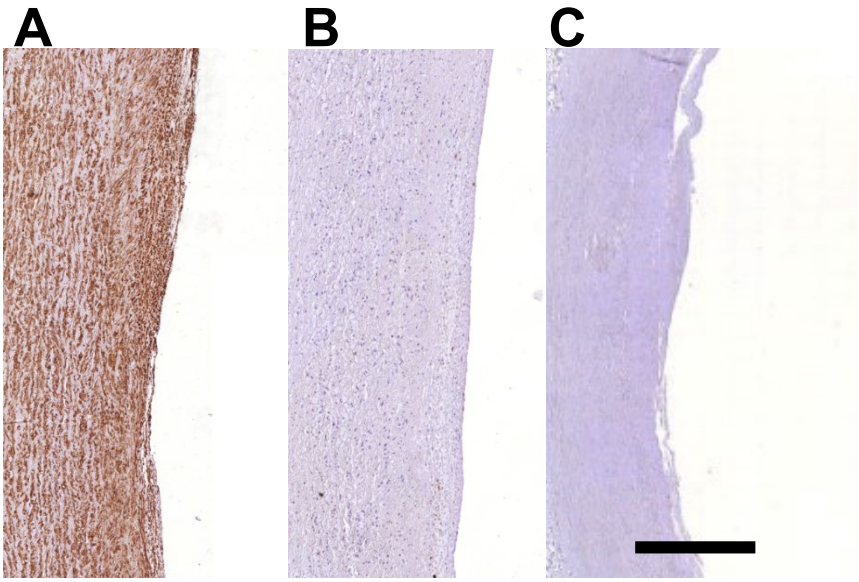


***FIGURE r-19*. BCL3 protein expression in grossly non-atherosclerotic (NP) carotid specimens.** Representative immunohistochemical staining of the NP for α-SMA (A) CD68 (B) and BCL3 (C).The lumen is located on the right side of the pictures. Bar= 200 μm.

TABLE r-1. **List of SNPs investigated in the Verona Heart Study for association with MI within stages of replication of GWAS MIGen Consortium.**

SNP	Chr	SNP type	Gene or nearest gene(s)	*P* *	OR†
rs646776	1p13	intergenic	*CELSR2, PSRC1, SORT1*	0.001	0.62 (0.47-0.80)
rs1122608	19p13	intronic	*SMARCA 4, LDLR*	0.005	0.70 (0.56-0.87)
rs3786722	19p13	intronic	*SMARCA 4, LDLR*	0.005	0.70 (0.56-0.88)
rs3764261	16q13	intergenic	*HERPUD1,CETP*	0.014	0.72 (0.58-0.90)
rs1333049	9p21	intergenic	*CDKN2B*	0.016	0.86 (0.70-1.00)
rs8192284	1q21	coding	*IL6R*	0.018	0.80 (0.65-0.99)
rs4977574	9p21	intergenic	*CDKN2B, CDKN2A*	0.019	0.83 (0.68-1.00)
rs4607103	9p14	intergenic	*ADAMTS9*	0.023	1.38 (1.10-1.74)
rs17482753	8p22	intergenic	*LPL*	0.041	0.69 (0.50-0.94)
rs16996148	19p13	intergenic	*CILP2, PBX4*	0.042	0.62 (0.42-0.91)
rs3890182	9q31	intronic	*ABCA1*	0.044	1.45 (1.08-1.95)
rs328	8p21	coding	*LPL*	0.065	0.71 (0.52-0.98)
rs10402271	19q13	intergenic	*BCAM, PVRL2*	0.068	1.27 (1.02-1.57)
rs16988929	20q12-13	intronic	*GDAP1L1, HNF4A*	0.078	0.55 (0.30-0.99)
rs6725887	2q33	intronic	*WDR12*	0.084	1.39 (1.03-1.87)
rs4430796	17cen-q21	intronic	*TCF2*	0.107	0.84 (0.68-1.03)
rs4675310	2q33	intronic	*NBEAL1*	0.120	1.36 (1.01-1.82)
rs11206510	1p32	intergenic	*BSND,PCSK9*	0.127	0.78 (0.61-0.99)
rs10415849	19p13	intronic	*GATAD2A*	0.142	0.71 (0.50-1.01)

SNP	Chr	SNP type	Gene or nearest gene(s)	*P* *	OR†
rs1867000	3p26	intergenic	*ITPR1*	0.172	1.21 (0.99-1.48)
rs7947046	11q12	intergenic	*C11orf66,SYT7*	0.172	0.69 (0.46-1.02)
rs20455	6p21	coding	*KIF6*	0.180	1.17 (0.95-1.44)
rs6429535	1p32	intronic	*PRNPIP*	0.208	0.88 (0.70-1.10)
rs9989419	16q21	intergenic	*CETP,HERPUD1*	0.218	1.19 (0.96-1.47)
rs780094	2p23	intronic	*BROX*	0.219	1.16 (0.95-1.42)
rs17782313	18q21	intergenic	*MC4R*	0.220	1.23 (0.97-1.57)
rs501120	10q11	intergenic	*CXCL12*	0.221	0.98 (0.74-1.30)
rs1746048	10q11	intergenic	*CXCL12*	0.225	0.95 (0.71-1.26)
rs4240934	1q41	intronic	*BROX*	0.251	1.23 (0.92-1.65)
rs2943634	2q36	intergenic	gene desert	0.258	0.85 (0.69-1.06)
rs3846663	5q14	intronic	*HMGCR*	0.273	1.19 (0.96-1.46)
rs2808630	1q23	intergenic	*CRPP1,CRP*	0.292	0.92 (0.74-1.16)
rs289742	16q21	intergenic	*CETP*	0.295	0.80 (0.60-1.08)
rs2338104	12q24	intronic	*KCTD10*	0.311	0.86 (0.70-1.06)
rs6922269	6q25	intronic	*MTHFD1L*	0.311	0.84 (0.67-1.05)
rs6458545	6p24	intronic	*PHACTR1*	0.312	0.87 (0.71-1.07)
rs10483778	14q24	intronic	*FUT8*	0.326	1.54 (0.66-3.63)
rs16969968	15q24	coding	*CHRNA5*	0.343	1.15 (0.94-1.42)
rs11774572	8p22	intergenic	*BLK*	0.352	1.02 (0.83-1.24)
rs10889353	1p31	intronic	*DOCK7*	0.366	1.17 (0.94-1.47)
rs17228212	15q22	intronic	*SMAD3*	0.383	0.85 (0.68-1.07)
rs6544366	22p23-24	intergenic	*APOB*	0.384	0.86 (0.69-1.08)

SNP	Chr	SNP type	Gene or nearest gene(s)	*P* *	OR†
rs6511720	19p13	intronic	*LDLR*	0.389	0.83 (0.61-1.12)
rs11591147	1p32	coding	*PCSK9*	0.391	1.57 (0.56-4.37)
rs12526453	6p24	intronic	*PHACTR1*	0.443	0.88 (0.71-1.08)
rs4835412	4q31	intronic	*EDNRA*	0.447	1.22 (0.89-1.66)
rs6589566	11q23	intronic	*ZNF259*	0.451	1.24 (0.87-1.75)
rs4783962	16q21	intergenic	*CETP*	0.454	1.04 (0.83-1.32)
rs10010131	4p16	intronic	*WFS1*	0.467	0.91 (0.74-1.11)
rs1051730	15q24	coding	*CHRNA5*	0.492	1.12 (0.91-1.38)
rs10923931	1p13-11	intronic	*NOTCH2*	0.504	0.99 (0.70-1.42)
rs10954284	7q32	intergenic	*KLF14*	0.512	0.89 (0.73-1.09)
rs12779790	10p13	intergenic	*CAMK1D,CDC123*	0.517	0.97 (0.75-1.25)
rs12802349	11q12	intergenic	*C11orf66, SYT7*	0.538	0.76 (0.45-1.30)
rs4242231	5q13	intergenic	*GPR98, FLJ42709*	0.548	1.10 (0.85-1.44)
rs4402960	3q27	intronic	*IGF2BP2*	0.553	1.07 (0.86-1.33)
rs7953249	12q24	intergenic	*TCF1*	0.554	0.89 (0.73-1.10)
rs864745	7p15	intronic	*JAZF1*	0.561	1.02 (0.84-1.25)
rs3798156	6q25	intronic	*SLC22A2*	0.575	1.16 (0.84-1.59)
rs562338	2p23-24	intergenic	*APOB*	0.586	0.91 (0.70-1.18)
rs1121980	16q12	intron	*FTO*	0.608	1.08 (0.88-1.32)
rs1367211	6q26	intron	*LPA*	0.628	1.09 (0.88-1.35)
rs4553696	18q12	intergenic	*DSC3, CDH2*	0.655	0.91 (0.73-1.12)
rs174547	11q12-13	intronic	*FADS1*	0.666	0.91 (0.73-1.13)
rs1421085	16q12	intronic	*FTO*	0.672	1.09 (0.89-1.33)

SNP	Chr	SNP type	Gene or nearest gene(s)	*P* *	OR†
rs17145738	7q11	intergenic	*TBL2*	0.676	0.96 (0.66-1.40)
rs7578597	2p21	coding	*THADA*	0.681	0.92 (0.67-1.27)
rs17465637	1q41	intronic	*MIA3*	0.687	1.10 (0.88-1.39)
rs3127575	6q25	intronic	*SLC22A2*	0.696	1.14 (0.82-1.60)
rs10483782	14q24	intronic	*FUT8*	0.707	1.45 (0.71-2.97)
rs9939609	16q12	intronic	*FTO*	0.709	1.05 (0.85-1.29)
rs7903146	10q25	intronic	*TCF7L2*	0.711	0.95 (0.76-1.18)
rs4703910	5q13	intronic	*ZNF366*	0.716	0.92 (0.71-1.18)
rs4939883	18q21	intergenic	*LIPG, ACAA2*	0.733	0.99 (0.75-1.31)
rs2271293	16q22	intronic	*NUFT2*	0.733	1.11 (0.82-1.51)
rs4398167	18q12	intergenic	*DSC3, CDH2*	0.766	0.93 (0.75-1.14)
rs693	2p24	coding	*APOB*	0.782	1.04 (0.86-1.28)
rs1501908	5q33	intergenic	*TIMD4, HAVCR1*	0.792	1.03 (0.84-1.25)
rs7583085	2p21	intergenic	*HAAO*	0.793	0.92 (0.72-1.18)
rs116843064	19p13	coding	*ANGPTL4*	0.801	0 (0-∞)
rs13266634	8q24	coding	*SLC30A8*	0.810	1.01 (0.80-1.27)
rs17321515	8q24	intergenic	*TRIB1*	0.827	0.96 (0.78-1.17)
rs6995374	8p23	intergenic	*MSRA, PRSS55*	0.830	1.08 (0.85-1.36)
rs9878602	3p14	intronic	*FOXP1*	0.836	0.96 (0.79-1.18)
rs867040	10p12	intronic	*KIAA1217*	0.842	0.96 (0.77-1.19)
rs973754	21q22	intergenic	*KCNE2*	0.861	1.07 (0.81-1.43)
rs4846914	1q41	intronic	*GALNT2*	0.863	1.06 (0.86-1.30)
rs9982601	21q22	intergenic	*KCNE2*	0.950	1.01 (0.75-1.36)

SNP	**Chr**	**SNP type**	**Gene or nearest gene(s)**	***P*** *	**OR†**
rs1800588	15q21	intergenic	*LIPC*	0.963	1.02 (0.79-1.30)
rs10014689	4q11	intergenic	*KDR*	0.967	0.81 (0.43-1.52)
rs769449	19q13	intronic	*APOE*	0.978	1.03 (0.72-1.48)

All the 91 SNPs are reported on the basis of their statistical significance, from the lowest to the highest P value. The 15 selected SNPs are reported in bold type.

* Sex- and age-adjusted P value for genotype distribution. † Sex- and age-adjusted Odds Ratios (ORs) for minor allele by additive model.

Chapter V:

Discussion and Concluding Remarks

DISCUSSION

Cardiovascular disease (CVD) is still one of the major causes of mortality and morbidity in the worldwide. Atherosclerosis is the common ground of several clinical manifestations of CVD, including coronary artery disease (CAD), myocardial infarction (MI), peripheral artery occlusive diseases, and stroke. Atherosclerosis is a multifactorial disease involving the interplay of genetic and environmental factors and the contribution of several cellular and plasmatic components of blood and the vessel wall. The research of my thesis was aimed to identify, by integrated approaches, genetic and molecular signatures in atherosclerosis. The experimental design involved investigation at DNA, RNA and protein level.

The study at DNA level includes the analysis of data obtained within a replication stage of GWAS –MIGen– and the genotyping of CAD patients and controls. At the RNA level, the transcriptome analysis involved microarray and qPCR in cultured VSMCs and carotid specimens. Immunohistochemical analysis was used for the investigation at protein level.

This experimental design was prompted by evidence from most recent studies showing that combined approaches, exploiting complementary methodologies, could lead to significant advances in understanding the mechanisms of multifactorial diseases, like CVD [56,72].

GWAS are powerful tools for unveiling the genetic components of CVD. Indeed, a number of GWAS have uncovered multiple genetic loci that are associated with

CAD and MI [reviewed in 73; 42]. However, the stringent level of statistical significance required in GWAS (P value $<5x10^{8}$) [44;74], necessary to exclude false positive results, may lead to discard genetic variants that could contribute to disease risk [44;75]. Recent observations suggest that GWAS databases likely include a number of "hidden" effective variants that up to now have only suggestive statistical evidence of association [76,77].

Moreover, the considerable overlap among clinical phenotypes of CVD leads several GWAS to cumulate indifferently CAD and MI in their analysis [53]. Recent studies have consistently demonstrated that some SNPs are associated more with CAD than with a specific predisposition to MI [78]. As matter of fact, in this thesis the second population study was represented by CAD-free controls and CAD patients without previous MI. Considering that the VSMC model refers principally to atherosclerosis pathways this second population study enabled to preferentially investigate this pathologic process rather than its acute thrombotic complication, like MI. Moreover, this second population avoided the potential biases due to survival after MI. The clear-cut definition of the clinical phenotype, with the exclusion from the control group of the subjects with subclinical but significant CAD, was obtained by an objective angiography-evaluation of the coronary artery vessels [95] in all enrolled subjects [62,63] .

The analysis of GWAS data (replication stage of GWAS –MIGen–) was conducted with low stringency criteria (P-value arbitrarily set at <0.10) in order to not exclude variant signals potentially associated with cardiovascular disease. This analysis provided us 71 genes proximal to 15 GWAS tag-SNPs. The functional association of these potential candidate genes (71) with atherosclerosis was inspected by two independent transcriptomic analyses in both VSMCs cultured from atherosclerotic and non-atherosclerotic tissue samples, and in entire carotid wall samples.

The VSMCs, which constitute the predominant cellular element of the vascular media are one of the accepted cellular models for detecting gene expression signatures of atherosclerosis. Mechanisms influencing the pathological behavior of VSMCs in atherosclerosis are object of extensive research and their contribution in plaque development and progression has been recently reviewed [25,26]. As the culture condition by itself may cause artificial shifts in transcriptome, not reflecting the *in vivo* changes of gene-expression, a microarray-based expression profiling of

the entire atherosclerotic and non-atherosclerotic carotid portions was performed to overcome this disadvantage.

Overall the transcriptome analyses, performed both at cellular and tissue level, compared atherosclerotic lesions (DP) with grossly non-atherosclerotic (NP) portions. The considered NP showed some histological alterations, like a thin thickened intima. Nonetheless, this grossly non-atherosclerotic portion appeared as a suitable endogenous control. Moreover, the study of couples of specimen would reduce variability in the gene expression profile.

A limitation in the transcriptional analysis could be represented by the small number of VSMC populations or whole specimens analyzed. However, human carotid specimens are quite unique samples, the NP portions are by definition very small in size and the RNA extraction from specimens is difficult due to the presence of abundant RNAses particularly in the atherosclerotic plaques [79,80]. The limitation due to small sample size in microarray has been tackled by the combination of subsequent and independent transcriptome analyses.

In microarray analysis, we did not perform adjustment for multiple comparisons. We attempted to overcome this drawback by two independent microarray-based transcriptome analyses, the former on cultured VSMCs, and the latter on atherosclerotic and non- atherosclerotic portions of carotid artery, setting a 10% False Discovery Rate (FDR) threshold in statistical analysis. It is worthy to note that, the non-adjustment for multiple comparisons is accepted in the field of studies integrating DNA, RNA, and protein evidences, with the aim to discover new molecular mechanisms in the pathophysiology of different diseases [81

The GWAS-related tag SNPs and two transcriptome analyses, in both cultured cells and tissue, enabled us to propose three candidate genes, *BCL3*, *PVRL2* and *ABCA1.* The last one has been extensively investigated in relation to the pathogenetic mechanisms of atherosclerotic vascular diseases [82,83].

The recognized role of the transporter *ABCA1* in CVD, strengthened by the observation that *ABCA1* mutations cause Tangier disease and premature atherosclerosis [84], supports the experimental approach of this thesis, aimed at detecting genetic components of atherosclerotic vascular diseases. The two other candidate genes (*BCL3* and *PVRL2*) were contiguously located on chromosome 19, in a region found associated with blood lipid traits, through a GWAS exploiting a magnetic bead-based technology [85]. In addition, this chromosome region,

marked by the intergenic rs10402271 nominally associate with CAD-MI, appeared of interest as suggested by the up regulation of *PVRL2,* observed in mouse aortic endothelium in response to atherogenic stimuli [86] .

All these reasons prompted us to further investigate specific SNPs for *BCL3* and the adjacent *PVRL2*, in a second case-control analysis considering CAD patients without MI history. This step produced coherent results: the carriership of *BCL3* rs2965169 G allele was associated to an increased risk of CAD independently from traditional atherosclerosis risk factors. The *BCL3* rs8100239 A allele, in moderate linkage, tended also to be more represented among CAD patients.

No significant associations were found for either *PVRL2* SNPs or *BCL3-PVRL2* intergenic variant.

It is worth noting that evidence concerning the functionality of *BCL3* is not reported. The allele specific contribution to *BCL3* mRNA expression level, which could provide insight for a functional effect and for the associated risk, is difficult to determine because of the SNPs localisation (5'UTR and intron 1) and the virtual absence of in linkage coding SNPs. On the other hand, the inspection of the data of Encyclopaedia of DNA Elements (ENCODE), for functional elements by H3K27 histone acetylation mark and by chromatin state segmentation analysis, both BCL3 rs2965169 and rs8100239 polymorphisms have been shown to lie within potentially active regulatory elements in endothelial-(HUVEC), B-lymphocyte-(GM12878) and fibroblast (NHLF) cell lines (see ***FIGURE r-13***). Interestingly, although no data on VSMCs were available, these observations have been obtained in several cell lines involved in atherosclerosis pathways.

Although until now classical GWAS approaches have failed to identify BCL3 as risk gene for CAD [73], a recent study combining GWAS results with expression quantitative trait loci (eQTLs) analysis identified BCL3 as a potential new risk gene for Crohn's disease [87]. Interestingly, this chronic inflammatory disorder is associated with an increased atherosclerotic risk unexplained by traditional cardiovascular risk factors [88]. Additional concern stems from a recent bioinformatics analysis of microarray data (GSE480609), reporting BCL3 up expression in blood samples from first-time acute MI patients (within 48-hours) as compared with normal controls [89].

With regard to the results at RNA level, our findings are also consistent with data from a transcriptome analysis by microarray on human carotid plaques (GEO

database, GSE 43292). In this study, which exploited an experimental design similar to ours, *BCL3* mRNA levels in the majority of atheroma plaques as compared to the macroscopically intact tissues [90].
Again with regard to the transcriptome studies, data obtained for BCL3 expression were non coherent between cultured VSMCs and in tissues. This could be explained by the complex transcriptional and post-transcriptional regulation of BCL3 that could be perturbed by in vitro culture state and thus differ from the in vivo condition [37]. In addition, the expression of genes implicated in the regulation of apoptosis, such as *BCL3*, could be perturbed by in vitro culture in itself.

The third level of this integrated study has been represented by the investigation of BCL3 protein expression in atherosclerotic vascular vessel. This analysis has not reported so far. BCL3 positivity was found only in atherosclerotic portions and in cells, like foam cells and VSMCs, which are well known to play crucial role in atherogenesis. Actually, very few positive VSMCs were observed, but this result is not surprising taking into account the lack of BCL3 positivity in most of tissues as reported in the Human Protein Atlas database. The BCL3 positivity in the DP portions supports the hypothesis of a role of BCL3 in CVD and fits with RNA data in carotid specimens. However, the very low BCL3 protein expression levels detected in DP portions might not favour to disentangle the interpretation of BCL3 role in atherosclerosis.

BCL3 is a transcriptional coregulatory and member of the inhibitor of nuclear factor-kB (IkB) family, which takes part to both positive and negative modulation of genes belonging to several pathways, including those of cell death/proliferation, inflammatory and stress responses [29;37;91].Biological information from Gene Ontology database locates BCL3 in multiple processes, such as immune response, apoptosis, biosynthetic process and TNF-a signalling which are known to be relevant in atherosclerosis (fig. xxxx).
Interestingly, it has been proposed that, through interaction with PGC-lα, ERRα and PPARα44, BCL3 takes part to the activation of several genes including those of energy metabolism [92], which further suggests a biological plausibility of its role in metabolic disorders and cardiovascular diseases. According with this hypothesis, in the present study the BCL3 rs8100239 A allele correlated with a less favourable lipid profile, in particular low HDL levels, and overweight. A trend of association was

noted also for hypertension. Such traits usually characterize the so-called metabolic syndrome, a still debated cluster of risk factors including altered energy metabolism [93,94].

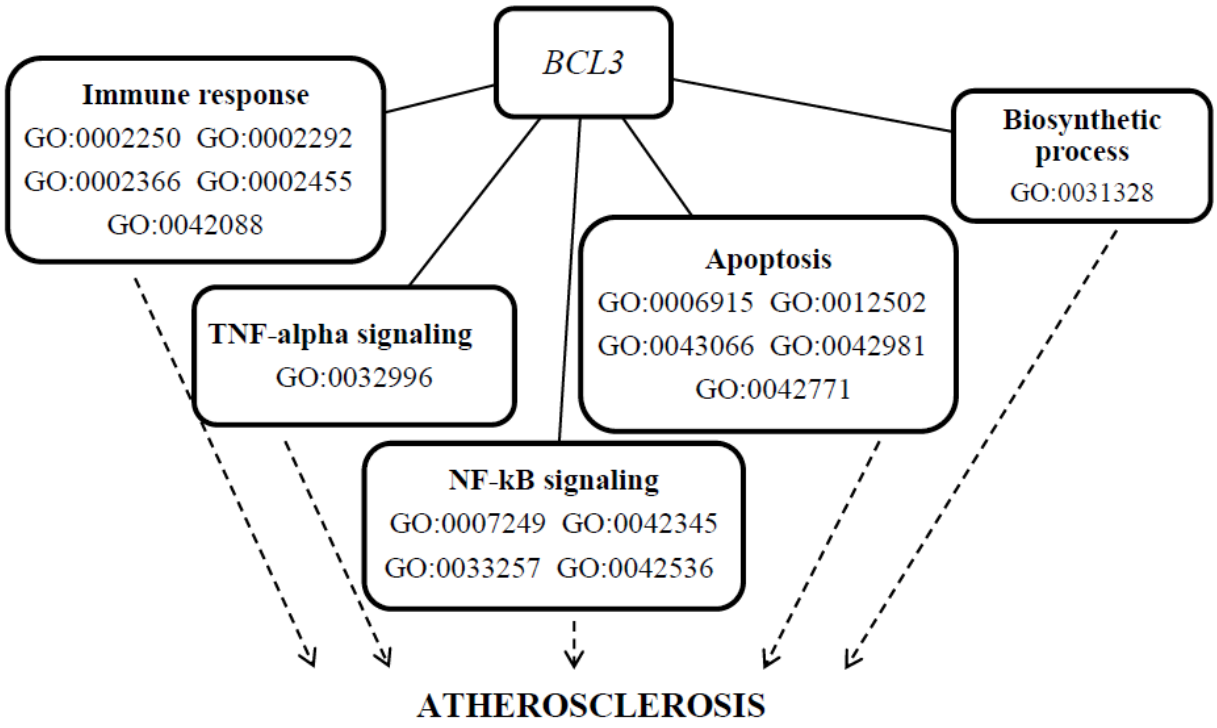


***FIGURE d-1*: Potential role of BCL3 in atherogenesis**. Functional annotations (GO terms) locate BCL3 in biological processes (boldface) known to be involved in atherosclerosis

CONCLUDING REMARK

This thesis shows that the integration of GWAS data with the downstream changes in the RNA and protein levels in human atherosclerotic and grossly non-atherosclerotic arterial wall portions, might be a powerful approach to unravel new genetic and molecular components involved in CVD.

The overlapping of consistent data from both genomic and transcriptomic analysis supports the hypothesis of a role for BCL3 in the pathways of atherosclerosis and coronary artery disease, which could be partly mediated through the influence on metabolic phenotypes.

Further studies in a larger number of samples, both at RNA and protein levels in cells and in tissue are needed to confirm data obtained. Moreover, experiments addressing functional implications would require to confirm the suggested role of BCL3 in the atherosclerotic process.

References

1 Perrotta I. Ultrastructural features of human atherosclerosis. *Ultrastruct Pathol*. 2013; 37(1):43-51.

2 Chistiakov DA, et al. Endothelial Barrier and Its Abnormalities in *Cardiovascular Disease*. Front Physiol. 2015; 6:365.

3 Lacolley P, et al. The vascular smooth muscle cell in arterial pathology: a cell that can take on multiple roles. *Cardiovasc Res*. 2012;95(2):194-204.

4 Alexander MR, Owens GK. Epigenetic control of smooth muscle cell differentiation and phenotypic switching in vascular development and disease. *Annu Rev Physiol*. 2012; 74:13-40.

5 Kwak BR, et al. Biomechanical factors in atherosclerosis: mechanisms and clinical implications. *Eur Heart J*. 2014;35(43):3013-20, 3020a-3020d.

6 Stary HC,et al. A definition of initial, fatty streak, and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Arterioscler Thromb*. 1994 ;14(5):840-56.

7 Stary HC, et al. A definition of advanced types of atherosclerotic lesions and a histological classification of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Arterioscler Thromb Vasc Biol*. 1995;15(9):1512-31.

8 Ghazalpour A, et al. Thematic review series: The pathogenesis of atherosclerosis. Toward a biological network for atherosclerosis. *J Lipid Res*. 2004 ;45(10):1793-805.

9 Libby P, et al. Progress and challenges in translating the biology of atherosclerosis. *Nature*. 2011;473(7347):317-25.

10 Yu XH, et al.Nuclear Factor-κB Activation as a Pathological Mechanism of *Lipid Metabolism and Atherosclerosis*. Adv Clin Chem. 2015;70:1-30.

11 Davis-Dusenbery BN, et al. Micromanaging vascular smooth muscle cell differentiation and phenotypic modulation. *Arterioscler Thromb Vasc Biol*. 2011;31(11):2370-7.

12 Virmani R,et al. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arterioscler Thromb Vasc Biol*. 2000;20(5):1262-75.

13 Abedin M,et al. Vascular calcification: mechanisms and clinical ramifications. *Arterioscler Thromb Vasc Biol.* 2004;24(7):1161-70.

14 Ge J,et al. Screening of ruptured plaques in patients with coronary artery disease by intravascular ultrasound. *Heart*. 1999;81(6):621-7.

15 Chen F, et al. Apoptosis and angiogenesis are induced in the unstable coronary atherosclerotic plaque. *Coron Artery Dis*.2005;16(3):191-7.

16 Lee RT, Libby P. The unstable atheroma. *Arterioscler Thromb Vasc Biol*. 1997;17(10):1859-67.

17 Mathiesen EB, Johnsen SH. Ultrasonographic measurements of subclinical carotid atherosclerosis in prediction of ischemic stroke. *Acta Neurol Scand Suppl*. 2009;(189):68-72.

18 Johnston SC, et al. Cognitive impairment and decline are associated with carotid artery disease in patients without clinically evident cerebrovascular disease. *Ann Intern Med*. 2004 Feb 17;140(4):237-47.

19 Romero JR, et al. Carotid artery atherosclerosis, MRI indices of brain ischemia, aging, and cognitive impairment: the Framingham study. *Stroke*. 2009;40(5):1590-6.

20 Nezu T,et al. Carotid Intima-Media Thickness for Atherosclerosis. J *Atheroscler Thromb*. 2016;23(1):18-31.

21 Michel JB, et al. Smooth muscle cells and vascular diseases. *Cardiovasc Res*. 2012;95(2):135-7.

22 Louis SF, Zahradka P. Vascular smooth muscle cell motility: From migration to invasion. *Exp Clin Cardiol*. 2010 ;15(4):e75-85.

23 Coen M, Bochaton-Piallat ML. Phenotype heterogeneity of smooth muscle cells. Implications for atherosclerosis. In: Molecular and Cellular Mechanisms Underlying Atherosclerosis. Edited by George S. and Johnson J. Wiley-VCH. Weinheim. pp.327-42, 2010

24 Doran AC, et al. Role of smooth muscle cells in the initiation and early progression of atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2008;28(5):812-9.

25 Johnson JL. Emerging regulators of vascular smooth muscle cell function in the development and progression of atherosclerosis. *Cardiovasc Res*. 2014;103(4):452-60.

26 Motterle A, et al. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. *Hum Mol Genet*. 2012; 21(18):4021-9.

27 Perisic L, et al. Profiling of atherosclerotic lesions by gene and tissue microarrays reveals PCSK6 as a novel protease in unstable carotid atherosclerosis. *Arterioscler Thromb Vasc Biol*. 2013;33(10):2432-43.

28 Stanzione R, et al. C2238/αANP modulates apolipoprotein E through Egr-1/miR199a in vascular smooth muscle cells in vitro. *Cell Death Dis*.2015;6:e2033.

29 Oeckinghaus A, Ghosh S. The NF-kappaB family of transcription factors and its regulation. *Cold Spring Harb Perspect Biol.* 2009;1(4):a000034.

30 Manavalan B,et al. Structure-function relationship of cytoplasmic and nuclear IκB proteins: an in silico analysis. *PLoS One*. 2010;5(12):e15782.

31 Martín-Ventura JL, et al. NF-kappaB activation and Fas ligand overexpression in blood and plaques of patients with carotid atherosclerosis: potential implication in plaque instability. *Stroke*. 2004;35(2):458-63

32 Pamukcu B,et al. The nuclear factor-kappa B pathway in atherosclerosis: a potential therapeutic target for atherothrombotic vascular disease. *Thromb Res.* 2011;128(2):117-23.

33 De Winther MP, et al. Nuclear factor kappaB signaling in atherogenesis. *Arterioscler Thromb Vasc Biol*. 2005;25(5):904-14.

34 Kanters E,et al. Hematopoietic NF-kappaB1 deficiency results in small atherosclerotic lesions with an inflammatory phenotype. *Blood*. 2004;103(3):934-40.

35 Moodie FM, et al. Oxidative stress and cigarette smoke alter chromatin remodeling but differentially regulate NF-kappaB activation and

proinflammatory cytokine release in alveolar epithelial cells. *FASEB J*. 2004;18(15):1897-9.

36 McKeithan TW, et al. BCL3 rearrangements and t(14;19) in chronic lymphocytic leukemia and other B-cell malignancies: a molecular and cytogenetic study. *Genes Chromosomes Cancer*. 1997;20(1):64-72.

37 Palmer S, Chen YH. Bcl-3, a multifaceted modulator of NF-kappaB-mediated gene transcription. *Immunol Res*. 2008;42(1-3):210-8.

38 Ohno H, et al. The candidate proto-oncogene bcl-3 is related to genes implicated in cell lineage determination and cell cycle control. *Cell*. 1990;60(6):991-7.

39 Nolan GP, et al. The bcl-3 proto-oncogene encodes a nuclear I kappa B-like molecule that preferentially interacts with NF-kappa B p50 and p52 in a phosphorylation-dependent manner. *Mol Cell Biol*. 1993 Jun;13(6):3557-66.

40 Wulczyn FG, et al. Candidate proto-oncogene bcl-3 encodes a subunit-specific inhibitor of transcription factor NF-kappa B. *Nature*. 199;358(6387):597-9.

41 Massoumi R, et al. Cyld inhibits tumor cell proliferation by blocking Bcl-3-dependent NF-kappaB signaling. *Cell*.2006;125(4):665-77.

42 Roberts R. Genetics of coronary artery disease. *Circ Res*. 2014;114(12):1890-903.

43 Bijnens AP, et al. Genome-wide expression studies of atherosclerosis: critical issues in methodology, analysis, interpretation of transcriptomics data. *Arterioscler Thromb Vasc Biol*. 2006;26(6):1226-35.

44 Jia P, Zhao Z. Network.assisted analysis to prioritize GWAS results: principles, methods and perspectives. *Hum Genet*. 2014;133(2):125-38.

45 Bampali K, et al. Genetics and coronary artery disease: present and future. *Hellenic J Cardiol*. 2014 Mar-Apr;55(2):156-63.

46 McPherson R,et al. A common allele on chromosome 9 associated with coronary heart disease. *Science*. 2007;316(5830):1488-91.

47 Helgadottir A, et al. A common variant on chromosome 9p21 affects the risk of myocardial infarction. *Science*. 2007;316(5830):1491-3.

48 Shen GQ, et al. Association between four SNPs on chromosome 9p21 and myocardial infarction is replicated in an Italian population. *J Hum Genet*. 2008;53(2):144-50.

49 Coronary Artery Disease Consortium, Samani NJ, et al. Large scale association analysis of novel genetic loci for coronary artery disease. *Arterioscler Thromb Vasc Biol*. 2009; 29(5):774-80.

50 Hinohara K, et al. Replication of the association between a chromosome 9p21 polymorphism and coronary artery disease in Japanese and Korean populations. *J Hum Genet.* 2008;53(4):357-9.

51 Hannou SA, et al. Functional genomics of the CDKN2A/B locus in cardiovascular and metabolic disease: what have we learned from GWASs? *Trends Endocrinol Metab*. 2015;26(4):176-84.

52 Myocardial Infarction Genetics Consortium, Kathiresan S, et al. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. *Nat Genet*. 2009;41(3):334-41.

53 Ding K, Kullo IJ. Genome-wide association studies for atherosclerotic vascular disease and its risk factors. *Circ Cardiovasc Genet*. 2009 Feb;2(1):63-72.

54 Thorgeirsson TE, et al. A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. *Nature*. 2008;452(7187):638-42.

55 Frayling TM, Et al. A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science*. 2007;316(5826):889-94.

56 Cappola TP, Margulies KB. Functional genomics applied to cardiovascular medicine. *Circulation*. 2011;124(1):87-94.

57 Malone JH, Oliver B. Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol*. 2011;9:34.

58 Raghavachari N. Microarray technology: basic methodology and application in clinical research for biomarker discovery in vascular diseases. *Methods Mol Biol*. 2013;1027:47-84.

59 Kolovou G, et al. Apolipoprotein E knockout models. *Curr Pharm Des*. 2008;14(4):338-51.

60 Okami N, et al. Laser microdissection-based analysis of hypoxia- and thioredoxin-related genes in human stable carotid plaques. *Cardiovasc Pathol*. 2009;18(5):294-300

61 Li P, et al. MicroRNA-638 is highly expressed in human vascular smooth muscle cells and inhibits PDGF-BB-induced cell proliferation and migration through targeting orphan nuclear receptor NOR1. *Cardiovasc Res.* 2013;99(1):185-93.

62 Girelli D, et al. Hyperhomocysteinemia and mortality after coronary artery bypass grafting. *PLoS One.* 2006 Dec 20;1:e83.

63 Bozzini C, et al. Influence of polymorphisms in the factor VII gene promoter on activated factor VII levels and on the risk of myocardial infarction in advanced coronary atherosclerosis. *Thromb Haemost*. 2004;92(3):541-9.

64 North American Symptomatic Carotid Endarterectomy Trial Collaborators. Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. *N Engl J Med*. 1991;325(7):445-53.

65 Skalli O, et al. A monoclonal antibody against alpha-smooth muscle actin: a new probe for smooth muscle differentiation. *J Cell Biol*. 1986 Dec;103(6 Pt 2):2787-96.

66 Aranda PS,et al. Bleach gel: a simple agarose gel for analyzing RNA quality. *Electrophoresis*. 2012;33(2):366-9.

67 Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. *Nucleic Acids Res*. 2007 ;35(Web Server issue):W43-6.

68 Bookout AL, Mangelsdorf DJ. Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. *Nucl Recept Signal*. 2003;1:e012

69 Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc*. 2008;3(6):1101-8.

70 Cui X, Churchill GA. Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol*. 2003;4(4):210.

71 Lake SL, et al. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. *Hum Hered*. 2003;55(1):56-65.

72 Shalhoub J, et al. Systems biology of human atherosclerosis. *Vasc Endovascular Surg.* 2014;48(1):5-17

73 Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. *Cell*.2012;148(6):1242-57.

74 Barsh GS, et al. Guidelines for genome-wide association studies. *PLoS Genet*. 2012;8(7):e1002812.

75 Marian AJ, Belmont J. Strategic approaches to unraveling genetic causes of cardiovascular diseases. *Circ Res.* 2011;108(10):1252-69.

76 Stahl EA,et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. *Nat Genet*. 2012;44(5):483-9.

77 Park JH, et al.Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nat Genet*. 2010;42(7):570-5.

78 Reilly MP, et al. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *Lancet*. 2011;377(9763):383-92.

79 Singh V, et al. Models to study atherosclerosis: a mechanistic insight. *Curr Vasc Pharmacol*. 2009;7(1):75-109.

80 Ahmed S, et al. Evaluation of optimal RNA extraction method from human carotid atherosclerotic plaque. *Cardiovasc Pathol*. 2015 ;24(3):187-90.

81 Sen SK, et al. Integrative DNA, RNA, and protein evidence connects TREML4 to coronary artery calcification. *Am J Hum Gen*et. 2014;95(1):66-76.

82 Oram JF, Vaughan AM. ATP-Binding cassette cholesterol transporters and cardiovascular disease. *Circ Res.* 2006;99(10):1031-43.

83 Van Eck M. ATP-binding cassette transporter A1: key player in cardiovascular and metabolic disease at local and systemic level. *Curr Opin Lipidol*. 2014;25(4):297-303.

84 Fitzgerald ML, et al. ABC transporters, atherosclerosis and inflammation. *Atherosclerosis*. 2010 Aug;211(2):361-70.

85 Talmud PJ, et al. Gene-centric association signals for lipids and apolipoproteins identified via the Human CVD BeadChip. *Am J Hum Genet*. 2009;85(5):628-42

86 Erbilgin A, et al. Gene expression analyses of mouse aortic endothelium in response to atherogenic stimuli. *Arterioscler Thromb Vasc Biol.* 2013;33(11):2509-17.

87 Fransen K,et al.Analysis of SNPs with an effect on gene expression identifies UBE2L3 and BCL3 as potential new risk genes for Crohn's disease. *Hum Mol Genet*. 2010;19(17):3482-8.

88 Miller AM, McInnes IB. Cytokines as therapeutic targets to reduce cardiovascular risk in chronic inflammation. *Curr Pharm Des*. 2011;17(1):1-8.

89 Gao Y,et al. Bioinformatics analyses of differentially expressed genes associated with acute myocardial infarction. *Cardiovasc Ther*. 2016 Jan 4. [Epub ahead of print]

90 Ayari H, Bricca G. Identification of two genes potentially associated in iron-heme homeostasis in human carotid plaque using microarray analysis. *J Biosci*. 2013;38(2):311-5.

91 Hinz M, et al. It takes two to tango: IκBs, the multifunctional partners of NF-κB. *Immunol Rev*. 2012;246(1):59-76.

92 Yang J, et al. Bcl3 interacts cooperatively with peroxisome proliferator-activated receptor gamma (PPARgamma) coactivator 1alpha to coactivate nuclear receptors estrogen-related receptor alpha and PPARalpha. *Mol Cell Biol*. 2009;29(15):4091-102.

93 Alberti KG,et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation*. 2009;120(16):1640-5.

94 Grundy SM. Metabolic syndrome: a multiplex cardiovascular risk factor. *J Clin Endocrinol Metab*. 2007;92(2):399-404.

Appendix

Publication related to this project of thesis:

Marchetti G, Girelli D, Zerbinati C, Lunghi B, Friso S, **Meneghetti S**, Coen M, Gagliano T, Guastella G, Bochaton-Piallat ML, Pizzolo F, Mascoli F, Malerba G, Bovolenta M, Ferracin M, Olivieri O, Bernardi F, Martinelli N.

An integrated genomic-transcriptomic approach supports a role for the proto-oncogene BCL3 in atherosclerosis.

Thromb Haemost. 2015;113(3):655-63.

An integrated genomic-transcriptomic approach supports a role for the proto-oncogene BCL3 in atherosclerosis

Giovanna Marchetti[1]; Domenico Girelli[2]; Carlotta Zerbinati[1]; Barbara Lunghi[3]; Simonetta Friso[2]; Silvia Meneghetti[1]; Matteo Coen[4]; Teresa Gagliano[5]; Giuseppe Guastella[1]; Marie-Luce Bochaton-Piallat[4]; Francesca Pizzolo[2]; Francesco Mascoli[6]; Giovanni Malerba[7]; Matteo Bovolenta[3]; Manuela Ferracin[8,9]; Oliviero Olivieri[2]; Francesco Bernardi[3]; Nicola Martinelli[2]

[1]Department of Biomedical and Specialty Surgical Sciences, University of Ferrara, Ferrara, Italy; [2] Department of Medicine, University of Verona, Verona, Italy; [3]Department of Life Sciences and Biotechnology, University of Ferrara, Ferrara, Italy; [4]Department of Pathology and Immunology, University of Geneva, Geneva, Switzerland; [5]Department of Medical Sciences, University of Ferrara, Ferrara, Italy; [6]Unit of Vascular and Endovascular Surgery, S. Anna University-Hospital, Ferrara, Italy; [7]Department of Life and Reproduction Sciences, Section of Biology and Genetics, University of Verona, Verona, Italy; [8]Department of Morphology, Surgery and Experimental Medicine, University of Ferrara, Ferrara, Italy; [9]Laboratory for Technologies of Advanced Therapies (LTTA), University of Ferrara, Ferrara, Italy

Summary

Data with border-line statistical significance, copiously generated in genome-wide association studies of coronary artery disease (CAD), could include functionally relevant associations. We propose an integrated genomic and transcriptomic approach for unravelling new potential genetic signatures of atherosclerosis. Fifteen among 91 single nucleotide polymorphisms (SNPs) were first selected for association in a sex- and age-adjusted model by examining 510 patients with CAD and myocardial infarction and 388 subjects with normal coronary arteries (CAD-free) in the replication stages of a genome-wide association study. We investigated the expression of 71 genes proximal to the 15 tag-SNPs by two subsequent steps of microarray-based mRNA profiling, the former in vascular smooth muscle cell populations, isolated from non-atherosclerotic and atherosclerotic human carotid portions, and the latter in whole carotid specimens. BCL3 and PVRL2, contiguously located on chromosome 19, and ABCA1, extensively investigated before, were found to be differentially expressed. BCL3 and PVRL2 SNPs were genotyped within a second population of CAD patients (n=442) and compared with CAD-free subjects (n=393). The carriership of the BCL3 rs2965169 G allele was more represented among CAD patients and remained independently associated with CAD after adjustment for all the traditional cardiovascular risk factors (odds ratio=1.70 with 95% confidence interval 1.07–2.71), while the BCL3 rs8100239 A allele correlated with metabolic abnormalities. The up-regulation of BCL3 mRNA levels in atherosclerotic tissue samples was consistent with BCL3 protein expression, which was detected by immunostaining in the intima-media of atherosclerotic specimens, but not within non-atherosclerotic ones. Our integrated approach suggests a role for BCL3 in cardiovascular diseases.

Keywords

BCL3, genomics, transcriptomics, coronary artery disease, metabolic disorders

Correspondence to:
Nicola Martinelli
Department of Medicine, University of Verona
37134 Verona, Italy
E-mail: nicola.martinelli@univr.it

Financial support:
This study was supported by grants from the Italian Ministry of University and Research; the Veneto Region; the Cariverona Foundation, Verona; the Carife Foundation, Ferrara; the Italian Ministry of Health (Finalized Research "Emilia Romagna Region").

Received: May 26, 2014
Accepted after major revision: October 8, 2014
Epub ahead of print: November 6, 2014

http://dx.doi.org/10.1160/TH14-05-0466
Thromb Haemost 2015; 113: 655–663

Introduction

Cardiovascular disease (CVD) is a leading health problem in the world, and inherited DNA sequence variants are well recognised to influence CVD risk (1, 2). Genome-wide association studies (GWAS) have identified a large number of CVD-associated loci (3, 4) with several of them containing genes not previously implicated in the traditional pathways of CVD. Functional explanations are still lacking for the majority of gene variants identified by GWAS, but required to elucidate the molecular mechanisms with potential application in clinical practice (5-11). Moreover, it is worthy to note that in spite of the abundance of results obtained by GWAS the identified loci explain only a limited proportion of the whole CVD heritability (1). The stringent level of statistical significance required in GWAS may lead to discard genetic variants that could contribute to disease risk (12). Recent observations suggest that GWAS databases likely include a number of "hidden" effective variants that up to now have only suggestive statistical evidence of association (13, 14).

Atherosclerosis is the common ground of several clinical manifestations of CVD, including coronary artery disease (CAD), myocardial infarction (MI), peripheral artery occlusive diseases, and

stroke (15, 16). Vascular smooth muscle cells (VSMCs), which are the only cell type of the media layer of normal vessel wall, play also a pivotal role in development and progression of atherosclerotic plaques by virtue of their remarkable plasticity and phenotypic modulation (17-22). VSMCs are recognised as one of the suitable cellular models to detect gene expression signatures underlying molecular processes in atherosclerosis, as well as for functional validation of genes identified/refined through GWAS (23, 24).

With the aim to unravel new potential genetic signatures of the atherosclerotic phenotype, we used an integrated approach, joining information from i) gene polymorphisms analysis in a case-control study of subjects with or without angiographically demonstrated CAD, ii) microarray-based gene expression analysis on human cultured VSMCs and on carotid artery specimens, and iii) immunohistochemical analysis in carotid artery specimens.

Methods

A schematic representation of the study design is summarised in Suppl. Figure 1 (available online at www.thrombosis-online.com).

Study population

This study was performed within the VHS framework, a regional survey designed for identification of new risk factors for CAD in subjects with objective angiographic documentation of their coronary vessels. Details about enrolment criteria have been previously described (25, 26). Subjects with proven CAD had at least one of the main epicardial coronary arteries affected (left anterior descending, circumflex, or right) with ≥1 significant stenosis (≥50%). CAD patients were classified into MI and non-MI groups on the basis of a thorough review of medical records including history, electrocardiogram, enzyme changes, and/or the typical sequelae of MI on ventricular angiography. CAD-free subjects had completely normal coronary arteries, being submitted to coronary angiography for reasons other than CAD, mainly valvular heart disease. These controls were also required to have neither history nor clinical or instrumental evidence of atherosclerosis in vascular districts beyond the coronary bed. All participants came from the same geographical area (North-East Italy). At the time of blood sampling, a complete clinical history was collected, including the assessment of cardiovascular risk factors such as obesity, smoking, hypertension and diabetes. As regards biochemical analysis, samples of venous blood were withdrawn from each subject, after an overnight fast. A detailed description of laboratory methods is provided in the Suppl. Methods (available online at www.thrombosis-online.com).

The first study population was represented by 510 patients with a history of MI before the age of 65 years and 388 CAD-free subjects who were included in the MIGen Consortium as replication population of a GWAS (27). The second study population analysis, addressing specifically the atherosclerotic (more than thrombotic) phenotype, included 442 CAD patients without MI history and 393 CAD-free subjects.

Within the second study population in a subgroup of non-diabetic subjects (n=395) with available data for fasting insulin concentration, the HOMA (homeostasis model assessment) score was calculated as surrogate of insulin resistance, using the following formula: [fasting glucose (mmol/l)*fasting insulin (mIU/l)/22.5] (28).

The study was approved by the Ethic Committee of Azienda Ospedaliera Universitaria Integrata, Verona, Italy. Written informed consent was obtained from all the participants after a full explanation of the study.

Table 1: General characteristics of the study populations, with or without coronary artery disease (CAD).

Characteristics	CAD-free (n=393)	CAD (n=442)	P
Age, years	58.9 ± 12.1	63.1 ± 9.1	<0.001 *
Male sex, %	62.8	72.9	0.002 †
BMI, kg/m²	25.4 ± 3.5	26.8 ± 3.4	<0.001 *
Hypertension, %	41.2	72.6	<0.001 †
Smoking, %	41.9	61.3	<0.001 †
Diabetes, %	7.1	20.4	<0.001 †
Glucose, mmol/l	5.47 (5.37–5.58)	5.70 (5.58–5.81)	0.007 *
Creatinine, mmol/l	89.1 (86.5–90.9)	90.0 (88.2–91.8)	0.587 *
Lipid-lowering therapies, %	5.9	31.6	<0.001 †
Total cholesterol, mmol/l	5.44 ± 1.09	5.45 ± 1.15	0.904 *
LDL-cholesterol, mmol/l	3.50 ± 0.93	3.62 ± 0.98	0.098 *
HDL-cholesterol, mmol/l	1.43 ± 0.42	1.22 ± 0.32	<0.001 *
Triglycerides, mmol/l	1.35 (1.28–1.40)	1.68 (1.62–1.77)	<0.001 *
High sensitivity- C-reactive protein, mg/l	1.97 (1.72–2.27)	3.22 (2.83–3.67)	<0.001 *

CAD patients were required to have no history of previous myocardial infarction. * by t-test. † by Chi² test.

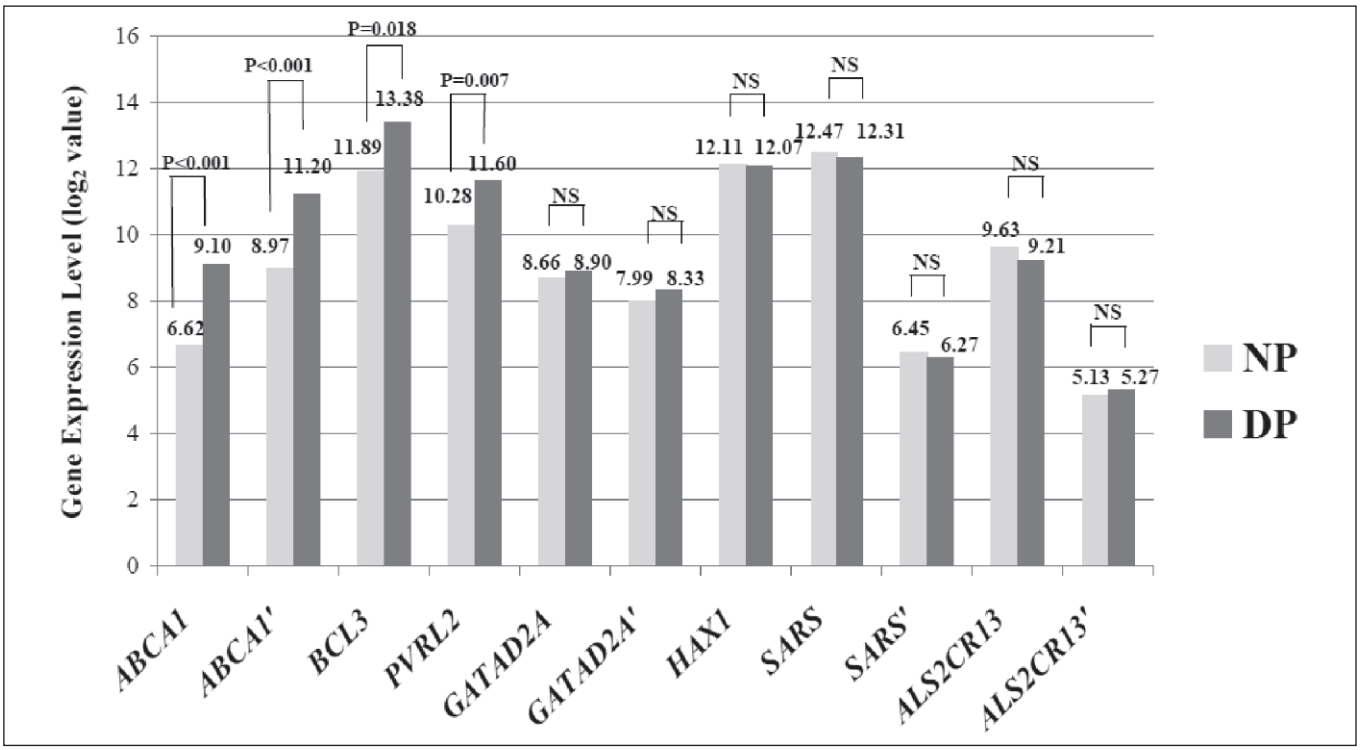


Figure 1: Tissue mRNA expression levels by microarray profiling of atherosclerotic diseased (DP) and non-atherosclerotic (NP) portions of carotid artery. The expression levels of the seven candidate genes, stemming from the preliminary screen in VSMC populations, are reported as log_2 value. Data obtained with two probes, present in the Agilent microarray, are reported for *ABCA1, SARS, GATAD2A* and *ALS2CR13*. NP: grossly non-atherosclerotic carotid portion; DP: atherosclerotic diseased portion; NS: not significant.

SNPs and genotyping

Genomic DNA was prepared from whole blood samples by phenol-chloroform extraction. The 91 SNPs investigated as replication stage of MIGen study, including intergenic rs10402271 polymorphism (*BCL3/PVRL2* locus), were genotyped using the iPLEX MassARRAY platform (Mass Array, Sequenom). PCR primers were designed by Sequenom Mass-Array-Assay-Design program, as previously described (27). The *BCL3* polymorphisms, rs2965169 (5' UTR, NM_005178.4: c.-892T>G) and rs8100239 (intron 1, NM_005178.4: c.256+801T>A) and the *PVRL2* polymorphisms, rs3810143 (5' UTR, NM_001042724.1: c.-381T>C or NM_002856.2: c.-381T>C) and rs1871047 (intron 1, NM_001042724.1: c.88+1876A>G or NM_002856.2: c.88+1876A>G) were genotyped by allele-specific real-time PCR (TaqMan SNP Genotyping Assays, Applied Biosystems, Foster City, CA, USA).

Collection of carotid specimens and histological-immunohistochemical analysis

Carotid artery specimens were obtained from 58 patients who underwent carotid endarterectomy (CEA) at the Unit of Vascular and Endovascular Surgery of S. Anna University-Hospital (Ferrara, Italy) for extracranial high-grade (>70%) internal carotid artery stenosis (29) by the same surgeon (FM). Restenotic lesions were excluded. The study was approved by the Ethic Committee of the University-Hospital and written informed consent was obtained from all patients.

Arteriotomy was performed on the common carotid artery and extended to the internal carotid artery (Suppl. Figure 2, available online at www.thrombosis-online.com). The CEA specimen was removed as a single piece and included the atheromatous area and two small adjacent areas, proximal and cranial, respectively, without clear evidence of atherosclerotic lesions. CEA specimen was cut transversally at the bifurcation and the portion towards the aortic arch was used. This segment, consisting of the common carotid artery, was further cut into: a proximal small portion without evidence of atherosclerotic lesions (grossly non-atherosclerotic portion, NP), and the distal one (cephalad), characterised by diffuse atherosclerosis (diseased portion, DP).

Although the NP showed some histological alterations, like a thin thickened intima, this portion was the available and suitable endogenous control.

NP and DP portions were processed for i) histology-immunohistochemistry analysis, ii) total RNA extraction, and iii) VSMC cultures. Portions for total RNA isolation were immediately placed into RNAlater (Ambion Inc., Austin, TX, USA) and then stored at -80°C.

Sections of formaldehyde-fixed paraffin-embedded NP and DP were stained with Masson's trichrome for histological analysis (Suppl. Figure 2, available online at www.thrombosis-online.com).

Immunostaining on paraffin sections was performed as previously described (30) by using the following primary mouse monoclonal antibodies: IgG2a for α-smooth muscle actin (α-SMA, clone 1A4) (31), IgG1 for CD68 (clone KP1; Dako, Glostrup, Denmark) and IgG2a for BCL3 (clone1E8, Abcam, Cambridge, UK). Before using the first antibody, immunoreactivity was intensified by pressure cooker (3 minutes [min]) for BCL3 or microwave treatment (750 W, 5 min) for α-SMA and CD68.

VSMC culture

Primary VSMC cultures were obtained from NP and DP carotid specimens collected at surgery, as previously reported (28). After luminal gentle scraping to remove endothelial tissue, the NP (intimal thickening and media) and the DP (plaque with underlying media) were cut into 3x3-mm pieces and plated on separate dishes with the abluminal side in contact with the culture dish. Tissue pieces were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA, USA), 100 U/ml penicillin, 100 μg/ml streptomycin and 2 mM L-glutamine, at 37°C with 5% CO_2. Explanted tissues were removed 7-10 days after the first VSMCs appeared. VSMC populations from both NP and DP were maintained in identical baseline culture conditions and studied at the third passage. SMC lineage was confirmed by immunofluorescence staining using a mouse IgG2a recognising α-smooth muscle actin (α-SMA, clone 1A4) (31) and rabbit polyclonal IgGs recognising both SMMHC types 1 and 2 (BT-562, Biomedical Technologies Inc, Stoughton, MA, USA).

Table 2: Distribution of *BCL3* and *PVRL2* genotypes in the study populations, with or without coronary artery disease (CAD).

Genotypes	CAD-free (n=393)	CAD (n=442)	P
rs2965169 (*BCL3*)			
TT	29.3	21.8	
TG	49.4	53.1	0.020 *
GG	21.4	25.2	
rs8100239 (*BCL3*)			
TT	33.1	27.1	
TA	45.8	50.5	0.146 *
AA	21.1	22.4	
rs10402271 (intergenic *BCL3-PVRL2*)			
TT	50.1	46.7	
TG	41.3	41.1	0.168 *
GG	8.6	12.2	
rs3810143 (*PVRL2*)			
TT	32.5	31.7	
TC	46.1	49.4	0.753 *
CC	21.4	18.9	
rs1871047 (*PVRL2*)			
AA	36.6	33.9	
AG	45.3	46.2	0.372 *
GG	18.1	19.9	

* by Chi^2 test for linear trend.

RNA expression study

Total RNA was isolated either from confluent cultured VSMCs or homogenised NP and DP portions with TRIzol Reagent (Invitrogen) according to the manufacturer's instructions and treated with RNase-free DNase (New England Biolabs, Ipswich, MA, USA). RNA quality was assessed on an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA) and low-quality RNA (RNA integrity number below 7) was excluded from further analyses.

Microarray profiling of cultured VSMCs

Labelled cRNA was synthesised from 500 ng of total RNA isolated from cultured VSMCs using the Low RNA Input Linear Amplification Kit (Agilent Technologies) in the presence of cyanine 3-CTP (Perkin-Elmer Life Sciences, Boston, MA, USA). Global gene expression was detected using Agilent Whole Human Genome microarray (Cat.No. G4112F, Agilent Technologies), which represents 41,000 unique human transcripts. RNA labelling and hybridisation were performed in accordance to manufacturer's indications. Feature Extraction 10.7 software (Agilent Technologies) was used to obtain the microarray raw-data. Microarray results were analysed using GeneSpring GX 12 software (Agilent Technologies). Data transformation was applied to set all negative raw values at 1.0, followed by a normalisation on 75th percentile.

Microarray profiling of whole carotid artery specimens

RNA from six specimens was hybridised on Agilent whole human genome oligo microarray (Cat.No. G4851A, Agilent Technologies) which represents 60,000 unique human transcripts. RNA labelling and hybridisation were performed in accordance to manufacturer's indications. Feature Extraction software v.10.7 (Agilent Technologies) was used to obtain the microarray raw-data. Microarray results were analysed using the GeneSpring GX 12 software (Agilent Technologies). Data transformation was applied to set all negative raw values at 1.0, followed by a quantile normalisation.

cDNA preparation and real-time quantitative polymerase chain reaction (qPCR)

cDNA was obtained from 1 μg (cultured cells) or 0.5 μg (carotid tissues) of total RNA by reverse transcription using SuperScript VILO cDNA Synthesis Kit (Invitrogen) according to the manufacturer's recommendations. Aliquots of diluted first-strand cDNA were amplified on CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA) using SsoFast EvaGreen Supermix (Bio-

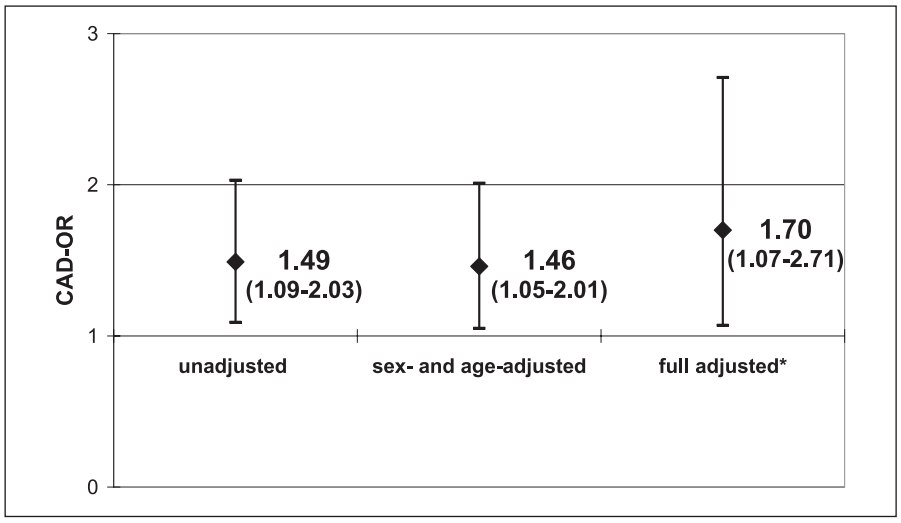


Figure 2: Association between the carriership of *BCL3* rs2965169 minor allele G and coronary artery disease (CAD). The strength of the association was estimated by means of ORs with 95% CI at univariate, sex- and age-adjusted, and full adjusted analysis. The homozygotes for the major allele (TT) were considered as reference. * by logistic regression analysis adjusted for sex, age, BMI, smoking, hypertension, diabetes, LDL- and HDL-cholesterol, triglycerides, creatinine, and hs-CRP.

Rad). PCR protocol was: 95°C for 30 seconds (sec), then 40 cycles of 5 sec at 95°C and 10 sec at 60°C.

Forward and reverse primers are reported in the Suppl. Table 1 (available online at www.thrombosis-online.com). Each reaction was performed in triplicate. The relative levels of mRNAs were calculated by the comparative CT method by using 18S rRNA as endogenous control in cell populations or ACTB in carotid specimens. Values were expressed as mean fold change ± standard error of the mean (SEM).

Statistical analysis

Differences in mRNA expression levels between VSMC populations by microarray-based transcriptome analysis were evaluated by using the log-transformed (log_2) ratio values (DP/NP) and assuming the log ratio equal to 0 when there is no difference of expression (32). A p-value <0.05 was considered significant without correction for multiple testing. Relationships between the fold-change (log_2 DP/NP ratio) and the p-value (negative log_{10}) were reported in a volcano plot.

Genes differentially expressed between three atherosclerotic and three non-atherosclerotic carotid specimens in microarray experiment were identified using GeneSpring GX software. A filter on fold change ≥1.5 was applied to the list of probes expressed in at least one sample. Then, a moderated t-test was used to identify the significantly modulated genes with p<0.05 and false discovery rate (FDR) =10%. Gene expression levels between VSMC populations or carotid portions in q-PCR analysis were compared by means of paired or unpaired t-test.

As regards population data, distributions of continuous variables in groups were expressed as means ± standard deviation (SD). Statistical analysis on skewed variables, like glucose, creatinine, triglyceride, and high-sensitivity C-reactive protein (hs-CRP), was computed on the corresponding log-transformed values. Thus, results are reported as geometric mean with 95% confidence interval (95% CI). Quantitative data were assessed using the Student's t-test or analysis of variance (ANOVA), with polynomial contrast for linear trend when indicated. Qualitative data were analysed with either Chi^2-test or Chi^2 for linear trend analysis when indicated. Within each group examined, the frequencies of the genotypes associated with each of the polymorphisms were compared by the Chi^2-test with the values predicted on the basis of the Hardy-Weinberg equilibrium.

In the first case-control population the strength of association with CAD/MI was assessed by means of a sex- and age-adjusted model. In the second case-control population the strength of association with CAD was estimated calculating the odds ratios (ORs) with 95% CIs by multiple logistic regression after sex- and age-adjustment and then after adjustment for all the traditional cardiovascular risk factors (i.e. sex, age, body mass index [BMI], smoking, hypertension, diabetes, LDL- and HDL-cholesterol, triglycerides, creatinine, and hs-CRP). A value of p<0.05 was considered statistically significant. Statistical power was estimated by Altman nomogram (33). Calculations were performed with IBM SPSS 20.0 statistical package (IBM Inc., Armonk, NY, USA).

Haplotype frequencies were estimated using the R software with haplo.stats package (R Foundation for Statistical Computing, Vienna, Austria; http: //www.R-project.org). The associations between haplotypes and laboratory and clinical outcomes were examined using a generalised linear model regression of a trait on haplotype effects, allowing for ambiguous haplotypes (haplo.glm function) (34). Statistical significance of associations was ascertained by randomly permuting the disease status in 1,000 replicates by Monte Carlo method.

Results

SNPs selection

Starting from a previous analysis of 91 SNPs in 898 subjects of the VHS population (388 CAD-free and 510 CAD with MI) within stages of replication of GWAS – MIGen – for cardiovascular risk (27), we selected 15 SNPs showing nominally an association with MI by means of a sex- and age-adjusted model with P values arbit-

rarily set at <0.10 (Suppl. Table 2, available online at www.thrombosis-online.com).

As more than half (8/15) of the selected SNPs localise within intergenic regions, the genomic regions tagged by the 15 SNPs were examined in the NCBI database for the presence of validated or putative coding gene sequences within a region of 400 kb centred on each SNP. This search identified 77 genes, listed in Suppl. Table 3 (available online at www.thrombosis-online.com).

Transcriptome analysis of cultured VSMCs and of carotid specimens

As a preliminary screen to identify, among the 77 genes, candidates for further evaluation, we performed a microarray-based transcriptome analysis of VSMC populations, cultured from NP: DP couples collected from three patients. The Agilent whole human genome microarray (44k) included probes for 71 of the 77 neighbouring genes. The analysis revealed significant differences (P<0.05) in the expression levels for seven genes (*ABCA1*, *ALS2CR13, PVRL2, BCL3, HAX1, SARS, GATAD2A,* Suppl.Table 4 and Suppl. Figure 3, both available online at www.thrombosis-online.com). Validation was conducted by qPCR analyses in independent cell populations (see Suppl. Results, available online at www.thrombosis-online.com).

To explore also in the vessel wall the differences in the expression levels detected in cultured VSMCs, we performed a second microarray experiment of whole atherosclerotic and non-atherosclerotic specimens, collected from four additional patients. The comparison (fold change ≥1.5, p<0.05, FDR 10%) of DP and NP tissues showed that three (*ABCA1*, *BCL3* and *PVRL2*) out the seven genes, identified in the first transcriptome screen, were significantly up regulated in DP (►Figure 1). This analysis revealed the most significant differences for the ATP-binding membrane cassette transporter A1 (ABCA1), whose role in lipid metabolism and in atherosclerosis pathways is well recognised.

We addressed then our attention to the *BCL3* and *PVRL2* genes, contiguously located on the chromosome region (19q13) marked by the intergenic rs10402271. With the aim to validate findings from microarray profiling of specimens, qPCR analyses were performed on carotid tissue samples from independent patients. *BCL3* mRNA levels were higher in DP than in NP specimens by comparison of i) DP: NP couples (five patients), with a mean fold change difference of 10.07 ± 2.87 (p=0.034) as well as ii) non-coupled DP and NP (14 patients), with a mean fold change difference of 6.81 ± 3.16 (p=0.013). PVRL2 expression levels were significantly higher in DP by comparison of couples (mean fold change 2.60 ± 0.46, p=0.026) and, as a trend, in non- coupled DP and NP (mean fold change 2.52 ± 0.80; p=0.060).

BCL3 SNPs genotyping for association with CAD

As the originally investigated rs10402271 maps in the intergenic 19q13 region between *BCL3* and *PVRL2*, intragenic SNPs, two for *BCL3* (rs2965169 and rs8100239, previously associated to oral clefts) (35), and two for *PVRL2* (rs3810143 and rs1871047), were selected for the second genetic association analysis. It is worthy to note that the low MAF value (<0.02, by NCBI) of cDNA SNPs, both in *BCL3* and *PVRL2*, makes them not eligible for this study.

We compared CAD-free controls (n=393) with a second population of CAD patients without previous MI (n=442). This cohort enabled us to preferentially investigate the atherosclerotic phenotype, and to avoid the potential biases due to selection/survival related to MI history.

The clinical and laboratory characteristics of the two groups are summarised in ►Table 1, while the data about genotype prevalence of the five SNPs are reported in ►Table 2. Genotypes had a distribution consistent with Hardy-Weinberg equilibrium in both cases and controls and showed a low or moderate degree of linkage disequilibrium (Suppl. Figure 4, available online at www.thrombosis-online.com), which were in accordance with Hap Map data.

While the *PVRL2* rs3810143 and rs1871047 and the intergenic rs10402271 did not show a significant association with CAD, the *BCL3* rs2965169 genotype distribution differed significantly between cases and controls (►Table 2). The carriership of the minor allele G was more frequent in CAD patients (►Table 2) and remained independently associated with CAD in different adjusted logistic regression models (►Figure 2).

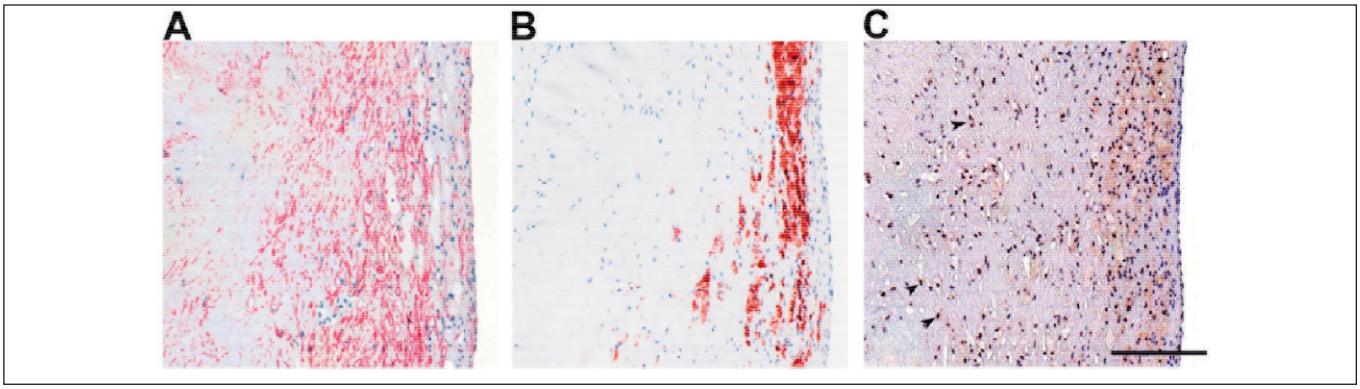


Figure 3: BCL3 protein expression in atherosclerotic (DP) carotid specimens. Representative immunohistochemical staining of the DP plaque for α-SMA (A), CD68 (B) and BCL3 (C). BCL3 is expressed by plaque foam cells and rare VSMCs (arrows). The lumen is located on the right side of the pictures. Bar = 200 µm.

Table 3: Distribution of *BCL3* haplotypes in the whole study population and in subjects with or without coronary artery disease (CAD).

	BCL3 polymorphisms		Haplotype frequency (%)		
	rs2965169	rs8100239	Whole study population (n=835)	CAD-free subjects (n=393)	CAD patients (n=442)
Haplotype 1	T	T	46.9	49.3	44.9
Haplotype 2 *	G	A	41.9	39.2	44.3
Haplotype 3	G	T	7.1	6.7	7.4
Haplotype 4	T	A	4.1	4.8	3.4

* subjects with G-A haplotype carried a greater risk of CAD than subjects with T-T haplotype (OR 1.25 with 95%CI 1.02–1.53, p=0.035; p=0.026 in a sex- and age-adjusted model after 1,000 permutations by the Monte Carlo method).

The carriership of the *BCL3* rs8100239 minor allele A appeared to be more represented in CAD patients, but did not reach a statistically significant difference (72.9% vs 66.9%, p=0.062). Further analyses showed association between rs8100239 and some intermediate metabolic phenotypes, with the A allele linked with insulin resistance-related conditions (see Suppl. Results and Suppl. Tables 5A and 6, all available online at www.thrombosis-online.com). On the other hand, associations of *BCL3* rs2965169 (Suppl. Table 5B, available online at www.thrombosis-online.com), *PVRL2* rs3810143 and rs1871047, and intergenic rs10402271 genotypes (data not shown) with lipid profile and metabolic phenotypes were not detected.The haplotype analysis based on *BCL3* rs2965169 and rs8100239 showed that the G-A haplotype, containing the minor alleles, was more represented in patients with CAD than in CAD-free (►Table 3, OR 1.25 with 95%CI 1.02-1.53 as compared with the most frequent T-T haplotype; p=0.035).

Expression of the BCL3 protein in carotid specimens

The expression of BCL3 protein was investigated in five NP and 10 DP specimens by immunohistochemical analysis. A positive staining for BCL3 (►Figure 3) was found in 20-40% of plaque foam cells and in rare intimal and medial VSMCs in nine of 10 DP samples. We did not detect any BCL3-expressing cells in the intima-media of all the NP samples (data not shown).

Discussion

GWAS are powerful tools for unveiling the genetic components of CAD and MI, but their rigorous approach may hide some potential limitations which should be taken into account. The extensive multiple testing correction in GWAS, necessary to exclude false-positive results, may simultaneously discard genetic loci which could contribute to pathophysiology (12-14). Moreover, the considerable overlap among clinical phenotypes of CVD, leads several GWAS to cumulate indifferently CAD and MI in their analysis (3). Recent studies have consistently demonstrated that some SNPs are associated more with CAD than with a specific predisposition to MI (36). Combined approaches exploiting complementary methodologies may help to overcome these limitations and lead to significant advances in understanding the mechanisms of multifactorial diseases, like CVD (37-39).

Our experimental design integrates analyses at DNA, RNA and protein levels.The GWAS-related tag SNPs and two transcriptome analyses, in both cultured cells and tissue, enabled us to propose three candidate genes, *BCL3*, *PVRL2* and *ABCA1*. The last one has been extensively investigated in relation to the pathogenetic mechanisms of atherosclerotic vascular diseases (40).The recognised role of *ABCA1* in CVD, strengthened by the observation that *ABCA1* mutations cause Tangier disease and premature atherosclerosis (41), supports our experimental approach aimed at detecting genetic components of atherosclerotic vascular diseases.

The two other candidate genes (*BCL3* and *PVRL2*) were contiguously located on chromosome 19, in a region previously associated to plasma lipid traits (42), thus increasing the interest for this genomic region. This prompted us to further investigate specific SNPs for *BCL3* and the adjacent *PVRL2*, in a second case-control analysis considering CAD patients without MI history. This step produced coherent results: the carriership of *BCL3* rs2965169 G allele was associated to an increased risk of CAD independently from traditional atherosclerosis risk factors and the *BCL3* rs8100239 A allele, in moderate linkage, tended also to be more represented among CAD patients. Consistently the rs2965169 G - rs8100239 A haplotype was more frequent in CAD patients than in CAD-free subjects.

Although until now classical GWAS approaches have failed to identify *BCL3* as risk gene for CAD (1), a recent study combining GWAS results with expression quantitative trait loci (eQTLs) analysis identified *BCL3* as a potential new risk gene for Crohn's disease (43), a chronic inflammatory disease associated with an increased atherosclerotic risk unexplained by traditional cardiovascular risk factors (44). With regard to the results at RNA level, our findings are also consistent with data from a transcriptome analysis by microarray on human carotid plaques (GEO database, GSE 43292), which shows higher *BCL3* mRNA levels in the majority of atheroma plaques as compared to the macroscopically intact tissues.

Evidence from our study for BCL3 protein expression in atherosclerotic vascular vessel wall, not reported so far, further supports the hypothesis of a role of BCL3 in CVD and fit with RNA data in carotid specimens. We found BCL3 positivity only in atherosclerotic portions and in cells, like foam cells and VSMCs, which are well known to play crucial role in atherogenesis. Actually, we observed very few positive VSMCs, but this result is not surprising taking into account the lack of BCL3 positivity in most of tissues as reported in the Human Protein Atlas database. On the other hand, the very low BCL3 protein expression levels in diseased tissue might not favour to disentangle the interpretation of BCL3 role in atherosclerosis.

BCL3 is a transcriptional coregulator and member of the inhibitor of nuclear factor-κB (IκB) family, which takes part to both positive and negative modulation of genes belonging to several pathways, including those of cell death/proliferation, inflammatory and stress responses (45, 46). Biological information from Gene Ontology database locates BCL3 in multiple processes which are known to be relevant in atherosclerosis (Suppl. Figure 5, available online at www.thrombosis-online.com).

BCL3 was also proposed as a key interface for regulatory crosstalk between inflammation and cellular energy metabolism (47). According with this hypothesis, in our study the *BCL3* rs8100239 A allele correlated with traits usually characterizing the so-called metabolic syndrome, a still debated cluster of risk factors including altered energy metabolism and whose core is thought to be insulin resistance (48, 49). The lack of waist circumference data in our population does not allow to define the conventionally stated diagnosis of metabolic syndrome (50). Nonetheless, we observed significant associations with both unfavorable lipid profile and overweight. Although assessed only in a relatively small group of subjects, the A allele was associated also with both hyperinsulinaemia and higher HOMA score, which is considered a surrogate measure of insulin resistance.

Our study has some limitations that warrant discussion. The allele specific contribution to *BCL3* mRNA expression level, which could provide insight for a functional effect and for the associated risk, is difficult to determine because of the SNPs localisation (5' UTR and intron 1) and the virtual absence of in linkage coding SNPs. On the other hand, at inspection (UCSC Genome Browser, hg19, ENCODE data) for functional elements by H3K27 histone acetylation mark and by chromatin state segmentation analysis, both *BCL3* rs2965169 and rs8100239 polymorphisms have been shown to lie within potentially active regulatory elements in endothelial-, B-lymphocyte- and fibroblast cell lines, which are of interest for atherosclerosis pathways.

The transcriptional analysis was obtained in VSMC populations or in whole specimens from a small number of patients. This limitation has been tackled by i) the study of couples of specimens from the same donor, which would reduce variability in the gene expression profile; ii) the combination of subsequent and independent transcriptome analyses, expected to increase robustness of the approach and to reveal hits of "*in vivo*" relevance; and iii) the validation of transcriptome changes by qPCR. We also disclose that the complex transcriptional and post-transcriptional regulation of *BCL3* (45) could be perturbed by in vitro culture state and thus differ from the *in vivo* condition.

Regarding the statistical analysis, the lack of adjustment for multiple comparisons is a further limitation of our work. We attempted to overcome this drawback not only by requiring case-control analyses in two CAD populations (with or without MI), but also by showing additional support for the association between BCL3 and atherosclerosis at mRNA (in cultured VSMC populations and in carotid tissues) and protein levels.

Other significant caveats of this study are related to the population analysis, including the retrospective case-control design, the relatively low number of subjects, and the lack of some clinical data. On the other hand, a remarkable strength of this study is represented by the angiography-evaluation of the coronary artery bed, which allows a clear-cut definition of the clinical phenotype and avoids the possibility to include in the control group subjects with subclinical, but significant CAD. Finally, considering the different prevalence of the carriership of rs2965169 G allele among subjects with or without CAD in the second case-control analysis, the statistical power of our study was about 80% by Altman nomogram with a significance level of 0.05 (33).

In summary, our integrated approach suggests for the first time the involvement of BCL3 in CVD, which could be partly mediated through the influence on metabolic phenotypes. Our results should be considered as hypothesis-generating and further epidemiological studies on larger samples, as well as basic investigations about BCL3 expression and function, are needed to confirm the role of BCL3 in the pathways of atherosclerosis.

What is known about this topic?

- The stringent level of statistical significance required in genome-wide association studies may lead to discard genetic variants that could contribute to disease risk.
- Integrated genomic and transcriptomic investigations might be a tool to unravel new genetic signatures for cardiovascular disease.
- The expression of the proto-oncogene *BCL3* in human atherosclerotic lesions has not previously been reported.

What does this paper add?

- *BCL3* mRNA and protein were differently expressed in atherosclerotic versus non-atherosclerotic portions of carotid artery.
- *BCL3* genotypes were associated with both coronary artery disease and metabolic derangements.
- The proto-oncogene *BCL3* may play a role in atherosclerosis.

Acknowledgements

We are very grateful to Prof. Sekar Kathiresan (Cardiovascular Research Center and Cardiology Division and Center for Human Genetic Research, Massachusetts General Hospital, Boston, Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA) for the ongoing cooperation and for providing the data about SNPs of

MIGen Consortium. We wish to thank Mrs. Maria Zoppi for her invaluable secretary help, and Dr. Patrizia Guarini, Diego Minguzzi and Patrizia Pattini for their excellent technical help.

Conflicts of interest
None declared.

References

1. Kathiresan S, Srivastava D. Genetics of human cardiovascular disease. Cell 2012; 148: 1242-1257.
2. Roger VL, et al. American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Executive summary: heart disease and stroke statistics-2012 update: a report from the American Heart Association. Circulation 2012; 125: 188-197.
3. Ding K, Kullo IJ. Genome-wide association studies for atherosclerotic vascular disease and its risk factors. Circ Cardiovasc Genet 2009; 2: 63-72.
4. O'Donnell CJ, Nabel EG. Genomics of cardiovascular disease. N Engl J Med 2011; 365: 2098-2109.
5. Bijnens APJJ, et al. Genome-wide expression studies of atherosclerosis-Critical issue in methodology, analysis, interpretation of transcriptomics data. Arterioscler Thromb Vasc Biol 2006; 26: 1226-1235.
6. Miller DT, et al. Atherosclerosis: the path from genomics to therapeutics. J Am Coll Cardiol 2007; 49: 1589-1599.
7. Cappola TP, Margulies KB. Functional genomics applied to cardiovascular medicine. Circulation 2011; 124: 87-94.
8. Park SW, et al. Post-transcriptional regulation of low density lipoprotein receptor protein by proprotein convertase subtilisin/kexin type 9a in mouse liver. J Biol Chem 2004; 279: 50630–50638.
9. Lambert G, et al. Molecular basis of PCSK9 function. Atherosclerosis 2009; 203: 1-7.
10. Musunuru K, et al. From noncoding variant to phenotype via SORT1 at the 1p13 cholesterol locus. Nature 2010; 466: 714-719.
11. Yang X. Use of functional genomics to identify candidate genes underlying human genetic association studies of vascular diseases. Arterioscler Thromb Vasc Biol 2012; 32: 216-222.
12. Marian AJ, Belmont J. Strategic approaches to unravelling genetic causes of cardiovascular diseases. Circ Res 2011; 108: 1252-1269.
13. Stahl EA, et al. Bayesian inference analyses of the polygenic architecture of rheumatoid arthritis. Nat Genet 2012; 44: 483-489.
14. Park JH, et al. Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat Genet 2010; 42: 570-575.
15. Ross R. Mechanisms of disease: Atherosclerosis-An inflammatory disease. N Engl J Med 1999; 340: 115–126.
16. Lusis AJ. Atherosclerosis. Nature 2000; 407: 233-241.
17. Campbell JH, Campbell GR. Smooth Muscle Phenotypic Modulation—A Personal Experience. Arterioscler Thromb Vasc Biol 2012; 32: 1784-1789.
18. Gomez D, Owens GK. Smooth muscle cell phenotypic switching in atherosclerosis. Cardiovasc Res 2012; 95: 156-164.
19. Hao H, et al. Arterial smooth muscle cell heterogeneity: implications for atherosclerosis and restenosis development. Arterioscler Thromb Vasc Biol 2003; 23: 1510-1520.
20. Doran AC, et al. Role of smooth muscle cells in the initiation and early progression of atherosclerosis. Arterioscler Thromb Vasc Biol 2008; 28: 812-819.
21. Lacolley P, et al. The vascular smooth muscle cell in arterial pathology: a cell that can take on multiple roles. Cardiovasc Res 2012; 95: 194-204.
22. Johnson JL. Emerging regulators of vascular smooth muscle cell function in the development and progression of atherosclerosis. Cardiovasc Res 2014; Epub ahead of print.
23. Motterle A, et al. Functional analyses of coronary artery disease associated variation on chromosome 9p21 in vascular smooth muscle cells. Hum Mol Genet 2012; 21: 4021-4029.
24. Perisic L, et al. Profiling of atherosclerotic lesions by gene and tissue microarrays reveals PCSK6 as a novel protease in unstable carotid atherosclerosis. Arterioscler Thromb Vasc Biol 2013; 33: 2432-2443.
25. Bozzini C, et al. Influence of polymorphisms in the factor VII gene promoter on activated factor VII levels and on the risk of myocardial infarction in advanced coronary atherosclerosis. Thromb Haemost 2004; 92: 541-549.
26. Martinelli N, et al. Polymorphisms at LDLR locus may be associated with coronary artery disease through modulation of coagulation factor VIII activity and independently from lipid profile. Blood 2010; 116: 5688-5697.
27. Kathiresan S, Myocardial Infarction Genetics Consortium. Genome-wide association of early-onset myocardial infarction with single nucleotide polymorphisms and copy number variants. Nat Genet 2009; 41: 334-341.
28. Matthews DR, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985; 28: 412-419.
29. North American Symptomatic Carotid EndarterectomyTrial Collaborators. Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. N Engl J Med 1991; 325: 445-453.
30. Coen M, et al. Calmodulin expression distinguishes the smooth muscle cell population of human carotid plaque. Am J Pathol 2013; 183: 996-1009.
31. Skalli O, et al. A monoclonal antibody against alpha-smooth muscle actin: a new probe for smooth muscle differentiation. J Cell Biol 1986; 103: 2787-2796.
32. Cui X, Churchill GA. Statistical tests for differential expression in cDNA microarray experiments. Genome Biol 2003; 4: 210-219.
33. Altman DG. Statistics and ethics in medical research. How large a sample? Br Med J 1980; 281: 1336-1338.
34. Lake SL, et al. Estimation and tests of haplotype-environment interaction when linkage phase is ambiguous. Hum Hered 2003; 55: 56-65.
35. Park BY, et al. Differential parental transmission of markers in BCL3 among Korean cleft case-parent trios. J Prev Med Public Health 2009; 42: 1-4.
36. Reilly MP, et al. Myocardial Infarction Genetics Consortium; Wellcome Trust Case Control Consortium. Identification of ADAMTS7 as a novel locus for coronary atherosclerosis and association of ABO with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. Lancet 2011; 377: 383-392.
37. Folkersen L, et al. BiKE and ASAP study groups. Association of genetic risk variants with expression of proximal genes identifies novel susceptibility genes for cardiovascular disease. Circ Cardiovasc Genet 2010; 3: 365-373.
38. Ramsey SA, et al. A systems biology approach to understanding atherosclerosis. EMBO Mol Med 2010; 2: 79-89.
39. Maouche S, Schunkert H. Strategies beyond genome-wide association studies for atherosclerosis. Arterioscler Thromb Vasc Biol 2012; 32: 170-181.
40. Oram JF, Vaughan AM. ATP-Binding cassette cholesterol transporters and cardiovascular disease. Circ Res 2006; 99: 1031-1043.
41. Fitzgerald ML, et al. ABC transporters, atherosclerosis and inflammation. Atherosclerosis 2010; 211: 361-370.
42. Talmud PJ, et al. Gene-centric association signals for lipids and apolipoproteins identified via the Human CVD BeadChip. Am J Hum Genet 2009; 85: 628-642.
43. Fransen K, et al. Analysis of SNPs with an effect on gene expression identifies UBE2L3 and BCL3 as potential new risk genes for Crohn's disease. Hum Mol Genet 2010; 19: 3482-3488.
44. Miller AM, McInnes IB. Cytokines as therapeutic targets to reduce cardiovascular risk in chronic inflammation. Curr Pharm Des 2011; 17: 1-8.
45. Palmer S, Chen YH. Bcl-3, a multifaced modulator of NF-kB-mediated gene transcription. Immunol Res 2008; 42: 210-218.
46. Hinz M, et al. It takes two to tango: IκBs, the multifunctional partners of NF-κB. Immunol Rev 2012; 246: 59-76.
47. Yang J, et al. BCL3 interacts cooperatively with peroxisome proliferator-activated receptor gamma (PPARgamma) coactivator 1alpha to coactivate nuclear receptors estrogen-related receptor alpha and PPARalpha. Mol Cell Biol 2009; 29: 4091-4102.
48. Alberti KG, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. Circulation 2009; 120: 1640-1645.
49. Grundy SM. Metabolic syndrome: a multiplex cardiovascular risk factor. J Clin Endocrinol Metab 2007; 92: 399-404.
50. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). J Am Med Assoc 2001; 285: 2486–2497.